UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Shaping the organization of classrooms: Cross-national studies of policy, institutions, and pedagogical complexity

Permalink
https://escholarship.org/uc/item/0q48692t

Author
Kim, Yoonjeon

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Shaping the organization of classrooms: Cross-national studies of policy, institutions, and pedagogical complexity

By

Yoonjeon Kim

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Education

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bruce C. Fuller, Chair
Professor Sophia Rabe-Hesketh
Professor Trond K. Petersen

Fall 2015

© 2015
Yoonjeon Kim
All Rights Reserved

ABSTRACT

Shaping the organization of classrooms: Cross-national studies of policy, institutions, and pedagogical complexity

by

Yoonjeon Kim

Doctor of Philosophy in Education

University of California, Berkeley

Professor Bruce C. Fuller, Chair

With globalization, models of educational systems, school organizations, and teaching practices are increasingly being compared internationally. Heightened global competition and visibility of nations' student achievement by international assessment has pushed nation-states to take the lead to enhance student achievement and improve formal schooling. Complex problem-solving skills and critical thinking are emphasized as necessary for the work, life, and citizenship in the 21st century. Ideas on how to equip the students with such skills in schools and classrooms travel around the world. Within this context, the present three-paper thesis uses an international comparative framework to study how classroom teaching and learning is interrelated with different macro-level factors, such as global institutional pressures, regional cultural factors, and national educational policies.

In each of the three papers, I attend to different macro-level factors and examine how they impact social organization of classrooms, especially complex and procedural tasks, and student- and teacher-centered structure, and student achievement. In the first paper, I take an ongoing theoretical debate on global convergence/divergence and use them as a lens to examine the change of classroom practices over time. It tests the explanatory power of macro-level factors such as global institutional pressures and nation-specific factors for classroom practice and finds that convergence pattern vary by classroom practice dimensions. The second paper attends to East Asian countries, a world region that is renowned for their high achievement, but also notorious for their didactic instructional practice. The paper takes a step back and examines whether and how countries within the East Asian region are different from and similar to other countries in the world in terms of their classroom practices and student achievement. In it, I show that association patterns of instructional practices and student achievement at the between-country and within-country levels are distinct, reflecting different causal mechanisms that warrant further investigation at each level. The third paper focuses on yet another macro-level factor, standards-based educational policies. I take apart three key components of standards-based educational policy and examine how each of the elements affects complex teaching and student achievement. The findings show that each of the policy tools has distinct direct and moderating effects on classroom practice and student achievement.

Data and methods

The Trends in International Mathematics and Science Study (TIMSS) data is used in all three of the papers. For research on classroom instructional practice, an essential advantage of TIMSS over other international surveys (such as PISA) is that it includes measures of classroom practices comparable across countries and years. In order to address the research questions, I compile the rich micro-level data on students, classrooms, and schools in TIMSS with macro-level information drawn from various sources. Information on nations' educational policy was derived from the TIMSS Mathematics Curriculum Questionnaire. Other country-level information such as economic development, linkage to world society, and regional location are derived from the Yearbook of International Organizations (Union of International Associations 2000), and World Development Indicators (World Bank 1999, 2007).

Paper 1

The first paper contrasts two theoretical arguments on globalization of educational phenomena and examines how classroom instructional practices are explained by the contrasting frameworks. World Culture scholars argue that global institutional pressures shape educational structures around the world. These scholars have found mass schooling, curricular structure, and textbooks converge around a global model in countries with tighter link to the world society. Another line of scholars argues that nation-specific factors are more powerful in shaping educational practice, and thus national and local variation persists. But little is known about how classroom instructional practice fits into the trend and which of the macro-level factors—globalizing institutional forces or nation-specific factors—are more influential in shaping instructional practices inside classrooms. This paper examines changes in the extent to which classrooms demonstrate complex instruction or student-centered pedagogies across twenty-three nations, from 1999 to 2007, drawing on the TIMSS data. It shows that over time countries converged towards more student-centered and less didactic form of pedagogy and classroom organization. However, factors that explain convergence were not limited to the countries' linkage to world society, as argued by world culture scholars. Distinct regional patterns persisted. Eastern European and Middle Eastern countries were most active in moving toward global models of instructional practice, while Asian countries were least susceptible to the global shift toward a more student-centered model. And countries with stronger central control of the curriculum tend to have smaller within-country variation.

Paper 2

The second paper builds on one of the findings of the first paper. It focuses on one world region, East Asia, and further unpacks the association between classroom instructional practice and student achievement. East Asian countries have received much attention from other countries for their high performance in international assessments. Thoughtful small scale qualitative case studies with selective samples of schools and classrooms have introduced and revealed classroom practices and the social and cultural nuances of classroom organization in East Asian countries. These study findings are sometimes exaggerated and sometimes distorted to feed the discourses of domestic educational reforms. This paper examines these circulating beliefs about East Asian classrooms: teacher-centered and procedural instruction (the so-called East Asian educational paradox) and within-country homogeneity of practices, including equitable distribution of practices. TIMSS 2007 data is used to assess these beliefs. The findings show that while East Asian classrooms tend to be higher in teacher-centered instruction and lower in complex instruction on average, classrooms within each country vary considerably in the degree to which they adopt certain practices. Within

each country, classrooms with more complex and student-centered instruction tend to show higher achievement; an opposite association was found when these phenomena are compared between countries. But when schools, classrooms, and student characteristics such as social class are taken into account, these positive effects of classroom instruction diminish or disappear.

These findings suggest a need to rethink the assumption of homogeneity of schooling within East Asian nations and to be cautious in generalizing about East Asian classrooms based on national average features. Often, classroom practices prevalent in East Asian countries are celebrated as best practices to emulate, criticized for their rigidity, or simply regarded as non-importable due to the peculiarity of East Asian culture. But as this paper shows, East Asian countries face educational challenges similar to those observed in the U.S. and elsewhere. Once we acknowledge the commonality, a variety of research questions arise, which may guide future research.

Paper 3

The third paper attends to another macro-level factor, national educational policy, and extends one of the results in the first paper. One major international movement in education policy is the push toward increasing standards-based reforms, especially ones that emphasize complex instruction, such as Common Core State Standards (CCSS) in the U.S. These national endeavors provoke some positive and negative reaction from historically decentralized nations. But little is known whether the national endeavors can increase complex teaching and learning in classrooms, a stated goal of many standards-based reforms. Standards-based educational policies tend to differ in their use of three levers – national curricular standards, prescribed instructional guides, and test-based controls. The third paper takes apart these three key components of standards-based educational policy and examines how each of the elements affects complex teaching and student achievement. I use data from 48 nations and 9,154 classrooms that participated in the 2007 TIMSS. Using national variation of the three policy components, and classroom practices and achievement, I show that nations that pressed uniform instructional guides tended to host classrooms with less complex instruction, and those with greater emphasis on national testing manifest lower achievement. Most importantly, the positive effects of complex instruction on student achievement were stronger when national curriculum standards and more specified instructional guides were in place.

The findings of this study have broader implication for our understanding of theory of action of standards-based policies. It suggests that a broad caricature of standards-based policy—either as strong or weak—may be misleading; in fact each component of the policy differentially affects the level and the effectiveness of complex instruction. The findings have several implications for the U.S. debate on CCSS. First, they temper the willingness to use existing cross-state studies in the U.S. as direct evidence of the potential impact of a national scale policy at play. Second, they provide insights into how different arrangements and doses of the policy components—national curricular standards, instructional guides, and test-based control—may cause different responses from classrooms around the world, including the U.S. Third, the results encourage research on the link between components of standards-based policies and social organization of classrooms.

Broader implications

How can we encourage more complex and student-centered instructions over procedural and didactic instructions? How can educational policy systematically enhance the desired practices within classrooms? The three thesis papers shed light on perennial questions in education by taking a cross-national comparative framework. The first paper shows that some dimensions of classroom practice

rooted in the societies' cultural or social conventions (e.g., student- and teacher-centered instructions) tend to be immune to interventions, while other dimensions are more malleable. In the third paper, the impact of standards-based accountability policy is examined through three distinct policy components and their underlying theory of actions. By moving toward a more theoretically based view of policy and classroom practice, as done in the papers, we will be able to develop an improved theory about how policy levers and implementation work and obtain a deeper understanding of the dimensions and quality of policy and practice that are linked to improved student learning.

The international and comparative approach helps disentangle the different components of policy and practice. Moreover, studying national educational policy making inherently requires an awareness of education policy abroad as the world is increasingly globalized. The increasing standardization of educational curriculum and growing emphasis on complex and student-centered instruction are just two examples of educational phenomena becoming more common in the course of globalization. Nations are also faced with similar educational challenges, including unequal distribution of practices and achievement as shown in the second paper.

Beyond the specific issues it addresses, I hope that this dissertation will help convince scholars of education of the relevance and importance of comparative and international research in education. The comparative perspective is essential to refine and extend existing theories of education and policy; a necessary step if we are to devise educational policies that can make a meaningful difference on the ground.

Table of Contents

Table of Contents ..... **i**

List of Figures ..... **iii**

List of Tables ..... **iv**

Acknowledgements ..... **v**

Chapter 1 Global convergence or national variation? Examining national patterns of classroom instructional practices **1**

1.1 Introduction 2

1.2 Literature review 3

1.3 Propositions and analytic strategy 8

1.4 Methods 9

1.5 Results 14

1.6 Discussion 27

Appendix A1 30

Chapter 2 An alternative look at classroom instructional practices in East Asian countries: Variations and effects within country **34**

2.1 Introduction 35

2.2 Literature review 36

2.3 Analytic strategy 40

2.4 Methods 40

2.5 Results 44

2.6 Discussion 50

Appendix A2 53

Chapter 3 Can standards-based policy tools raise complex instruction and student achievement? A cross-national examination 58

3.1 Introduction 59

3.2 Literature review 60

3.3 Research questions and analytic strategy 62

3.4 Methods 63

3.5 Results 66

3.6 Discussion 72

Appendix A3 75

References 76

List of Figures

Figure 1.1 The relationship between country mean mathematics scores and within-country standard deviations for computational and conceptual instructions 16

Figure 1.2 The relationship between country mean mathematics scores and within-country standard deviations for teacher-centered and student-centered instructions 17

Figure 2.1 The relationship between country average and coefficient of variation of instructional practices 45

Figure 3.1 Moderating effects of the three policy tools on complex classroom instruction 60

Figure 3.2 Country average complex instruction and national curriculum 69

Figure 3.3 Country average complex instruction and degree of specification of instructional guides 69

Figure 3.4 Country average complex instruction and degree of test-based control.....69

Figure 3.5 The average slopes of complex instruction among countries with different levels of policy tools 71

List of Tables

Table 1.1 Descriptive statistics for all the variables used in the analyses 12

Table 1.2 Changes in levels and variances of instructional practices 18

Table 1.3 Effects of world linkage, national factors and world regional locations on computational and conceptual instructions 20

Table 1.4 Effects of world linkage, national factors and world regional locations on teacher- and student-centered instructions 22

Table 1.5 Effects of world linkage, national factors and world regional locations on within-country heterogeneity of computational and conceptual instructions 25

Table 1.6 Effects of world linkage, national factors and world regional locations on within-country heterogeneity of teacher- and student-centered instructions 26

Table 2.1 Descriptive statistics 41

Table 2.2 Relationship between instructional practices and student achievement in Japan 47

Table 2.3 Relationship between instructional practices and student achievement in Singapore 48

Table 3.1 Policy tools, complex instruction, and mathematics score across 48 countries 67

Table 3.2 Correlation matrix for country-level variables 68

Table 3.3 Effects of educational policy tools on the slope of complex instruction 70

Table 3.4 Variances of intercept and complex instruction slope estimated from the multilevel models 72

Acknowledgements

I would never have completed this dissertation without the support of many people. First, I am deeply grateful for the guidance and commitment of my advisor, Professor Bruce Fuller. He has patiently provided the vision, encouragement, and advice necessary for me to navigate through the most challenging task of my life. I thank Professor Sophia Rabe-Hesketh for her instruction and statistical advice on the dissertation analysis that helped shape and guide its direction. I also thank Professor Trond Petersen for his critical feedback and for providing outside perspective.

I would like to thank Professor Erin Murphy-Graham for her encouragement and appreciation of international comparative education research. I am grateful to the late Professor Norton Grubb and Professors Mark Wilson and John Lie for serving as important mentors earlier in my graduate school career.

I would also like to thank the current and former members of the Policy Implementation Research Group (PIRG) as well as its initial faculty sponsor, Professor Cynthia Coburn, who have reviewed numerous versions and antecedents of this dissertation. The research group was an invaluable part of my graduate school education.

I thank my parents, Yong Duk Kim and Hee Jun Kim, whose unconditional love, encouragement, and care touch every aspect of my life. I owe my sister, Minkyoung, and brother, Bumjoon, so much for their good wishes. And finally, I would like to thank my loving husband, Seung Han Kim, whose incredible support, patience, and encouragement throughout my graduate studies helped herd this project to completion.

Chapter 1 Global convergence or national variation? Examining national patterns of classroom instructional practices

Abstract

Little is known about the extent to which global institutional pressures and nation-specific factors shape instructional practices inside classrooms. The *world culture* framework, stemming from neo-institutional theory, posits a convergence in the organization of education across the world. But scholars observing teaching practices and life inside classrooms find significant variation among nations and regions of the world. This study reports changes in the extent to which classrooms demonstrate complex instruction or student-centered pedagogies across twenty-three nations, from 1999 to 2007, drawing on teacher reports from the Trends in Mathematics and Science Study (TIMSS). Means and within-country variances (heterogeneity) are then estimated among nations, predicted from standard measures of *world society linkage* and *country-specific factors* related to wealth, centralization, and culture, after controlling for teacher attributes. Despite evidence of convergence in classroom organization worldwide, distinct regional differences persist and nation-specific factors continue to exert significant explanatory power. Some advanced industrial nations show a reversion to teacher-centered instruction. Regional distinctions persist in East Asia, favoring teacher-centered classrooms.

1.1 Introduction

As global pressures shape the social organization of school systems, are instructional practices inside classrooms also converging toward an isomorphic model? *World culture* scholars argue that not only has the administrative structure of mass schooling spread around the world, but other domains of schooling are becoming more similar over time.[1] However, the evidence is mixed as to whether classrooms across countries are converging toward a Westernized model. World culture scholars explain that there are universal elements in most educational systems today that lead to convergence in education policies, formal and intended curricula, subject matter, school organization, and curriculum content (Ramirez 2012). In contrast, other researchers argue that structural, historical, and sociocultural factors at the regional and national levels are more powerful in shaping classroom practices.[2] According to this view, national patterns of teaching will maintain their own trajectory rather than converge toward a common model.

According to the world culture perspective, we would expect convergence toward more learner-centered and complex classroom instruction in countries with closer linkage to world society. By contrast, if the nation-state perspective has greater validity, we would expect to find distinct regional and nation-level differences in instructional practices that are unresponsive to or buffer globalizing forces. This paper enters the debate over the prevalence of global and national models in explaining instructional practices, asking the following questions: Have instructional practices across countries converged toward a global model in recent decades? Are the shifts in national patterns of instructional practices associated with the countries' embeddedness in world society or other country-specific factors? To answer these questions, I first investigate the changes in the levels of and within- and between-country variations in instructional practices over time. Second, I assess the extent to which the changes in classroom practices are explained by world culture and regional/national factors.

This study draws on three waves of the Trends in Mathematics and Science Study (TIMSS) data to assess levels of contemporary convergence of instructional practices. A cross-sectional analysis is not sufficient for testing a global convergence, since isomorphic trends are by definition processes that unfold over time. Moreover, this study traces the within-country variances over time in order to explore whether convergence occurs among teachers within countries. This provides a general description of how local actors respond to intervention at the global and national levels. Within-country variation has been overlooked by previous studies. In studies that took the world culture perspective, a strong assumption of homogeneity or convergence led to little attention being paid to the heterogeneity within nation-states. In studies that took the nation-state perspective, study designs that included small numbers of classrooms that sometimes only involved "highly effective" cases did not allow for gauging the heterogeneity within countries.

I begin by reviewing the two theoretical perspectives on what forces shape change in or reinforce static instructional practices across nations. This leads to the empirical research questions. The data set and analytic model are described, followed by the empirical findings.

---

[1] See, e.g., Boli, Ramirez, and Meyer (1985); Meyer et al. (1992); Baker and LeTendre (2005); Benavot and Braslavsky (2007); Meyer, Bromley, and Ramirez (2010); Ramirez (2012).

[2] See, e.g., Anderson, Ryan, and Shapiro (1989); Rohlen and LeTendre (1996); Cummings (1999); Stigler and Hiebert (1999); Anderson-Levitt (2002); Givvin et al. (2005), Clarke and Xu (2008), Tobin, Hsueh, and Karasawa (2009), Straubhaar (2014).

1.2 Literature review

World culture perspective: Global convergence in instructional practices

The world culture[3] perspective arose from an empirical puzzle: the rise and expansion of mass schooling irrespective of a nation's particular economic, political, or societal context. The worldwide growth of primary schooling after World War II could not be explained by either functional requirements of the society as a whole or stratification of society, which were common explanatory variables of educational phenomena in functionalist and class reproduction perspectives (Meyer et al. 1977; Boli, Ramirez, and Meyer 1985; Meyer et al. 1992). In order to explain the commonality across countries, scholars of neo-institutional orientation posit an international society in which nation-states are embedded. According to this view, nation-states expanded schooling to enact a proper nation-state identity and to be regarded as legitimate in world society. This search for legitimacy, they argue, has led countries to seek a common global model of mass schooling regardless of the country's own level of economic development, political regime, or local culture (Meyer et al. 1992; Ramirez 2012, 2003).

Earlier studies on the expansion of mass education were followed by studies that examined what went on within education systems, primarily by examining curricula and textbooks, moving a step closer to the core of schooling. Cross-national longitudinal studies of intended and formal school curricula reported increasing worldwide homogenization. Global standardization of curricula was observed in educational aims, curricular structure, emphasis on certain subjects, and time allocation (Benavot et al. 1991; Meyer et al. 1992; Frank et al. 2000; Cha 2007; Fiala 2007).

The direction of convergence suggested by world culture scholars generally reflects Western-centric ideals of individuals and schooling (Ramirez and Boli 1987; Meyer and Ramirez 2000; Baker 2009). For example, spread of human rights education (Suarez 2006), environmentalism (Frank, Hironaka, and Schofer 2000), and elimination of discrimination against women (Wiseman 2008) were documented by world culture scholars. These global trends are reported to have influenced growing emphasis of student-centeredness, democracy, citizenship, human rights and environmentalism in school curricula and textbooks (Benavot and Braslavsky 2007; Meyer, Bromley, and Ramirez 2010; Bromley, Meyer, and Ramirez 2011a, b).

Although world culture scholarship documents the convergence of codified elements of instruction such as textbooks and national curricula, little attention has been paid to the classrooms' social organization or the micro-structure of instructional practices. One exception is Gerald K. LeTendre et al.s' (2001) study, which investigated whether the effects of a global model or a national model predominate in various dimensions of teachers' work in Germany, Japan, and the United States. By means of analyses of variance with cross-sectional data, they report more similarity in teachers' instructional practices and beliefs and more difference in non-instructional work cross-nationally. Based on this finding, they argue that instructional practices are institutionalized at a global level. However, David L. Stevenson and David P. Baker (1991), in their study of the effect of state control of curricula on mathematics classroom instruction in fifteen countries, demonstrated that cross-

[3] "World culture" refers to a "rationalized world institutional and cultural order" (Meyer et al. 1997, 151). Other authors refer this as "world society" (Ramirez 2003), "institutionalism" (LeTendre et al. 2001), "neo-institutionalism" (Levin 2001), or "world systems theory" (Cummings 1999)

national variations in the content of mathematics curricula and instruction do exist and that the degree of centralization of curriculum control explains the variation in the curricular material taught in classrooms among countries.[4] As can be seen, even in the few studies on classroom instructional practices by world culture scholars, there is no conclusive evidence on global convergence.

While not empirically tested by world culture scholars, in the discussion of cultural influence on educational practices within the larger comparative education community, there has been some agreement on the existence and spread of global models of classroom pedagogy (Anderson-Levitt 2003; Ramirez 2003).[5] Learner-centered pedagogy or progressive teaching orientation is one such global models (McEneaney and Meyer 2000; Vavrus and Bartlett 2012; Schweisfurth 2013). This global model of instructional practice emphasizes problems that require higher-order thinking and active participation of students, and deemphasizes routine computational problems and didactic lecturing. However, whether classroom practices have converged toward this global model has yet to be assessed with quantitative design pursued by world culture scholars.

World culture scholars argue that the organizational form of mass schooling continues to converge, but their empirical work often demonstrates otherwise, revealing considerable regional and national variability. For example, clear regional patterns were observed in student-centrism in textbooks, with Latin America, Sub-Saharan Africa, and Asia showing high levels of student- centrism compared to the West (Bromley, Meyer, and Ramirez 2011a). Potential regional effects were also reported on curriculum structures in secondary education (Kamens, Meyer, and Benavot 1996) and on emphasis on social science instruction (Wong 1991). Nevertheless, these studies rarely inquire about the presence of multiple models at regional or national levels.[6] Some scholars have called for further examination of other national institutional characteristics that may explain national differences in classroom practices (Stevenson and Baker 1991).

Global convergence is by definition a process and can only be captured by historical data. Therefore, a standard research design in world culture studies has been to track historical trends in the national-level outcome of interest, demonstrating a lack of association between such outcomes and nation-specific factors and the presence of common trend lines across nations.[7][8] However, research that involves classroom instructional practices has yet to incorporate the time perspective (Stevenson and Baker 1991; LeTendre et al. 2001). Moreover, there has been little attention paid to variation within countries. The question on what happens beyond the moment of isomorphism at the between-country level is often truncated by the invocation of "loose coupling."[9] Most studies mainly focus on between-country variation, or cite a larger within-country variation compared to between-country variation in cross-sectional data as evidence of global convergence (LeTendre et al. 2001). But with growing recognition that fields are often populated by multiple and competing models (Bartlett 2003; Straubhaar 2014), it is worth taking a more systemic approach for analyzing heterogeneity within

[4] Similar findings were reported in Schmidt and Prawat (2006).

[5] See Carney, Rappleye, and Silova (2012) for critical appraisal of the trajectory of the discussion.

[6] Cummings (1999) suggests that there are multiple institutions of education, each of which develops along a distinct trajectory. Institutions are formed as a result of factors such as historical origin, representative schools, definition of the ideal person, and learning theory.

[7] See Schneiberg and Clemens (2006) for details of the study design.

[8] See, e.g., Meyer et al. (1977), Benavot et al. (1991), Kamens, Meyer, and Benavot (1996), and Bromley, Meyer, and Ramirez (2011b).

[9] World culture scholars explain the weak connection between formal structures and actual work activities by invoking the concept of loose coupling (Meyer and Rowan 1977). They do recognize the possibility of loose coupling between the stated activities of a country and the local implementation, but rarely incorporate assessment of local heterogeneity in the research design.

countries. This paper extends prior work done by world culture scholars by placing the cross-national data within a historical trend and investigates how patterns of instructional practices change within and among countries.

The nation-state perspective: Regional and national variation in instructional practices

The world culture perspective in instructional practice can be contrasted with what might be termed the *nation-state perspective.*[10] While world culture studies were triggered by a phenomenon of worldwide similarity, the studies coming from a nation-state perspective focus on differences among nations and regions of the world. It should be noted that the studies categorized under the nation-state perspective are less uniform than those in the world culture tradition in terms of the disciplinary backgrounds of the scholars, methodologies, countries studied, and attention to education or classroom practices. However, a unifying theme in this literature is that the studies' findings or reasoning suggest patterns of classroom practices that are relatively similar within nations and different across nations, and they do not assume a homogenizing global model.

The relationship between national education policy and classroom instruction is a popular topic in educational policy studies. Some quantitative studies have attempted to link education policy at the state-level with classroom practices, using large-scale international data. These studies report that teachers in countries with centralized curriculum control tend to have smaller variation in classroom practice than those who teach in decentralized countries (Stevenson and Baker 1991; Westbury and Hsu 1996; Astiz, Wiseman, and Baker 2002; Schmidt and Prawat 2006). Nonetheless, these studies have not tested how the degree of control over the curriculum might be associated with social organization and emphasis on certain types of activities in classrooms.

In discussions of educational reform in the US, scholars report that educational policies with a top-down approach worked better when aiming for classrooms that emphasize basic skills and didactic instruction than when aiming for classrooms that focus on complex learning and student-centered pedagogy (Rowan 1990; Cohen and Barnes 1993). It is argued that a decentralized policy is more suitable when the goal is to promote complex classrooms and student-centered instruction (Cohen and Spillane 1992; Fuller 2008). However, there is a scarcity of studies that test these theories using data at the classroom level that are comparable across educational systems (Spillane 2007).

Nations' economic development is another common factor that may affect educational practice. Classic theories of economic development depict a division of labor in the world economy where nation-states occupy certain roles and different types of production (e.g., agriculture, manufacturing, or knowledge-based industries) based on the countries' position in the world (Wallerstein 1974). With globalization and the proliferation of multinational corporations, international outsourcing has become common, wherein more developed countries outsource lower-skilled activities to less developed countries (Irwin 1996; Feenstra 1998). From this perspective, one would imagine that the types of labor and skills favored in a country might depend on the economic development of the country. Studies show that education reforms in developed countries stress conceptual understanding, creativity, problem-solving, and interpersonal skills over memorizing facts or

[10] Note that LeTendre et al. (2001), Anderson-Levitt (2002a), and Givvin et al. (2005) contrast the world culture perspective with what they term a "national culture perspective." In this study, the term "nation-state perspective" is used to encompass the effects of the nation's education policy and economic development on classroom practices in addition to the national culture.

procedural learning (Hargreaves 2003; Sahlberg 2006).

On a different note, economically dependent countries may be more susceptible to educational and economic reforms promoted by international development organizations and teacher educators. Studies have shown that educational practices favored by international organizations—especially learner-centered pedagogy—affected educational reform in developing countries such as Botswana (Tabulawa 2003), China (Carney 2009), Guinea (Anderson-Levitt and Diallo 2003), and Tanzania (Vavrus and Bartlett 2012), at least at the level of policy.

Another body of cross-national classroom research finds unique national cultural values to be the main factor that shapes schooling and instructional practices.[11][12] For example, studies of Japanese education have reported academically motivated students, high expectations held by parents, child-centered practices, and an emphasis on problem-solving in classrooms (Tobin, Wu, and Davidson 1989; Stevenson and Stigler 1992; Stigler and Hiebert 1999). Researchers argued that classroom structures and instructional practices in Japan differ from those of other nations because assumptions about the nature of mathematics, the nature of learning, and the role of the teacher are different across cultures (Stigler and Hiebert 1999).

Similar cultural arguments are advanced by studies that contrast education and schooling in different regions of the world. Studies attempting to explain the so-called East Asian paradox, wherein repetitive practices and memorization (which are generally not favored in the West) seem to be effective in East Asian classrooms, argue that different ideas about memorization and learning processes in Asian and Western cultures explain the puzzle (Kember 1996; Marton, Dall'Alba, and Tse 1996; Dahlin and Watkins 2000). More recently, the Nordic region received much attention with the Finnish success in education reform. Finnish teachers were characterized as having strong control of learning activities, keeping professional distance from students and their families, and conducting whole-class lectures (Norris et al. 1996, Simola 2005, Andrews 2013). Studies explain that the effectiveness of such practices was made possible by the social capital and trust engendered by the benevolence of the Nordic welfare state as well as the high social status and quality of teachers (Simola 2005; Sahlberg 2010).

While studies of the relationships among education policy, economic development, and classroom practices are rare, studies of the relationship between national culture and classroom practices are abundant. In arguing for national patterns of teaching, these studies assume that instructional practice within countries is more consistent than teaching across countries, but the designs of the studies do not allow for this to be demonstrated. In most cases, studies are not based on nationally representative samples and/or do not deal with variability within countries. Without representativeness and attention to variability, it is hard to make a case that what was observed is typical of the country.[13] Moreover, usually only a handful of countries are compared, many of which include high-achieving countries and the United States; lower-achieving countries are rarely

[11] See, e.g., Anderson et al. (1989), Tobin et al. (1989), Stevenson and Stigler (1992), Stigler and Hiebert (1999), Alexander (2000), Anderson-Levitt (2002b), Hiebert et al. (2003), Andrews (2007), Clarke and Xu (2008), and Hayashi and Tobin (2014).

[12] This body of literature has often been contrasted with the world culture perspective. See, e.g., LeTendre et al. (2001), Anderson-Levitt (2002a), and Givvin et al. (2005).

[13] See Tobin et al. (1989, 2009) for an exceptional case that incorporates both between and within country variability using a video-cued multivocal ethnography.

included.[14] The small number and selectivity of countries may have led to overstating the variation in classroom practices across countries and only allowed for inconclusive speculation about associations between regional/national contexts and micro classroom practices. Including a larger number of countries with a broader spectrum of backgrounds may be one way to address this limitation (Buchmann 2011).

As in the case of the world culture perspective, studies in the national culture tradition rely on cross-sectional data and rarely consider how classroom practices may change over time. These studies leave us uninformed about whether and how the national patterns of instructional practices subsist, diverge, or converge toward a global model over time. There are, however, exceptions. Tobin, Hsueh, and Karasawa's (2009) study analyzing preschools in China, Japan, and the US at two points in time is an exemplary analysis that incorporates a time component into a cross-national ethnographic classroom study. A more recent study by Rolf Straubhaar (2014) that examines how historically dominant model of instruction resonate in current classroom practices is another study that incorporates the time perspective

Reflecting a lack of common framework on features of classroom instruction in the broader discipline of teaching and learning (Hiebert and Wearne 1993; Grossman and McDonald 2008), studies of both world culture and nation-state perspectives use different language and measures in describing features of classroom instruction.[15] Moreover, while ethnographic studies within the broader comparative education literature do recognized that certain features of classroom practice are more malleable or resistant to change than others,[16] differential convergence pattern in dimensions of instructional practice has not been examined in previous research.

This study chose to focus on two dimensions of teaching practice that are commonly differentiated in teaching and learning literature; task and student-teacher relationship (Bossert 1979; Doyle 1981; Hiebert and Wearne 1993; Stein, Grover, and Henningsen 1996).[17] Tasks or problems given to students may differ in cognitive demand and induce different kinds of learning (Doyle, 1988). For example, some mathematical tasks may have multiple solution strategies and require students' complex thinking and reasoning, while other tasks may require memorization or use of routine procedures. Student-teacher relationship depicts the degree of teacher control or participation of students in classroom activities. The global model of classroom instruction, learner-centered and complex instruction, can be gauged along the two dimensions. Classrooms that follow the global model of teaching may favor conceptual instruction over computational instruction, and emphasize student-centeredness over teacher-centeredness.[18] Specific operationalizations of the two

[14] See, e.g., Hiebert et al. (2003), Stigler et al. (1987), Givvin et al. (2005), and Stigler and Hiebert (1999).

[15] For example, Hiebert et al. (2003) use "lesson signature", Schmidt et al. (2002) use "characteristic pedagogical flow", and Stigler and Hiebert (1999) use "cultural script" to name a few. Most quantitative studies used individual survey items (Schaub and Baker 1991; LeTendre et al. 2001).

[16] Hayashi and Tobin (2014) introduce the concept of implicit/explicit pedagogical practice, and explain that implicit practices are more resilient and resistant in the face of pressure to change. Implicit pedagogical practice can be conceptually related to what Anderson-Levitt (2002) terms "professional knowledge rooted in national classroom culture."

[17] Different terms are used by different authors. Task is referred as "academic task" or "types of problems". What I term student-teacher relationship is referred as "discourse" or "authority structure."

[18] Discussion of tasks in mathematics classroom, especially with regards to conceptual tasks, has been central to reform in mathematics education in the US and other countries (Kilpatrick, Swafford, and Findell 2001; Desimone et al. 2005; Common Core State Standards Initiative 2010; Andrews 2013) and is established as a key objective of mathematical literacy in the Programme for International Student Assessment (PISA). Student-centeredness in education has also been a major theme in comparative education, hypothesized to be a global idea that spread across the world in conjunction with a view of children as active learners (Bromley, Meyer,

dimensions are described in the Methods section below.

1.3 Propositions and analytic strategy

Given these competing perspectives, the present study derives a list of propositions on the worldwide convergence pattern of classroom instructional practices and key explanatory forces of the change. The first two propositions state the worldwide trends of classroom instructional practices hypothesized by world culture and nation-state perspectives. Note that the phenomena are in direct opposition: convergence and non-convergence.

Proposition 1: Country average classroom practices converge to a global model of classroom instructional practice that is learner-centered and complex.

Proposition 2: Country average classroom practices do not converge to a global model of classroom instructional practice that is learner-centered and complex.

Next set of propositions consider the national characteristics that may predict the nation's level of learner-centered and complex classroom instruction. A country's connection to the wider world has been the key explanatory variable for world culture theorists (Proposition 3). I also consider the possible impact of nation-specific factors emphasized by the nation-state perspective (Propositions 4a, 4b, and 4c).

Proposition 3: Countries with greater linkage to world society have higher level of learner-centered and complex classroom instruction on average.

Proposition 4a: Countries with different degrees of centralization in education policy have different levels of learner-centered and complex classroom instruction on average.

Proposition 4b: Countries with different levels of wealth have different levels of learner-centered and complex classroom instruction on average.

Proposition 4c: Countries in different world regions have different levels of learner-centered and complex classroom instruction on average.

In addition, I develop propositions to be examined for within-country variations. As both world culture and nation-state perspectives are less assertive about the patterns of within-country variation of instructional practices, these propositions are exploratory in nature. But if the logic of isomorphism extends to the within-country level, I expect to find teachers report on their classroom practices to converge within countries.[19]

---

and Ramirez 2011a; Vavrus and Bartlett 2012).

[19] World culture scholars do recognize local variation using the concept of loose coupling. But it is at least hard to argue for convergence toward a global model of instruction when there is great and/or growing heterogeneity within country.

Proposition 5: Within-country variations of classroom practice decrease over time.

Proposition 6: Countries with greater linkage to world society have smaller within-country variation of classroom practice.

Based on the world culture perspective, classrooms around the world will move toward a global model of teaching, and the variations across and within countries will decrease over time. Also, linkage to world society will be predictive of such changes, net of other national variations. Note that, while world culture and nation-state perspectives emphasize the importance of different variables, they do not imply contradictory arguments. Thus, the presence of association between a national factor suggested in one perspective and instructional practice does not necessarily contradict the other perspective.

To assess these working hypotheses, I analyze three waves of TIMSS survey data, for 1999, 2003, and 2007. Combining cross-sectional data over time allows for an assessment of the changes in and convergence of instructional practices. Furthermore, by combining classroom-level data with country-level information, I can gauge the explanatory power of factors advocated by the world society perspective and nation-state perspective for explaining the changes over time.

1.4 Methods

Data

Assessing the working hypotheses requires detailed data about classroom instruction in educational systems that differ in world linkage and national factors over a period of time. The TIMSS, undertaken by the International Association for the Evaluation of Educational Achievement (IEA), provides this type of data. TIMSS was first carried out in 1995 and has been repeated every four years thereafter. In addition to nationally representative fourth- and eighth-grade students' mathematics and science assessment scores, TIMSS contains a wealth of information on schools, curricula, instruction, and the lives of teachers and students.

I use teacher survey data collected from teachers of nationally representative eighth grade mathematics classes. The original TIMSS 1999, 2003, and 2007 eighth-grade mathematics data include 7,495, 10,018, and 12,591 teachers from 38, 51, and 57 countries and benchmark participants, respectively. For the purpose of my analysis, I will only use 3,419, 3,628, and 4,418 teachers with complete data in 23 countries that participated in TIMSS in all three time periods.[20]

For research on classroom instructional practice, an essential advantage of TIMSS over other international surveys (such as PISA) is that it includes measures of classroom practices comparable

[20] The small number of countries may raise a concern about the likely unstable results of the study (e.g., whether the results would differ if those excluded countries could be included). To assess the potential impacts of excluding many countries, results from supplementary analyses that included countries that participated in the two adjacent years were compared with the results based on 23 countries. The results (not shown) based on 32 countries in 1999-2003 period and 36 countries in 2003-2007 period were very consistent with the result presented in this study.

across countries and years. The scarcity of large-scale cross-national research on classroom instructional practice is partly due to data limitations, since cross-national data with comparative measures of classroom practices have been rare. The availability of TIMSS data provides an excellent opportunity to examine longitudinal trends in national patterns of instructional practices.

In addition to TIMSS surveys, data from the TIMSS Mathematics Curriculum Questionnaire (Mullis et al. 2000), the Yearbook of International Organizations (Union of International Associations 2000), and World Development Indicators (World Bank 1999) are used to derive country-level variables.

Dependent variables

Classroom instructional practices are captured by two particularly robust dimensions: tasks and student-teacher relationship. The first dimension, task, is measured by computational and conceptual instructions, and the second dimension, student-teacher relationship, is measured by teacher- and student-centered instructions.

Computational and conceptual instruction

A question in the three waves of the TIMSS survey asked about the frequency of use of computation as a method of teaching mathematics: "How often do you ask students to practice computational skills?"[21] Response options ranged from never (1) to always (4). Four items focused on conceptual instruction, asking how often teachers ask the students to perform the following tasks:[22] (a) explain the reasoning behind an idea, (b) represent and analyze relationships using tables, charts, or graphs, (c) work on problems for which there is no immediately obvious method of solution, and (d) write equations to represent relationships. The response options were the same as those for computational instruction. The mean across all four items was used as a conceptual instruction variable. Variables were standardized to have a mean of zero and standard deviation of ten.

Teacher- and student-centered instruction

Another set of questions asked the teachers write in the percentage of time they spend on a list of activities, which should sum up to 100%. Three among the seven items were used to derive the teacher-centered instruction variable: (a) lecture-style presentation by teacher, (b) teacher-guided student practice, (c) re-teaching and clarification of content and/or procedures. The sum of the three items was used as a teacher-centered instruction variable. A question about the percentage of time spent on student independent practice was used for a student-centered instruction variable. Variables were standardized to have a mean of zero and standard deviation of ten.

Within-country variation

As a measure of within-country heterogeneity, the standard deviation of each of the instructional practices for each country in a given year was calculated. This generated three values of standard deviation from the three waves—1999, 2003, and 2007—for each country.

[21] The wording of the question was slightly different in different years. Despite slight modifications of questions throughout the three waves, the theoretical concepts captured by the questions essentially remained the same.

[22] Results from factor analyses showed that there is only one factor with an eigenvalue greater than 1.

Independent variables

Key country-level independent variables are derived from various sources, as mentioned above. These variables represent values at or closest to the baseline year, 1999.[23] For better comparison, all country-level independent variables are standardized to have zero mean and one standard deviation.

World linkage variables

Three standard measures of linkage to world society are employed: national memberships in international nongovernmental organizations, national memberships in international governmental organizations (Union of International Associations 2000), and trade as a percentage of GDP (World Bank 1999). These measures are used to assess propositions 3 and 6.

Centralization of curriculum policy

Central control over curriculum, the focus of proposition 4a, is measured by a composite variable based on factor analysis of eight items in the TIMSS 1999 curriculum survey that asked about the central control over curriculum policymaking and implementation.[24] While two components had eigenvalues greater than one, a scree plot displayed an obvious break between the first and second components. Thus, only the first component, which had an eigenvalue of 3.78, was used for the analyses. Cronbach's alpha for the overall measure was 0.77.

Economic development

A standard measure of national economic development, log GNP/capita, is included (World Bank 1999). This measure is used to explore proposition 4b.

World region

Countries' location in world regions is included, as is common in cross-national studies. Grouping countries by world region is useful for seeing overall patterns across countries that share similar historical, cultural, political, or religious traditions. However, it should be noted that there may be considerable variation within each region. In this study, countries are classified in the following world regions: eastern Europe, the Middle East and North Africa, Asia, and the West (which includes Western Europe, North America, and Australia).[25] These measures are used to assess proposition 4c.

---

[23] Note that country variables from 1999 are used to predict the levels and changes of instructional practices in both 1999-2003 and 2003-2007 periods. This approach can be justified under the assumption that country variables are stable from 1999 to 2007. Preliminary analyses show that country measures changed only minimally during 1999 and 2007 (correlations range from 0.97 to 0.99), and that models using 2003 country measures for 2003-2007 period provide similar results to those using 1999 country measures. I also re-ran all models using average of the country variables over the two periods involved in the models and find that the results largely replicate the central findings reported. Thus, I use 1999 country measures in all analyses for simplicity.

[24] The items asked (1) whether decisions about curricula are made at the national or local level; (2) whether instructional time for mathematics is specified in the intended curriculum; and whether methods of implementation and monitoring included (3) pre-service teacher education, (4) in-service teacher education, (5) mandated or recommended textbooks, (6) instructional or pedagogical guides, (7) ministry notes and directives, and (8) systems of school inspection or auditing.

[25] Eastern European countries are Bulgaria, Hungary, Lithuania, Romania, the Russian Federation, and Slovenia; Middle Eastern and North African countries are Iran, Israel, Jordan, Morocco, and Tunisia; Asian countries are Chinese Taipei, Hong Kong, Indonesia, Japan, Korea (South), Malaysia, and Singapore; Western countries are Australia, Cyprus, England, Italy, and the United States.

Control variables

The average socioeconomic status of students in the classroom is controlled, because teachers may use different classroom practices for students of different backgrounds. The number of books and possessions in students' homes are averaged to the classroom level and used as proxies for classroom average socioeconomic status. A measure of teachers' years of experience was included as a proxy for teachers' content knowledge and pedagogical content knowledge. Although the TIMSS question regarding teachers' educational levels and college majors could have also served as a proxy for teachers' knowledge, data on this question were not available for several countries (e.g., Hungary), while information on the "years of teaching experience" question was available for all countries. Class size is known to affect overall classroom environment and teachers' work (Rice 1999; Pong and Pallas 2001) and is thus included as a control variable as well. Table 1.1 provides the means, standard deviations, and ranges of all teacher- and country-level variables used in the analyses (See Table A1.1 and A1.2 for descriptive statistics by country).

Table 1.1 Descriptive statistics for all the variables used in the analyses

	Mean	SD	Minimum	Maximum
Teacher level(N=11,465)				
Instructional practices				
Computational Instruction	0.00	10.00	-20.62	10.10
Conceptual Instruction	0.00	10.00	-33.56	34.52
Teacher-centered Instruction	0.00	10.00	-39.34	32.71
Student-centered Instruction	0.00	10.00	-15.01	63.66
Control variables				
Books at home	1.98	0.73	0.00	4.00
Possession at home	3.41	0.62	0.11	4.00
Teacher experience(x 5)	3.25	2.09	0.00	10.00
Class size(log)	3.26	0.44	0.00	5.27
Country level(N=23)				
Linkage to world society				
IGO memberships	47.04	18.74	0.00	72.00
INGO memberships	1,419.87	789.58	80.00	3,467.00
Trade(% of GDP)	94.78	77.62	19.00	343.00
National characteristics				
Centralization index	0.00	1.95	-7.05	1.08
GNP/capita(log)	8.75	1.24	7.01	10.55
World regions				
Western Europe and North America	0.22	0.42	0	1
Eastern Europe	0.26	0.45	0	1
Middle East and Africa	0.22	0.42	0	1
Asia	0.30	0.47	0	1

Analyses

Two sets of analyses were conducted to test the propositions. The first set of analyses takes a multilevel modeling approach. The analyses test the working hypotheses from a world culture perspective; namely, an overall movement toward conceptual and student-centered instruction, decrease of between-country variation over time, and a positive effect of world linkage on such changes, net of other national factors. The modeling approach also allows for an investigation of change patterns in within-country variation of the instructional practices over time. The second set of analyses takes an Ordinary Least Squares (OLS) regression approach. The analyses examine how country-level factors explain the levels and changes in within-country variation.

Multilevel modeling

The data include cohorts of teachers in 23 countries from 1999, 2003, and 2007. The data structure is longitudinal with respect to countries but cross-sectional with respect to each separate cohort of teachers. I use a hierarchical linear model to acknowledge this data structure. The modeling approach is similar to that which Stephen W. Raudenbush (1989) used to model successive cohorts of students from the same schools. Teacher data are pooled for adjacent years, 1999 and 2003, and 2003 and 2007.

The longitudinal hierarchical linear model involves two sets of equations for the teacher and country levels. The teacher-level model is a regression of instructional practices on teacher-level control variables within each country and at each point in time:

$$(Instructional\ Practice)_{ijt} = \beta_{0jt} + \beta_1 X_{ijt1} + \cdots + \beta_k X_{ijtk} + R_{ijt}$$

for teacher i (i=1,…,n_j) in country j (j=1,…,23) at year t (t=1, 2). Since 23 countries are observed at two time points, the equation above specifies 46 separate regressions. The β_{0jt} represent the instructional practice for each country j at year t after adjusting for the covariates in the model. Coefficients β_1 to β_k lack j and t subscripts because the effects of teacher-level control variables are assumed to be constant across countries and across time. Within-country error terms for year 1 and year 2, R_{ij1} and R_{ij2}, are independent and allowed to have different variances in each year.

The country-level model regresses countries' mean instructional practice, β_{0jt}, on D_2 (a dummy variable for year 2), country-level variable P_j, and an interaction between D_2 and P_j:

$$\beta_{0jt} = \Phi_{00} + \Phi_{01} D_2 + \Phi_{02} P_j + \Phi_{03} D_2 P_j + V_{j1}(1-D_2) + V_{j2} D_2$$

D_2 is coded as 0 in year 1 and 1 in year 2. Hence, Φ_{00} represents the mean instructional practice across all countries at year 1, and Φ_{01} represents the change in instructional practice over the two time points for countries with country characteristics corresponding to the average among the 23 countries (i.e., country-level variables are centered around the grand mean). Φ_{02} represents the effect of country characteristics on instructional practice in year 1, and Φ_{03} represents the change in the

effect of country characteristics on instructional practice between year 1 and year 2.

V_{j1} is the random intercept for year 1 and V_{j2} is the random intercept for year 2, and V_{j1} and V_{j2} are correlated. The effects (V_{j1} V_{j2}) are assumed to have a bivariate normal distribution with zero means, variances (τ_{00}, τ_{11}), and covariance τ_{01}. The combined equations for year 1 and year 2 are as follows:

Year 1 ($D_2 = 0$):

$$(Instructional\ Practice)_{ij1} = \Phi_{00} + \beta_1 X_{ij11} + \cdots + \beta_k X_{ij1k} + \Phi_{00} P_j + V_{j1} + R_{ij1}$$

Year 2 ($D_2 = 1$):

$$(Instructional\ Practice)_{ij2} = \Phi_{00} + \Phi_{01} + \beta_1 X_{ij21} + \cdots + \beta_k X_{ij2k} + (\Phi_{02} + \Phi_{03}) P_j + V_{j2} + R_{ij2}$$

Likelihood ratio tests are employed to choose between the presented models and nested models imposing a common between- or within-country variance between the two adjacent years. This allows for testing whether the changes in variances at the between- and within-country levels are statistically significant. Prior work by world culture scholars has focused on comparing the magnitude of between- and within-country variation at a given time point, largely ignoring changes of the variations over time. Moreover, patterns in within-country variation were rarely investigated.

Modeling within-country variation

The second set of analyses takes an Ordinary Least Squares (OLS) regression approach and regresses the within-country standard deviations of instructional practices on country-level independent variables. Since each country has three values of standard deviation from the three waves, which may be highly correlated, I use cluster-robust standard errors (sandwich estimator).

1.5 Results

Are there converging patterns of instructional practice?

Do country's average instructional practices move towards more conceptual and student-centered instruction, and less computational and teacher-centered instruction? How do variations of the instructional practices between- and within-county change over time? Figures 1.1 and 1.2 provide initial answers, and Table 1.2 shows the statistical significance of changes. The graphs plot the within-country standard deviations against the country means of the instructional practices over the three time periods (see Tables A1.3 and A1.4). Each year's overall mean of country average instructional practices and within-country standard deviation are represented by dashed lines.

Convergence patterns differed by instructional practices. For computational and conceptual instruction, no clear pattern of changes in the overall mean across countries was detected (Figure 1.1). This supports proposition 2. Countries have diverged in terms of their average computational

instruction. And within-country variances have increased over time in both computational and conceptual instruction. Change patterns were more visible in teacher- and student-centered instruction and in support of propositions 1 and 5 (Figure 1.2). An overall decrease of country average teacher-centered instruction, decrease of between-country variation, and decrease of within-country variation can be observed from Figure 1.2. Also, the between- and within-country standard deviations of student-centered instruction have decreased over time.

Observations from the figures are substantiated by Table 1.2 that shows the statistical significance of changes. The two columns under each instructional practice represent the analyses of the adjacent years, 1999–2003 and 2003–2007. Tests for the fixed effects show whether average levels of instructional practices change significantly between the adjacent years. Tests for the random part of the model indicate whether the between- and within-country variances for adjacent years are significantly different. This tests whether instructional practices are converging or diverging among nation-states and/or within nation-states.[26] If classrooms around the globe were converging toward a dominant model of instruction or social organization, we would see a decline in these variances over the nine-year period. Prior work by world culture scholars has focused on the directionality of fixed effects[27] while largely ignoring the changes in both between- and within-country variances over time, as discussed in the previous sections.

[26] Likelihood ratio tests are employed between the presented models and nested models that impose a common between- or within-country variance between the two adjacent time points.

[27] See, e.g., Meyer et al. (2010), Bromley et al. (2011a), and Bromley et al. (2011b).

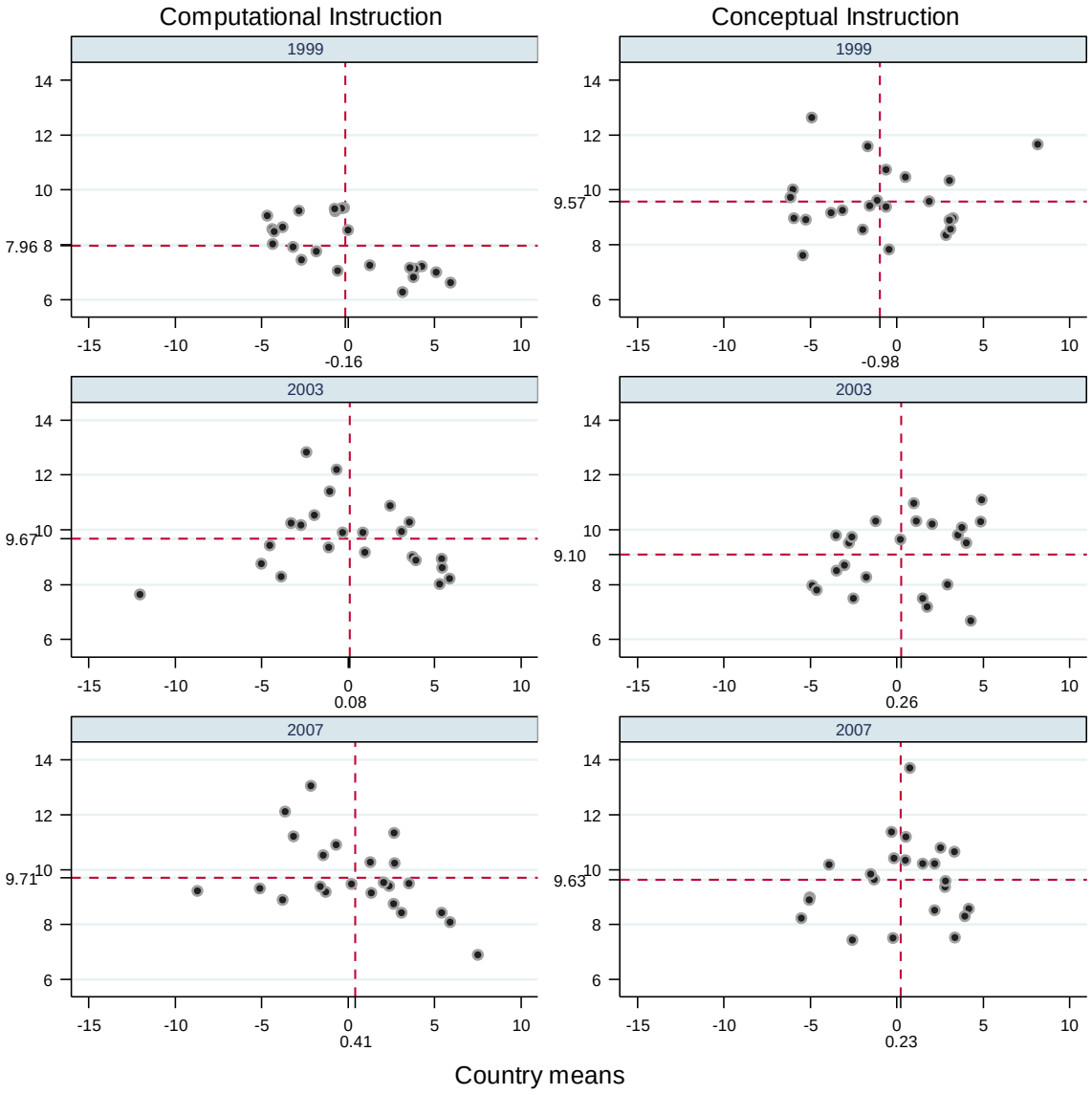


Figure 1.1 The relationship between country mean mathematics scores and within-country standard deviations for computational and conceptual instructions

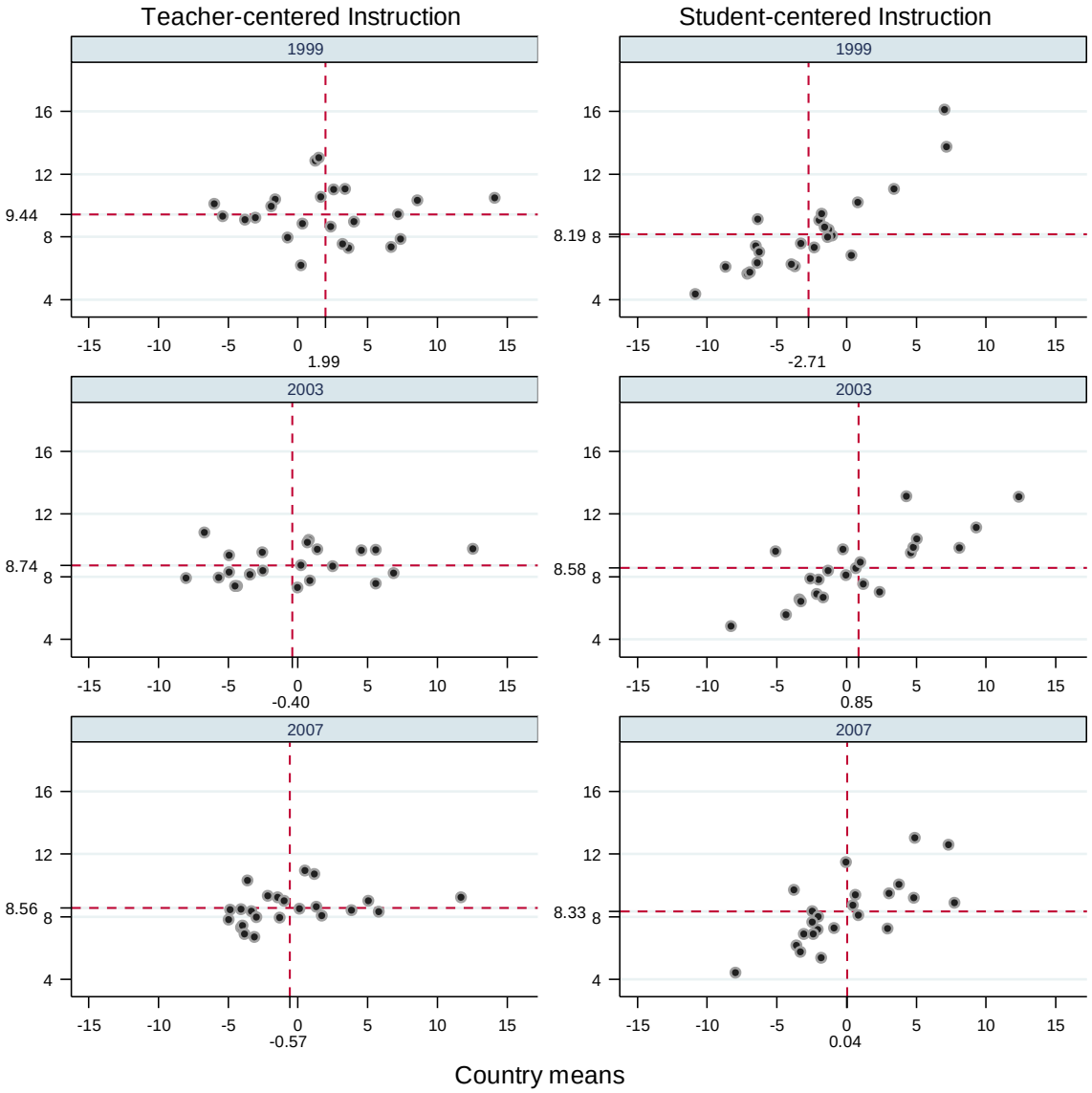


Figure 1.2 The relationship between country mean mathematics scores and within-country standard deviations for teacher-centered and student-centered instructions

Table 1.2 Changes in levels and variances of instructional practices

	Computation		Conceptual		Teacher-centered		Student-centered	
	1999-2003	**2003-2007**	**1999-2003**	**2003-2007**	**1999-2003**	**2003-2007**	**1999-2003**	**2003-2007**
Fixed								
Intercept	-0.16(0.71)	0.06(0.9)	-0.99(0.79)	0.35(0.67)	1.99(0.99)*	-0.39(1.04)	-2.71(0.95)**	0.9(1.03)
Year[a]	0.22(0.9)	0.38(0.34)	1.27(0.7)$^{+}$	-0.17(0.32)	-2.39(0.9)**	-0.2(0.53)	3.55(0.85)***	-0.84(0.36)*
Random								
Country-level								
SD_{Year1}[b]	3.33	4.25	3.72	3.11	4.68	4.94	4.51	4.89
SD_{Year2}[bc]	4.28	3.84	3.18	2.91	4.98	4.05^{+}	4.92	3.91**
$Corr_{Year1, Year2}$[b]	0.43	0.96	0.60	0.93	0.63	0.88	0.66	0.97
Teacher-level								
SD_{Year1}[b]	8.11	9.68	9.65	9.16	9.69	8.78	9.06	8.98
SD_{Year2}[bd]	9.68***	9.79	9.16**	9.53*	8.78***	8.51^{+}	8.98	8.62**
Intra-class correlation								
ICC_{Year1}[b]	14%	16%	13%	10%	19%	24%	20%	23%
ICC_{Year2}[b]	16%	13%	11%	9%	24%	18%	23%	17%

NOTE—a. Year indicates year 2003 for the models that pooled data for years 1999 and 2003, and indicates year 2007 for the models that pooled data for years 2003 and 2007.
b. Year1 and Year2 indicate the first and second years in the pooled data. For example, in the 1999-2003 models, Year1 and Year2 indicate years 1999 and 2003.
c. Significance indicates likelihood-ratio test comparing the presented models that allow between-country variances of the adjacent years to differ with nested models that impose a common between-country variance for the two adjacent years.
d. Significance indicates likelihood-ratio test comparing the presented models that allow within-country variances of the adjacent years to vary and nested models that impose a common within-country variance for the two adjacent years.
*+p<.1 *p<0.05 **p<0.01 ***p<0.001*

The convergence patterns of teacher- and student-centered instructions detected in Figure 1.2 were statistically significant. Teacher-centered instruction decreased and student-centered instruction increased during the period from 1999 to 2003. Likelihood ratio tests show that the decrease in between-country variances during the 2003–2007 period was statistically significant for both teacher- and student-centered instruction. Within-country variance also decreased throughout the nine-year period, which may signal a convergence within each country.[28]

As seen in Figure 1.1, no clear pattern of convergence was detected in computational and conceptual instruction, except for a marginal increase in conceptual instruction during the 1999–2003 period. As indicated by likelihood ratio tests, changes in between-country variation were not statistically significant for either computational or conceptual instruction. Furthermore, within-country variance of computational instruction increased significantly, especially during the 1999–2003 period, which indicates a divergence within each country. Changes in the within-country variances of conceptual instruction were less consistent, decreasing in 1999–2003 and then increasing in 2003–2007.

There was greater variation in the instructional practices of teachers within nation-states than in mean instructional practices across nation-states. This is indicated by the intra-class correlations at the bottom of Table 1.2. Intra-class correlations for teacher- and student-centered instruction were larger (ranging from 17% to 24%) than for computational and conceptual instructions (ranging from 9% to 16%). This indicates that not more than one-fourth of the variance in instructional practices and social organization of the classroom can be attributed to between-country factors. The bulk of variance remains attributable to factors that operate among schools and classrooms within nation-states. This leads to our next set of research questions, addressing which factors help to account for between- and within-country variability.

What factors explain the levels and changes of instructional practices?

Results from the multilevel model analyses are detailed in Tables 1.3 and 1.4. Blocks of predictors related to a nation's global linkage, nation-specific factors, and indicators of regional membership are entered into a null model that includes teacher-level control variables one block at a time. Only significant main effects and the corresponding interaction terms are presented for each model. Next, full models are reported that retain all significant variables and interactions terms. Initially, teachers' years of experience, class size, class mean of books in students' homes, and class mean of possessions in students' homes were included as control variables. The coefficients of teachers' years of experience and students' books at home were significant and were thus included in all estimation models.

[28] Analysis of pooled data of 1999 and 2007 shows that within-country variance of teacher- and student-centered instructions decreased significantly between 1999 and 2007.

Table 1.3 Effects of world linkage, national factors and world regional locations on computational and conceptual instructions

	Computational Instruction							Conceptual Instruction							
	1999-2003			2003-2007				1999-2003				2003-2007			
	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8	Model9	Model10	Model11	Model12	Model13	Model14	Model15
	National factors	World regions	Full	World linkage	National factors	World regions	Full	World linkage	National factors	World regions	Full	World linkage	National factors	World regions	Full
Teacher-level															
Intercept	-0.13	-1.41*	-1.14+	0.01	0.03	-1.35	-0.81	-1.06	-1.07	-1.84*	-2.01**	0.51	0.5	-0.12	0.14
Year[a]	0.15	0.02	0.3	0.35	0.38	0.17	0.13	1.49*	1.52*	1.66*	2.13**	-0.08	-0.01	-0.17	0.06
Teacher experience(x5)	-0.12*	-0.13*	-0.13*	-0.27***	-0.27***	-0.27***	-0.28***	-0.07	-0.07	-0.07	-0.08	-0.28***	-0.28***	-0.28***	-0.28***
Books at home	-0.74***	-0.86***	-0.79***	-0.38*	-0.34+	-0.40*	-0.36*	1.43***	1.44***	1.40***	1.41***	1.71***	1.75***	1.70***	1.73***
Country-level															
IGO memberships				1.5			0.66	1.57*			1.78*	1.58*			1.19+
IGOmemberships xYear				-0.13			-0.07	-0.07			-0.57	0.41			0.37
GNP/capita(log)	-1.84**		-1.12+		-2.79***		-1.97*		-0.13		1.16		-1.53*		-0.84
GNP/capita(log) x Year	-0.86		-0.96		0.1		0.25		-1.25+		-1.89*		-0.37		-0.38
Eastern Europe		4.90***	3.85**			5.23**	3.26+			2.99+	3.59*			2.4	1.36
Eastern Europe x Year		0.32	-0.59			0.54	0.76			-0.7	-2.32			0.28	-0.27
Variance explained at country level (%)[b]															
Year1[c]	27%	41%	47%	8%	38%	26%	44%	15%	-6%	10%	27%	20%	18%	6%	26%
Year2[c]	36%	26%	43%	8%	44%	41%	59%	20%	16%	6%	24%	39%	37%	13%	60%

NOTE—a. Year indicates year 2003 for the models that pooled data for years 1999 and 2003, and indicates year 2007 for the models that pooled data for years 2003 and 2007.
b. Percentage of the country-level variance explained by each model compared to a null model that includes teacher-level variables.
c. Year1 and Year2 indicate the first and second years in the models. For example, in the 1999-2003 models, Year1 and Year2 indicate years 1999 and 2003.
$^{+}$ *p<.1;* * *p<.05;* ** *p<.01;* *** *p<.001*

Table 1.3 presents results for computational and conceptual instruction. We see that a nation's stronger linkage to the global organizational networks was predictive of higher means for conceptual, but not computational, instruction. Membership in IGOs was significantly predictive of conceptual instruction during both the 1999–2003 and 2003–2007 periods (models 8 and 12). Note that 15% to 39% of between-country variance in conceptual instruction could be accounted for by countries' membership in IGOs. These results support proposition 3.

Associations between GNP per capita and computational and conceptual instruction were significantly negative, suggesting that at least middle-income nations tend to emphasize both computational and conceptual instruction more than wealthy countries. And the negative interaction between GNP per capita and Year, seen in model 9, suggests that wealthier countries may have deemphasized conceptual instruction amidst a worldwide increase in conceptual instruction in the 1999–2003 period, perhaps due to the rise of school accountability regimes. GNP per capita does explain a fairly large portion of between-country variance in computational instruction (27% to 44%) while accounting for a somewhat smaller portion for conceptual pedagogy (16% to 37%). The results support proposition 4b.

One consistent finding is that eastern European countries showed markedly different patterns of instruction and classroom organization compared with other regions of the world (proposition 4c).[29] Eastern European countries displayed significantly higher levels of computational and conceptual instruction relative to all other regions, including Asian nations. Note that 26% to 41% of between-country variance in computational instruction was explained by this regional location. This topic deserves further research in light of international interest in the ostensibly more rote or hierarchical classroom organization in east Asian nations. It appears that, to the contrary, eastern European educational systems place greater emphasis on computational and conceptual instruction in mathematics.

To examine the relative strength of predictors, models 3, 7, 11, and 15 include all significant independent variables (identified earlier), including world linkage, national factors, and regional location. GNP per capita and location in eastern Europe retain their significant influence on computational and conceptual instruction (models 3, 7, and 11). The world linkage factors continue to be significantly associated with higher levels of conceptual instruction, even after controlling for nation-level and social-class factors (models 11 and 15).

Table 1.4 presents results for teacher- and student-centered instruction. In the 1999–2003 period, the main effect of IGO memberships on student-centered instruction was negative, while the interaction between IGO membership and Year was positive. This indicates that countries with more IGO linkage employed less student-centered instruction in 1999, but they increased student-centered instruction at a much more rapid pace by 2003. Another indicator of a nation's linkage to world society, membership in NGOs, had a positive association with mean levels of student-centered instruction. No significant association between world-linkage factors and teacher-centered pedagogy was observed. These results support proposition 3.

[29] More fine-grained regional comparison of eastern Europe, the Middle East and North Africa, and Asia versus a reference category of western Europe, North America, and Australia was examined. The eastern European difference was most outstanding compared with other regions.

Table 1.4 Effects of world linkage, national factors and world regional locations on teacher- and student-centered instructions

	Teacher-centered						Student-centered						
	1999-2003			2003-2007			1999-2003				2003-2007		
	Model1 National factors	Model2 World regions	Model3 Full	Model4 National factors	Model5 World regions	Model6 Full	Model7 World linkage	Model8 National factors	Model9 World regions	Model10 Full	Model11 National factors	Model12 World regions	Model13 Full
Teacher-level													
Intercept	2.01*	1.01	1.51	-0.42	-2.26*	-1.9	-2.76**	-2.75**	0.49	-1.64	0.97	2.10+	1.70
Year[a]	-2.44**	-3.28**	-3.44**	-0.25	0.10	0.36	3.68***	3.67***	1.77	1.06	-0.77*	-0.70	-0.72
Teacher experience(x5)	0.10+	0.10+	0.10	0.12*	0.12*	0.12*	-0.11+	-0.11+	-0.11*	-0.11*	-0.03	-0.03	-0.03
Books at home	-0.32+	-0.31	-0.32	-0.46**	-0.44**	-0.45**	0.73***	0.72***	0.72***	0.71***	0.85***	0.85***	0.84***
Country-level													
IGO memberships							-2.38*			-2.18+			
IGO memberships x Year							1.73+			1.97			
INGO memberships							1.93*			1.37			
INGO memberships x Year							-1.12			-1.43			
Centralization	1.30		0.77	2.26*		0.69		-2.25**		-0.88	-2.10*		-1.33
Centralization x Year	0.99		-0.07	-0.36		0.29		0.35		-1.19	0.04		0.11
GNP/capita(log)	1.80+		1.35	1.98+		0.65					-0.62		0.04
GNP/capita(log) x Year	0.16		-0.74	0.50		1.04					-0.43		-0.37
Eastern Europe									-3.05	-0.67			
Eastern Europe x Year									4.77*	5.36+			
Middle East									-5.14+	-0.92			
Middle East x Year									2.43	1.90			
Asia		3.30	1.65		5.98**	4.87+			-4.36+	-2.45		-3.66+	-2.40
Asia x Year		2.73	3.28		-1.01	-2.00			0.40	2.62		-0.40	-0.20
Variance explained at country level (%)[b]													
Year1[c]	6%	8%	2%	15%	29%	23%	18%	24%	6%	16%	8%	10%	6%
Year2[c]	15%	29%	23%	29%	29%	32%	-6%	13%	13%	8%	18%	21%	18%

NOTE—a. Year indicates year 2003 for the models that pooled data for years 1999 and 2003, and indicates year 2007 for the models that pooled data for years 2003 and 2007.
b. Percentage of the country-level variance explained by each model compared to a null model that includes teacher-level variables.
c. Year1 and Year2 indicate the first and second years in the models. For example, in the 1999-2003 models, Year1 and Year2 indicate years 1999 and 2003.
+ p<.1; p<.05;** p<.01;*** p<.001*

Classrooms situated in nations with stronger centralized control of the curriculum tend to be more teacher-centered and less student-centered in their instructional organization (proposition 4a). Association between GNP per capita and teacher-centered instruction was significantly positive, suggesting that classrooms in wealthier countries tend to be more teacher-centered and less student-centered compared with those in developing countries (proposition 4b). Note that 6% to 29% of between-country variance in teacher-centered instruction and 8% to 24% of between-country variance in student-centered instruction could be accounted for by countries' degree of centralized control of curriculum and GNP per capita.

Asian nations did show distinct patterns: classrooms in the region remained quite teacher-centered, while teachers and classrooms in non-Asian classrooms were moving away from teacher-centered instruction throughout the study period. Other regional patterns also emerged from this analysis, as shown in model 9. Indicators of world regions show that classrooms in non-Western countries were less student-centered compared with classrooms in Western countries in the baseline year, 1999. While countries in all regions tended to move toward more student-centered instruction during the period of 1999–2003, the greatest increase in student-centered instruction was observed in eastern European countries, as indicated by the significant positive interaction term. Classrooms in Asian countries moved most slowly toward student-centered instruction, with mean levels remaining the lowest of any region in the world by 2007. These results provide strong support for proposition 4c.

Again, the significant predictors identified in the estimation models are presented to determine which operated in true independent fashion to account for variation in instructional practices (models 3, 6, 10, and 13). The most resilient factor in shaping teacher-centered instruction is the country indicator for Asian regional membership, after taking into account all other predictors. For student-centered instruction, IGO memberships and location in eastern Europe retained robust associations during the 1999–2003 period.

Coefficients for teacher-level control variables show that teachers with fewer years of experience organized higher levels of computational and conceptual instruction, compared with teachers reporting more years of experience. Teachers with longer tenures employed more teacher-centered and less student-centered instruction on average, compared with less experienced teachers. Classrooms populated by students from more affluent families experienced less computational and teacher-centered instruction and more conceptual and student-centered instruction.

What factors explain the levels of and changes in within-country variation in instructional practices?

Prior research in the world culture tradition emphasizes change in the organization of mass schooling among countries. But few studies have examined whether within-country variability is narrowing and whether nation-states conform to global expectations regarding the social organization of schooling, as discussed above. Tables 1.5 and 1.6 present results from multiple regression of within-country standard deviations in instructional organization on world linkage, national factors and world regional locations. The selection of predictors was handled as the multilevel models described above. That is, small blocks of country-level predictors and interaction terms were included in regression models, introduced one block at a time. The main effects and the corresponding interaction terms are then reported. Significant predictors are then brought forward into the full model.

Table 1.5 shows that the within-country variations increased over time for computational instruction, while within-country variation for conceptual instruction showed a less consistent pattern, significantly decreasing over the 1999–2003 period and then rising over the 2003–2007 period. Nations that have stronger linkage to the global economy showed smaller increases in within-country variation for computational instruction compared to nations that have less global economic exposure (proposition 6). Wealthier countries had more within-country heterogeneity in computational instruction. Eastern European nations displayed the lowest within-country heterogeneity for both computational and conceptual instruction. Middle Eastern and African countries displayed significantly larger within-country variation for conceptual instruction, perhaps as a result of shifts in dominant forms of instructional organization. The full models for computational instruction (models 4 and 8) reveal that within-country heterogeneity in computational instruction is explained by all three factors: linkage to world economic society, GNP per capita, and regional location for many nation-states. For conceptual instruction, the world-region indicators explain about 30% of the variation among nations in their within-country heterogeneity as indicated by the adjusted R^2 statistics (models 10 and 13).

A decrease in within-country variation was observed over time for teacher-centered instruction (Table 1.6). For countries with stronger links to the global economy, the decrease in within-country variability of teacher-centered instruction was smaller. We see that attributes of nation-states were significantly associated with the mean within-country standard deviation of student-centered instruction and explained notable portions of the standard deviation as indicated by the adjusted R^2 statistics in models 6 and 9. Countries with stronger centralization of curriculum controls display smaller and decreasing variances, while wealthier countries tend to display larger within-country variation in student-centered instruction. Eastern European nations again stand out, showing lower within-country variances for teacher- and student-centered instruction. However, these countries showed no significant change in their within-country variation in teacher-centered instruction—differing sharply from all other regions, which showed decreasing within-country standard deviations in teacher-centered instruction. For student-centered instruction, eastern European nations actually display rising levels of within-country standard deviation between 1999 and 2003. These patterns may be understood better in conjunction with the eastern European nations' prominent increase in levels of student-centered instruction, reported in the previous section. Asian countries showed significant decline in within-country variation in student-centered instruction over the 2003–2007 period.

Table 1.5 Effects of world linkage, national factors and world regional locations on within-country heterogeneity of computational and conceptual instructions

	Computational								Conceptual					
	1999-2003				2003-2007				1999-2003			2003-2007		
	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8	Model9	Model10	Model11	Model12	Model13	Model14
	World linkage	National factors	World regions	Full	World linkage	National factors	World regions	Full	World linkage	World regions	Full	World linkage	World regions	Full
Intercept	7.96***	7.96***	8.17***	8.08***	9.67***	9.67***	9.98***	9.91***	9.57***	9.69***	9.77***	9.1***	9.16***	9.25***
Year[a]	1.71***	1.71***	1.81***	1.83***	0.03	0.03	0.17	0.11	-0.47*	-0.52+	-0.52+	0.53*	0.44	0.51
IGO memberships									0.29		0.3	0.33		0.3
IGO memberships xYear									0.33		0.3	0.61*		0.58*
Trade(%ofGDP)	-0.02			-0.08	-0.63*			-0.68**						
Trade(%ofGDP)xYear	-0.63*			-0.68**	-0.44+			-0.53**						
GNP/capita(log)		0.45*		0.38		0.36		0.29						
GNP/capita(log)xYear		0.36		0.29		0.67*		0.53*						
Eastern Europe			-0.79+	-0.45			-1.18*	-0.92+		-1.17**	-1.31***		-1.18+	-1.32*
Eastern Europe xYear			-0.39	-0.47			-0.51	-0.3		-0.01	-0.01		0.09	-0.04
Middle East										0.85	0.64		1.11*	0.89+
Middle East xYear										0.26	0.25		0.3	0.1
N	46	46	46	46	46	46	46	46	46	46	46	46	46	46
Adjusted-R^2	42%	40%	42%	53%	11%	10%	19%	40%	3%	32%	34%	10%	32%	41%
Δ Adjusted-R^{2b}	7%	5%	7%	18%	13%	12%	21%	42%	2%	31%	33%	8%	30%	39%

NOTE—a. Year indicates year 2003 for the models that pooled data for years 1999 and 2003, and indicates year 2007 for the models that pooled data for years 2003 and 2007.
b. Δ Adjusted-R^2 is the change in Δ Adjusted-R^2 compared to a null model that includes Year only.
*+ p<.1; *p<.05; **p<.01; ***p<.001*

Table 1.6 Effects of world linkage, national factors and world regional locations on within-country heterogeneity of teacher- and student-centered instructions

	Teacher-centered					Student-centered					
	1999-2003			2003-2007		1999-2003			2003-2007		
	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8	Model9	Model10	Model11
	World linkage	World regions	Full	World linkage	Full	National factors	World regions	Full	National factors	World regions	Full
Intercept	9.44***	9.96***	9.96***	8.74***	8.74***	8.19***	8.53***	8.39***	8.58***	8.71***	8.36***
Year	-0.71*	-1.14**	-1.14**	-0.18	-0.18	0.39	-0.07	-0.12	-0.25	0.02	0.14
Trade(%ofGDP)	0.41		0.41	0.36*	0.36*						
Trade(%ofGDP) x Year	0.36*		0.36*	0.05	0.05						
Centralization						-1.18*		-1.23+	-1.04*		-1.2*
Centralizatio x Year						-0.83+		-0.75	-0.91*		-0.79
GNP/capita(log)						0.68*		0.52			
GNP/capita(log) x Year						0.52		0.78**			
Eastern Europe		-1.98***	-1.98***				-1.31	-0.77			
Eastern Europe x Year		1.66*	1.66*				1.79**	1.99*			
Asia										-0.42	0.74
Asia x Year										-0.89+	-1.28+
N	46	46	46	46	46	46	46	46	46	46	46
Adjusted-R^2	7%	20%	24%	0%	0%	24%	-3%	24%	15%	-2%	13%
Δ Adjusted-R^2	3%	16%	20%	2%	2%	26%	-1%	26%	17%	0%	15%

NOTE—a. Year indicates year 2003 for the models that pooled data for years 1999 and 2003, and indicates year 2007 for the models that pooled data for years 2003 and 2007.
b. Δ Adjusted-R^2 is the change in Δ Adjusted-R^2 compared to a null model that includes Year only.
*+ p<.1; * p<.05; ** p<.01; ***p<.001*

1.6 Discussion

In this study, I asked whether classroom instructional practices across the world converge toward global models of instruction and whether the countries' embeddedness in world society explains the convergence, as hypothesized by world culture scholars. Rather than taking a cross-sectional approach and focusing on educational organizations, this study examines the cross-national patterns of teacher instructional practices in a longitudinal scale. By doing so, this study assesses patterns at the supranational level—such as convergence—more accurately and unpacks what factors explain the changes as well as the levels of instructional organization. Moreover, the study contributes to our understanding of institutionalization patterns in instructional practices across countries by examining within-country variability, which has been overlooked in most previous research. My analysis reveals that instructional practices across countries moved toward a more student-centered and less didactic form of pedagogy and classroom organization. However, factors that explain convergence were not limited to the countries' linkage to world society, as argued by world culture scholars. Moreover, convergence patterns and explanatory factors differed according to the dimensions of instructional practices.

To be sure, convergence was prominent in certain dimensions of classroom instructional practice. Student-teacher relationship shifted toward increased emphasis on individual agency and child-centered curricula, as suggested by world culture scholars (Bromley et al. 2011). Moreover, emphasis on student-centeredness was closely related to a country's exposure to the wider world and hence to the world model of child-centered pedagogy. On the other hand, the overall decrease in teacher-centered instruction could not be explained by countries' linkage to world society, which suggests that this movement may be due to other factors. In addition, more centralized countries tend to be resistant to the overall shifts toward emphasis on student-centered instruction, maintaining lecture-oriented classrooms and teacher control over classroom activities. This may support David K. Cohen's (1990) hypothesis about the congruence between the pedagogy of policy and pedagogy in classrooms. Future cross-national case studies may build upon the present work by comparing the social organization of classrooms in countries with different degrees of curriculum control. More specifically, unpacking the relationship between the top-down nature of education policy and modes of classroom instruction could contribute to our knowledge of educational policy implementation and thereby provide useful guidance for educational policy and reform.

With respect to instructional tasks in classrooms, countries were similar in the degree to which teachers emphasized computational and conceptual tasks. Also, no significant changes in the mean levels or convergence patterns were detected. It is possible that current observations are a result of global convergence due to institutionalization at the global level. The significant association between the world-linkage factor and conceptual instruction after controlling for nation-specific characteristics further supports this conclusion. Analysis of classroom instruction data from a broader time span may provide additional evidence of this convergence.

Another finding that deserves more discussion is the presence of distinct patterns in different regions of the world. Eastern European and Middle Eastern countries were most active in moving toward global models of instructional practice, while Asian countries were resistant to global trends. Distinct world regional effects were also found in previous studies of curricula and textbooks

(Kamens, Meyer, and Benavot 1996; Bromley, Meyer, and Ramirez 2011a, b). World culture scholars often explain the uneven incorporation of global models across world regions as resulting from the relative strength or weakness of international communication and the various regions' educational heritage (Meyer and Ramirez 2000; Bromley, Meyer, and Ramirez 2011a). Previous studies on eastern European countries document a rapid change toward Western-centric educational policy models (Baker, Köhler, and Stock 2007) or a more complex transformation processes within eastern European countries (Silova 2009) after the fall of the Soviet system. The faster incorporation of conceptual and student-centered instruction in eastern European classrooms may be reflecting the political, economic, and social environment within these countries.

Consistent with the smaller case studies of Asian classrooms, Asian countries maintained teacher-centered pedagogy, being the least susceptible to the global shift toward a more student-centered model. This study shows that centralized control over curricula partly explains the distinctive patterns in Asian classrooms. It is possible that a tradition of authoritarian teachers and deeply rooted authority structures in schools explain the teachers' resistance to an overall shift toward child-centered instruction. The high-stakes exam system of Asian countries may have led the teachers to concentrate on mathematic problems that appear in the exams and favor lecture-oriented or teacher-centered mode of instruction. Alternatively, a close alignment of classroom practices, educational policies, educational beliefs of teachers, and general national cultural values may have fortified teacher-centered pedagogy against the global tide (Hayashi and Tobin 2014).

This study focused on within-country variability and thereby shed light on global convergence patterns of instructional practice. Institutionalization of instructional practices at the global level may not necessarily lead to the uniform adoption of such practices within countries. Convergence of teacher- and student-centered instruction occurred within countries, and centralized educational policies partly explained the decrease in within-country variation. However, teachers within countries diverged in their emphasis on computational instruction over time. Persisting within-country variation in teacher practices has been continuously highlighted by cross-national case studies that focus on the local enactment of world-level models (Anderson-Levitt 2003; Vavrus and Bartlett 2012). Acknowledging inconsistency and variation in teacher instructional practices within countries, along with differences between countries, may provide a more holistic depiction of the process of globalization. More attention to within-country heterogeneity is warranted in future studies.

Given the short period of time and the limited number of countries analyzed, the present study is exploratory, and its findings are only suggestive. To fully understand the worldwide convergence patterns of instructional practices would require data from a larger number of countries over a longer period of time. Still, the analytical approach of this paper can be modeled upon by future studies. The multilevel model used in this paper acknowledges a data structure that is longitudinal with respect to countries but cross-sectional with respect to each separate cohort of teachers. It also allows for testing the statistical significance of change in variation at both between- and within-country levels, which has been largely ignored in previous studies.

Teachers' own reports may not directly reflect how they actually teach in the classroom and may be subject to social desirability bias (Burstein et al. 1995). However, anonymous surveys such as TIMSS are known to be less susceptible to such bias than interviews or focus groups (Desimone and LeFloch 2004; Aquilino 1998). Moreover, even if social desirability bias is present, the survey at least

provides insight into what teachers think they should be doing in classrooms. That teachers across countries tend to perceive more student-centered instruction as desirable suggests that a global norm of student-centered instruction is influencing teachers' outlook. Future studies may attempt to parse the relationship between teachers' perception of their teaching and actual classroom practice.

Appendix A1

Table A1.1 Descriptive statistics of teacher-level variables by country

Country	Books at home		Possession at home		Teacher experience		Class size		# of Teachers
	mean	SD	mean	SD	mean	SD	mean	SD	
Australia	2.60	0.56	3.81	0.22	15.74	10.15	25.75	5.04	516
Bulgaria	2.33	0.84	3.15	0.60	21.44	9.19	22.15	5.64	478
Cyprus	2.00	0.38	3.69	0.22	11.64	9.49	25.24	3.84	377
England	2.29	0.82	3.75	0.28	13.88	10.65	23.54	8.58	508
Hong Kong	1.45	0.48	3.62	0.24	12.17	9.38	36.19	8.71	372
Hungary	2.61	0.69	3.64	0.34	21.11	9.63	20.78	6.96	508
Indonesia	1.08	0.36	2.49	0.46	12.93	7.59	48.88	33.35	281
Iran	1.18	0.63	2.16	0.77	14.10	8.73	28.10	8.32	440
Israel	2.33	0.68	3.82	0.29	16.05	9.63	29.76	9.08	565
Italy	2.06	0.56	3.64	0.23	22.18	8.79	21.03	4.23	626
Japan	1.99	0.39	3.70	0.19	15.23	8.78	33.92	7.08	431
Jordan	1.53	0.39	2.85	0.52	10.83	7.85	35.55	9.45	410
Korea	2.35	0.50	3.81	0.22	12.78	8.49	39.00	7.68	559
Lithuania	1.91	0.58	3.37	0.46	21.15	9.97	24.42	5.55	557
Malaysia	1.53	0.51	3.27	0.39	10.96	8.41	36.70	7.91	395
Morocco	1.07	0.46	2.29	0.57	16.99	6.53	38.75	25.16	301
Romania	1.75	0.72	2.73	0.79	23.70	11.31	22.39	6.13	495
Russian Federation	2.39	0.57	3.21	0.45	22.14	10.71	22.81	5.97	620
Singapore	1.84	0.56	3.78	0.22	10.77	11.60	37.11	4.85	708
Slovenia	1.96	0.56	3.74	0.25	18.38	9.43	17.37	6.56	629
Taiwan	1.87	0.50	3.71	0.24	13.92	9.67	36.66	6.23	401
Tunisia	1.27	0.47	2.73	0.54	12.30	8.72	32.91	4.62	237
United States	2.18	0.72	3.66	0.30	15.03	10.45	23.81	13.27	1,051

Table A1.2 Country-level variables by country

Country	IGO memberships	INGO memberships	Trade (% of GDP)	Centralization	GNP /capita(log)	World Region
Australia	51	2306	39	-3.63	9.94	West
Bulgaria	52	1372	117	0.52	7.06	Eastern Europe
Cyprus	48	831	102	0.33	9.25	Asia
England	0	80	54	-2.62	9.82	West
Hong Kong	11	1173	249	1.08	10.13	West
Hungary	55	2185	132	0.34	8.41	Asia
Indonesia	53	1084	63	0.89	7.01	Eastern Europe
Iran	44	647	36	1.08	7.48	Asia
Israel	41	1895	71	-0.22	9.69	Mid East and Africa
Italy	72	3467	47	-0.19	9.91	Mid East and Africa
Japan	61	2246	19	0.89	10.55	West
Jordan	54	566	104	0.89	7.33	Asia
Korea	51	1387	71	1.08	9.26	Mid East and Africa
Lithuania	36	915	87	-0.41	7.72	Asia
Malaysia	50	1131	218	1.08	8.42	Eastern Europe
Morocco	68	848	56	1.08	7.14	Asia
Romania	56	1483	61	1.08	7.25	Mid East and Africa
Russian Federation	63	1901	69	1.08	7.89	Eastern Europe
Singapore	33	1115	343	1.08	10.40	Eastern Europe
Slovenia	41	1307	99	0.37	9.19	Asia
Taiwan	11	1081	41	0.37	9.49	Eastern Europe
Tunisia	69	779	78	0.89	7.65	Mid East and Africa
United States	62	2858	24	-7.05	10.28	West

Table A1. 1 Country means of instructional practices

	Computational			Conceptual			Teacher-centered			Student-centered		
	1999	**2003**	**2007**	**1999**	**2003**	**2007**	**1999**	**2003**	**2007**	**1999**	**2003**	**2007**
Australia	-4.25	-4.53	-3.76	-5.23	-4.61	-5.05	-1.63	-6.69	-3.60	7.17	12.37	7.32
Bulgaria	-0.22	5.45	5.42	3.25	1.51	4.20	6.69	4.59	0.12	-2.31	0.69	0.49
Cyprus	-2.70	0.87	2.08	-0.60	-1.19	2.21	0.23	-5.68	-4.07	-7.10	-2.12	-3.29
England	-4.36	-1.11	-1.61	-5.42	-2.49	-2.55	1.51	0.82	1.18	7.06	4.33	4.90
Hong Kong	0.02	-12.03	-8.72	-5.93	-3.49	-5.50	2.57	5.59	3.85	-1.54	-0.25	-2.03
Hungary	3.89	3.95	1.35	2.86	4.29	3.37	2.38	-4.93	-4.86	-1.02	8.10	4.85
Indonesia	3.16	-3.85	-1.28	-3.12	-3.01	-0.15	-5.37	0.87	-4.06	-1.29	-2.57	-0.88
Iran	-0.74	3.11	2.69	-4.90	4.93	2.54	-1.91	-3.43	-3.97	-6.50	-1.99	-2.04
Israel	-2.84	-2.72	-3.14	-1.67	1.02	-0.30	-3.05	-4.50	-2.96	3.41	4.82	3.07
Italy	-0.76	-1.95	-0.67	3.07	4.03	3.35	3.20	-0.02	-1.30	-3.92	-3.39	-3.57
Japan	-1.81	-2.41	-2.14	8.16	1.16	0.54	14.13	12.55	11.72	-6.36	-5.07	-3.76
Jordan	5.12	3.74	3.53	-1.13	3.78	2.81	-6.00	-2.48	-3.09	-1.95	-3.26	-1.81
Korea	-4.67	-1.06	-3.62	1.88	3.54	2.20	8.58	2.51	5.07	-1.77	4.64	0.86
Lithuania	5.94	0.98	2.40	-1.96	1.77	-0.20	3.63	-4.93	-4.96	-1.32	9.30	7.75
Malaysia	4.27	5.41	3.11	0.53	0.24	0.53	1.28	-2.53	-2.15	-6.38	-0.03	-3.06
Morocco	-0.59	-0.66	0.22	-6.13	-2.57	0.78	4.02	1.40	-1.43	-10.84	-4.31	-2.46
Romania	3.78	5.89	7.51	3.05	4.87	2.84	0.34	5.60	1.72	-3.68	-1.65	-2.37
Russian Federation	3.59	5.31	5.91	-0.44	2.96	3.94	-0.74	-4.36	-3.81	0.38	2.41	2.97
Singapore	-4.37	-5.02	-5.09	-6.00	-4.88	-5.02	1.64	0.71	-0.97	-3.27	-1.32	-2.46
Slovenia	1.27	-0.30	2.65	3.11	-1.75	-1.29	7.36	0.25	1.35	-6.24	5.05	3.76
Taiwan	-3.17	2.46	2.67	-1.54	-3.47	-3.91	7.21	6.89	5.83	-6.92	-8.27	-7.96
Tunisia	-3.79	3.57	1.31	-3.79	-2.75	-1.48	3.40	-8.02	0.53	-8.66	1.20	-0.04
United States	-0.38	-3.28	-1.42	-0.59	2.05	1.53	-3.78	-3.44	-3.30	0.81	0.99	0.64
Total	**-0.16**	**0.08**	**0.41**	**-0.98**	**0.26**	**0.23**	**1.99**	**-0.40**	**-0.57**	**-2.71**	**0.85**	**0.04**
Between country variation	**14%**	**16%**	**13%**	**13%**	**11%**	**9%**	**19%**	**24%**	**18%**	**20%**	**25%**	**17%**

Table A1. 2 Within-country standard deviations of instructional practices

	Computational			Conceptual			Teacher-centered			Student-centered		
	1999	**2003**	**2007**	**1999**	**2003**	**2007**	**1999**	**2003**	**2007**	**1999**	**2003**	**2007**
Australia	8.46	9.43	8.90	8.90	7.79	8.89	10.37	10.82	10.30	13.73	13.09	12.58
Bulgaria	9.35	8.61	8.43	8.95	7.49	8.56	7.33	9.69	8.49	7.32	8.54	8.71
Cyprus	7.43	9.89	9.52	9.38	10.31	10.21	6.16	7.94	8.46	5.62	6.89	5.73
England	8.02	9.35	9.39	7.61	7.50	7.44	13.04	10.35	10.72	16.11	13.14	13.01
Hong Kong	8.52	7.63	9.21	8.97	9.78	8.23	11.02	9.70	8.40	8.61	9.75	7.99
Hungary	7.11	8.87	9.15	8.35	6.67	7.52	8.63	8.28	8.43	8.09	9.84	9.19
Indonesia	6.26	8.29	9.18	9.25	8.70	10.41	9.29	7.76	7.31	8.44	7.88	7.27
Iran	9.21	9.92	10.23	12.63	11.08	10.79	9.93	8.17	7.42	7.40	7.81	7.17
Israel	9.24	10.17	11.20	11.59	10.97	11.37	9.20	7.40	7.95	11.03	9.87	9.49
Italy	9.30	10.53	10.90	10.33	9.51	10.64	7.54	7.29	7.93	6.24	6.53	6.16
Japan	7.75	12.81	13.06	11.64	10.31	11.19	10.48	9.78	9.23	9.10	9.62	9.69
Jordan	6.99	9.01	9.49	9.60	10.08	9.36	10.10	8.37	6.70	9.06	6.41	5.37
Korea	9.04	11.39	12.11	9.57	9.80	8.52	10.30	8.65	9.00	9.47	9.52	8.09
Lithuania	6.61	9.17	9.39	8.53	7.19	7.50	7.29	9.36	7.81	7.97	11.14	8.88
Malaysia	7.21	8.93	8.43	10.47	9.65	10.34	12.85	9.54	9.33	6.32	8.08	6.88
Morocco	7.05	12.19	9.47	9.72	9.73	13.71	8.97	9.74	9.22	4.34	5.57	8.35
Romania	6.81	8.22	6.90	8.89	10.30	9.58	8.83	7.57	8.04	6.12	6.67	6.87
Russian Federation	7.16	8.02	8.08	7.83	8.00	8.30	7.94	7.40	6.89	6.79	7.01	7.22
Singapore	8.56	8.76	9.31	10.00	7.95	8.99	10.54	10.19	8.99	7.56	8.38	7.64
Slovenia	7.23	9.90	8.76	8.57	8.26	9.63	7.86	8.72	8.62	7.02	10.39	10.05
Taiwan	7.92	10.87	11.33	9.41	8.49	10.18	9.44	8.22	8.29	5.72	4.83	4.41
Tunisia	8.64	10.28	10.27	9.16	9.51	9.84	11.04	7.90	10.94	6.09	7.52	11.47
United_States	9.33	10.24	10.52	10.73	10.19	10.21	9.08	8.12	8.33	10.18	8.92	9.40
Total	**7.96**	**9.67**	**9.71**	**9.57**	**9.10**	**9.63**	**9.44**	**8.74**	**8.56**	**8.19**	**8.58**	**8.33**

Chapter 2 An alternative look at classroom instructional practices in East Asian countries: Variations and effects within country

Abstract

East Asian schools have received much attention from researchers and reform advocates for the comparatively high achievement of their students. Western scholars at times advance broad claims about the rote character of instruction or the complexity of classroom practice, at times generalizing to an entire nation. Yet little is known about the variation in and impact of classroom instructional practices within East Asian countries. This study first compares the prevalence and within-country variation of classroom instructional practices in Chinese Taipei, Hong Kong, Japan, Korea, and Singapore with the other 45 countries that participated in the Trends in Mathematics and Science Study (TIMSS) in 2007. Second, the study examines the effect of varying instructional practices on student achievement within each country using data from Japan and Singapore. I find that while East Asian classrooms tend to be higher in teacher-centered instruction and lower in complex instruction on average, classrooms within each country vary considerably in the degree to which they adopt certain practices. Within each country, classrooms with more complex and student-centered instruction tend to show higher achievement; an opposite association found when these phenomena are compared between countries. But when schools, classrooms, and student characteristics such as social class are taken into account, these positive effects of classroom instruction diminish or disappear. These findings suggest a need to rethink the assumption of homogeneity of schooling within East Asian nations and to be cautious in generalizing about East Asian classrooms based on national average features. The different association patterns of instructional practices and student achievement at the between-country and within-country levels indicate that different causal mechanisms are at play, which warrant further investigation in future research.

2.1 Introduction

The repeated success of East Asian students in international assessments has triggered much curiosity about their schools and the social organization of their classrooms (Stevenson and Stigler 1992; Stigler and Hiebert 1999; Hiebert et al. 2003). Generalities about allegedly nationwide teaching practices and instructional organization abound. Secondary classrooms in East Asia, for example, are often characterized as having teachers that use more didactic approaches, passive students, and an emphasis on procedural knowledge and whole-class instruction (Watkins and Biggs 1996; Huang and Leung 2004). These characteristics are in sharp contrast with what is believed to be conducive to learning in the West. In addition, the modal instructional practices in East Asian countries are interpreted as driving the high performance of their students. Calls for a more uniform curriculum and national testing in the U.S. often invoke national average characteristics of East Asian schools and classrooms.

Cross-national qualitative studies have provided us with a deeper understanding of the historical, social, and cultural value and meaning of the instructional practices prevalent in East Asian classrooms (Tobin, Wu, and Davidson 1989; Paine 1990; Rohlen and LeTendre 1996; LeTendre 1999; Tobin, Hsueh, and Karasawa 2009). They elaborate on how teachers make sense of their instruction, how the instruction is received, and in turn, how it may contribute to high performance. Still, this work falls short of considering variation in classroom instruction that may exist within a country or of testing the presumed effect of instructional practice on achievement.

This study addresses this gap by using the 2007 Trends in International Mathematics and Science Study (TIMSS) to delve into (1) the within-nation variability in the social organization of instruction among classrooms in East Asia in addition to the country average features, (2) the extent to which complexity or regimented simplicity accounts for higher student achievement within the East Asian countries, and (3) the extent to which schools or school systems nonrandomly assign students into more or less complex classrooms, which conditions the truly independent effect of instructional organization on pupil achievement.

Overall, the results show that on average East Asian classrooms do have less complex instruction and more teacher-centered instruction compared to classrooms in other countries, but they display considerable variation across classrooms. The within-country analyses show that, classrooms with more complex and student-centered instruction show higher achievement; an opposite direction of association is found when the between-country pattern is examined. Moreover, the within-country associations diminish or disappear when class average socioeconomic status or school characteristics are taken into consideration. These findings yield a more textured understanding of how variability in instructional organization ties to achievement within East Asian nations and how these relationships are conditioned by features of classroom and school structures and social-class stratification of families.

2.2 Literature review

This section reviews three lines of literature on classroom organization and teaching practices in East Asia. Earlier studies, based on either observation or large-scale data, focused on gross national comparison of distinctive patterns of classroom practices. Other qualitatively oriented studies were geared toward more complex and contextualized descriptions of schooling and the cultural meaning underlying surface differences. A smaller set of quantitative studies began exploring the link between classroom practices and student achievement in East Asian countries. These three lines of work inform us on East Asian classroom instruction at different levels but rarely intersect with each other. Informed by this body of work, this paper attempts to extend the line of inquiry by investigating the variation and effect of classroom practices *within* the East Asian countries.

Gross national comparison of classroom teaching

Many earlier studies on East Asian schools were prompted by the superior academic performance of East Asian students on international measures. Researchers, often led by Western scholars, have attended to distinct national patterns of classroom organization and pedagogy in East Asian classrooms (Schmidt et al. 1996; Stigler and Hiebert 1999; Hiebert et al. 2003). The primary goal of these observational studies was to identify the weaknesses of schooling in the United States. Teaching patterns that are rarely observed in American classrooms but are present in East Asian classrooms were pointed out as potential areas of improvement for American schools. Harold Stevenson, James Stigler, and their associates' cross-national studies, which include Japanese, Taiwanese, German, and American classrooms, take this gross national comparison approach (Stevenson and Stigler 1992; Stigler and Hiebert 1999). They highlight behavioral patterns in Japanese or Taiwanese classrooms, such as prevalence of whole-class academic activities in elementary classrooms or devotion to recitation and memorization strategies, and contrast them with those in American classrooms. The observed cross-national differences are assumed to stem from distinct national cultures that are believed to be relatively stable across countries and homogenous within countries (Stigler and Baranes 1988; Stevenson and Stigler 1992).

Using a similar gross comparison approach, other researchers take the unit of analysis up to the cross-cultural level and compare schooling in Eastern and Western countries (Watkins and Biggs 1996; Leung 2001; Li 2002; Tweed and Lehman 2002). The distinctive schooling patterns and the high achievement in East Asian countries are attributed to Confucian cultural heritage shared by these nations (Leung 1995). Studies that employ the East–West comparative framework assume that different cultural models of teaching and learning "work" in these different cultural systems. Dahlin and Watkins (2000) reported that East Asian students tend to think of understanding as a long process that requires mental effort in the part of the learner, whereas Western students tend to view learning as a process of sudden insight and attribute academic success to less controllable factors such as inherent ability. Pragmatic orientation and instrumental approaches to learning in East Asian countries were also reported as part of the Confucian conception of education (Tweed and Lehman 2002).

International surveys such as the Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), which collect information on classrooms and teaching practices, have allowed scholars to make gross national comparisons of

instructional practice for a broader range of countries (Stevenson and Baker 1991; Schmidt et al. 2001; Desimone et al. 2005), often with the direct intention of understanding the high achievement of East Asian countries (Leung 2002; Shen 2005). Based on cross-national comparison of various features of national educational systems, William H. Schmidt and his colleagues found that instructional topics tend to be more scattered in American classrooms, leading to disconnected learning opportunities for students (Schmidt et al. 1999; Schmidt et al. 2001). By contrast, classroom practices in high-performing East Asian countries such as Japan and Korea were reported to be more coherent with the national curriculum (Schmidt et al. 1999). Some studies that intended to do an East–West comparison using TIMSS data failed to find a consistent set of variables that explained the high achievement of East Asian countries and ended up resorting to "cultural effect." (Leung 2002; Shen 2005). Other studies found that country average classroom practices were quite similar in terms of time spent on various instructional activities (LeTendre et al. 2001) and the emphasis on conceptual or computational instruction (Desimone et al. 2005).

Complex and contextualized portrayals of East Asian classrooms

While gross comparative studies provide a broad-brush description of classroom instruction in East Asian countries, a substantial body of literature presents a more contextualized and nuanced picture. The most prominent and well-documented research has taken place in Japan (Rohlen 1983; Tobin, Wu, and Davidson 1987; Lewis 1995; Rohlen and LeTendre 1996; LeTendre 1999; Fukuzawa and LeTendre 2001). Refuting the monolithic image of child-centered or drill-oriented Japanese classrooms, field studies find that classroom structure and pedagogical approach change over the course of schooling as a result of the onset of high-stakes national tests (Baker and LeTendre 2000). Scholars also explain that the educational system as a whole is built in line with Japanese assumptions about human maturation and that classroom organization at different levels of schooling reflects the developmental expectation for each stage (Rohlen and LeTendre 1996). Studies further place the static portrait of Japanese classrooms into context. Scholars explain how school organizational structure interacts with or reinforces instruction in classrooms (Yang 1999; LeTendre et al. 2001) and how various non-academic activities in schools function to balance out the intense and text-oriented classroom instruction (Fukuzawa 1996).

Another line of scholarship has focused on the so-called "East Asian classroom paradox," which refers to the seemingly incompatible phenomena of the high achievement of East Asian students and classroom features that are assumed to be less conducive to learning, such as teacher-centered instruction and rote learning (Watkins and Biggs 1996; Leung 2006; Mok 2006; Zhou, Lam, and Chan 2012). Some suggest that the highly structured classroom activities and teachers with strong subject knowledge compensate for practices that are deleterious to learning and eventually lead to higher student achievement (Ma 1999; Wang and Paine 2003; An, Kulm, and Wu 2004). Other studies report that carefully designed official curricula and textbooks and teacher collaboration in planning the lessons further assist in producing well-structured lessons (Fan and Kaeley 2000; Lewis, Perry, and Murata 2006). Researchers of mathematics education and learning theory discovered a common instructional pattern in East Asian classrooms that involves different levels and types of variations and explains how students acquire understanding in teacher-dominated classrooms. They argue that the use of this instructional strategy enables students to discern critical dimensions of the learning topic (Lianghuo et al. 2004; Fang and Gopinathan 2009).

Ethnographic studies take even a deeper look into East Asian classrooms. Scholars point out that classroom practices hold different values and meanings within each social, historical, and cultural setting and have different educational implications. For example, in East Asia, memorization is viewed as complementing deeper understanding, as opposed to the general Western notion that memorization equates with rote learning (Kember 1996; Marton, Dall'Alba, and Tse 1996). Moreover, studies report that what translated as a content-oriented and teacher-dominated lesson from a Western perspective was perceived as a student-centered one from an East Asian perspective, in that the teacher developed a framed experience in consideration of how his students think and learn (Mok 2006; Clarke and Xu 2008). Other studies find that teachers may appear authoritarian or student centered depending on the setting and that such contingent shifts are natural in the Confucian cultural context. For example, Chinese teachers tend to keep absolute control over the situation in the classroom but interact with students on more friendly terms outside the classroom (Paine 1990; Ho 2001).

These rich observational studies begin to break away from the conventional assumption of within-country homogeneity prevalent in gross national comparison studies by exploring variability within countries. Joseph Tobin and his colleagues' studies of preschools in China, Japan, and the United States have examined variation within cultures by showing videos of instructional practices in preschools to members within and outside of the preschool community and collecting the varied responses to the video cues (Tobin, Wu, and Davidson 1989; Tobin, Hsueh, and Karasawa 2009). Earlier Japanese studies also acknowledged the vastly different teaching approach observed across grade levels (Rohlen and LeTendre 1996) and school types (Kariya and Rosenbaum 1999). But little is known about whether classroom practices differ across classrooms in the same grade level and how the variation in classroom practices associates with varied student achievement.

Relationship between instructional organization and achievement within country

Finally, a number of studies have interrogated the relationship between classroom practices and student achievement using large-scale data (Schaub and Baker 1991; Paik 2004; House and Telese 2008). Using a common cross-national quantitative analytic strategy that compares the effect of variable of interest on student achievement in different national settings, these studies have explored the association between the learning environments and student achievement within East Asian countries. Studies report that instructional strategies that use real-world examples and independent learning activities are positively related to mathematics achievement in Japan and Taiwan (House 2002, 2009). By contrast, a study on Japan finds a negative association between student individual or group work and mathematics achievement and a positive association between teachers' time spent on whole-class lecturing and mathematics achievement (Schaub and Baker 1991). Other studies report no statistically significant association between classroom instructional organization and student achievement (Paik 2004; Mohammadpour 2013). The mixed results may be partly attributable to the different data and items involved in the analyses. Setting aside this matter, because most of the studies have failed to control for student, classroom, or school background variables, it is likely that the estimated associations between classroom practice and achievement are biased. Moreover, most studies used OLS regression analyses when multilevel modeling would have been appropriate given the nested structure of the data. Failing to recognize the hierarchical structure may have resulted in misestimating the standard errors of the instructional practice coefficients (Raudenbush and Bryk 2002).

Recognizing heterogeneity and increasing diversity in schools, classrooms, and students

The three lines of literature discussed thus far has rarely taken into consideration the variation of classroom practices that may exist across classrooms within the East Asian countries. The assumption of within-country homogeneity has been more readily made about East Asian countries because of their relatively small and/or less diverse populations and their long-maintained standardized and tightly controlled educational systems (Schmidt et al. 2001). Or it was argued that any initial variation in student ability or background becomes irrelevant because schools and classrooms are so effective at closing the achievement gap (White 1987; Stevenson and Stigler 1992). But as continually argued by the research community, using national averages of educational characteristics (e.g., national averages of teacher-centered instruction) without consideration of the variability within nations may only lead to a superficial understanding of schooling in other countries (Baker 1997; Torney-Purta and Amadeo 2013).

Changes in East Asian education in recent decades further necessitate taking within-country diversity into consideration in researching classroom instructions. Since the early 2000s, many East Asian nations have been experimenting with curricular decentralization (Pennington 1999; Bjork 2006; Waldow, Takayama, and Sung 2014) and incorporating educational reform measures such as ability grouping within schools, school choice, and decentralization of school funding (Bjork 2009; DeCoker and Bjork 2013; Park 2013). As the organization of schooling becomes more diversified, it is natural to ask how the different types of instructional strategies are distributed within a country and how the effects of classroom practices on achievement are confounded by diversity of forms of schooling.

This study extends the previous literature on East Asian classrooms by considering heterogeneity of schooling within East Asian countries and asks the following research questions:

1. To what extent do classroom instructional practices in East Asian countries differ from those in other countries?
    A. Are classroom instructional practices in East Asian countries more procedural and teacher-centered and less complex and student-centered compared to other countries?
    B. Are classroom instructional practices more homogeneous among classrooms within East Asian countries compared to other countries?

2. How are classroom instructional practices associated with student achievement within East Asian countries?

3. How does the association between classroom instructional organization and student achievement change after controlling for student, classroom, and school background variables?

2.3 Analytic strategy

To address the research questions, I use the Trends in International Mathematics and Science Study (TIMSS) 2007 data to define the four key aspects of instructional organization, the prevalent use of complex and procedural tasks, and student- and teacher-centeredness. Firstly, I examine how the country means and within-country variation of the four aspects of classroom instructional organization in the five East Asian countries compare with those in the other 45 nations in the sample. Secondly, I narrow my focus to the two of the East Asian countries that have vastly different school structures, Japan and Singapore, to examine the within-country variability of classroom practices and how they relate to student achievement. Thirdly, I examine whether and how classroom instructional organization maintains effects on student achievement after taking into account diversity of schools, classrooms, and students.

2.4 Methods

Data

The data set used is the Trends in International Mathematics and Science Studies 2007 (TIMSS 2007), conducted by the International Association for the Evaluation of Educational Achievement (IEA). TIMSS provides mathematics and science assessment scores of nationally representative fourth- and eighth-grade students and a wealth of information on schools, curricula, teachers, and students in over 50 countries. In each country, a two-stage stratified clustered sampling design is conducted, with schools sampled in the first stage and intact classrooms sampled in the second stage. Typically, countries sampled 150 schools and one or two intact classrooms. TIMSS 2007 eighth grade mathematics data contains information on about 240,000 students and about 12,000 teachers from 50 countries and seven benchmark participants.

To analyze country average instructional practices and within-country variation of instructional practices, a sample of 8,864 teachers in 50 countries with complete information on their teaching is used. In estimating the relationship between instructional practice and student achievement, it was essential to link teachers to the class or group of students they taught. While most classes in Japan and Singapore were exposed to one mathematics teacher, some classes were taught by different teachers. In cases where classrooms are administrative units and not necessarily learning units, I have treated the multiple teachers within classrooms as separate units. In estimating the effect of instructional practice on student achievement in Japan and Singapore, I used an operational sample of 140 teachers and 2,831 students in Japan and 284 teachers and 2,950 students in Singapore. All statistics reported use the appropriate student, teacher, and school sampling weights, as recommended in the TIMSS 2007 User Guide (Foy and Olson 2009); methods for sampling weights are detailed below.

Measures

Table A2.1 in the Appendix provides detailed descriptions of all variables used in this study. Descriptive statistics of the variables are provided in Table 2.1 and the intercorrelations of the key variables are presented in Tables A2.2 and A2.3.

Table 2.1 Descriptive statistics

Variables	JPN				SGP			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Student level		***N=2831***				***N=2950***		
Mathematics Score (PV1)	577.2	84.72	283.7	857.75	594.75	89.08	283.4	835.47
Mathematics Score (PV2)	579.9	86.14	285.9	891.69	595.62	89.91	297.8	833.72
Mathematics Score (PV3)	579.2	86.16	236.5	878.76	595.68	90.3	282.2	819.53
Mathematics Score (PV4)	578.8	85.74	219.8	858.98	594.98	89.94	280.5	826.39
Mathematics Score (PV5)	577.9	85.77	230.8	851.96	595.96	89.21	278.6	828.42
Male	0.47	0.50	0.00	1.00	0.49	0.50	0.00	1.00
Foreign born parent	0.02	0.14	0.00	1.00	0.29	0.45	0.00	1.00
Language at home	0.99	0.11	0.00	1.00	0.49	0.50	0.00	1.00
Books at home	102.4	102.4	5.00	300.00	91.72	96.65	5.00	300.00
Parent education	0.64	0.48	0.00	1.00	0.49	0.50	0.00	1.00
Teacher/Class level		***N=140***				***N=284***		
Instructional Practice								
Complex instruction	2.34	0.47	1.25	4.00	2.28	0.49	1.00	4.00
Procedural instruction	2.74	0.66	1.00	4.00	2.56	0.51	1.33	4.00
Student-centered instruction	1.16	1.05	0.00	5.00	1.37	0.83	0.00	4.30
Teacher-centered instruction	4.38	0.94	2.50	7.50	3.46	1.38	0.40	7.00
Teacher Background								
Male teacher	0.57	0.50	0.00	1.00	0.36	0.48	0.00	1.00
Years of experience	15.93	9.47	1.00	34.00	7.98	9.49	0.00	41.00
Teacher's Degree in mathematics or mathematics degree	0.85	0.36	0.00	1.00	0.77	0.42	0.00	1.00
Class Characteristics								
Class size	33.23	7.41	4.00	47.00	37.91	3.92	20.00	45.00
Different ability of students as limiting instruction	0.30	0.46	0.00	1.00	0.28	0.45	0.00	1.00
Uninterested students as limiting instruction	0.08	0.27	0.00	1.00	0.24	0.43	0.00	1.00
Class average number of books at home	101.7	33.46	31.73	209.41	85.37	46.96	5.00	255.50
Class proportion of parents with college degree	0.63	0.18	0.23	1.00	0.46	0.25	0.00	1.00
School level		***N=123***				***N=152***		
School size	339.2	223.0	20.00	1102.0	1336.5	299.8	248.0	3000.0
Student from affluent family>50%	0.43	0.50	0.00	1.00	0.16	0.37	0.00	1.00
School resource	0.53	0.50	0.00	1.00	0.91	0.29	0.00	1.00
School climate	0.07	0.26	0.00	1.00	0.22	0.41	0.00	1.00
School type (Japan)								
Public school in large city	0.27	0.44	0.00	1.00	-	-	-	-
Public school in small city	0.69	0.46	0.00	1.00	-	-	-	-
Private school	0.04	0.20	0.00	1.00	-	-	-	-
School performance rank (Singapore)	-	-	-	-	15.42	6.64	1.00	30.00

Key explanatory variables: Instructional practices

I used prior research on classroom instruction to sort the teacher survey items into categories of instruction. I select four aspects of classroom instructional practice that are commonly referred to in depicting East Asian classrooms: complex instruction, procedural instruction, teacher-centered instruction, and student-centered instruction. *Complex instruction* denotes classroom practices that emphasize real-world problem solving, elicit student reflection, and involve working with problems with no obvious solution (also called "higher-order instruction," "authentic pedagogy," or "conceptual instruction") (e.g., Carpenter et al. 1989; Raudenbush, Rowan, and Cheong 1993; Newmann, Marks, and Gamoran 1996; Desimone et al. 2005). In mathematics, it is believed that complex instruction allows students to develop a deeper understanding of mathematical concepts and logic. A greater emphasis on complex instruction is widely advised in the United States and is a core component of current state and national education policies (National Council of Teachers of Mathematics 2000; Common Core State Standards Initiative 2010). To determine teachers' emphasis on complex instruction, I used four items that asked how often a teacher asks students to explain the reasoning behind an idea, relate the learning topic to their daily lives, work on problems with no immediately obvious method of solution, and decide on their own procedures for solving complex problems. Response options ranged from never (1) to always (4). An average of the four items formed the complex instruction variable. Cronbach's alpha is 0.66.

Procedural instruction refers to classroom practices that emphasize computational skills and involve memorizing facts and procedures (also called "basic skills instruction" or "traditional instruction") (e.g., Hamilton and Martinez 2007). Too much reliance on procedural instructional practices is generally thought to hinder students from developing a deeper understanding of the mathematical concepts behind the procedures (Romberg 1988; National Council of Teachers of Mathematics 2000; Baroody and Benson 2001). I used three items that asked how often teachers ask students to practice adding, subtracting, multiplying, and dividing without using a calculator; memorize formulas and procedures; and apply facts, concepts, and procedures to solve routine problems, which I categorize as procedural instruction. An average of the three items is used as the procedural instruction variable. Cronbach's alpha is 0.52 for the procedural instruction scale.

To determine teachers' emphasis on teacher- and student-centered instruction, I used another set of teacher survey questions that ask teachers to write in the percentage of time they spend on a list of activities, which should add up to 100%. Two of the seven items were used to derive the teacher-centered instruction variable: (a) lecture-style presentation by teachers and (b) re-teaching and clarification of content/procedures. The sum of the percentages of time divided by 10 was used as the teacher-centered instruction variable. For the student-centered instruction variable, the percentage of time spent on student independent practice divided by 10 was used.

Dependent variable: student mathematics achievement

For the student achievement measure, I used the student mathematics test scores, which consist of five plausible values. TIMSS conceptualizes student mathematics achievement as a latent construct measured with uncertainty and provides five separate imputed scores. Any analysis that involves student achievement should be replicated five times, and the results should be aggregated to obtain the final estimate as well as the imputation error (Foy and Olson 2009). Thus, I use all five plausible values for the analysis in this paper.

Classroom and school characteristics

To take into account possible nonrandom assignment of students to classrooms that emphasize different types of instructional practices, I included school and classroom characteristics that may be simultaneously correlated with classroom instruction and class average achievement. Classroom aggregates of students' socioeconomic background variables, number of books at home, and parent education capture the effect of class socioeconomic status on average class achievement over and above the effect of individual socioeconomic status on student achievement within a classroom. I also included teachers' reports on whether different abilities of students and uninterested students limited their instruction. These variables reflect teachers' perceptions of class characteristics, which may have affected their choice of instructional strategies. At the school level, I included school size, a dummy variable that indicates whether 50% or more of students are from affluent families, a dummy variable for high mathematics resource, and a dummy variable for high school climate. I also included school characteristics variables that were reported by each TIMSS participant country for stratified sampling of schools. School type information (public and private schools) and school performance ranking were used in the analyses of Japan and Singapore, respectively.

Control variables

At the individual student level, I include five control variables: number of books in students' homes, parent education, student's sex, whether parents are foreign born, and frequency of speaking the test language at home. The first two variables are proxies for the student's socioeconomic background. Because class mean socioeconomic status variables are included in the model, these variables indicate the effects of individual socioeconomic status after controlling for the class mean, i.e., effects of socioeconomic status within the classrooms. I included student sex because gender differences in math achievement are traditionally prevalent in many countries. At the teacher level, teachers' years of teaching experience and educational degree were included as proxies for teachers' content knowledge and pedagogical content knowledge. Teachers' sex is also controlled.

Estimation approach

The analysis is divided into two parts. First, in order to address the first set of research questions, I compare the prevalence and within-country variation of instructional practices in East Asian countries with the other 45 countries. Specifically, I examine whether classrooms in East Asian countries have more procedural and teacher-centered instruction and less complex and student-centered instruction, and whether there is less within-country variation in classroom instructional practices in East Asian countries. Prevalence of an instructional practice is measured by the country-level mean, and within-country variation is measured by the within-country coefficient of variation (standard deviation divided by mean, multiplied by 100). The measures are derived by using mathematics teacher weight variable (MATWGT). Using MATWGT allows for describing the instructional practices faced by nationally representative samples of students (Foy and Olson 2009). Note that the sampling design of TIMSS, where the second-stage sampling units are classrooms and not teachers, places a constraint on the interpretation of teacher data analysis. Inferences should be made at the student level and not the teacher level.

Second, I focus on two East Asian countries, Japan and Singapore, to examine the effects of instructional practices on student achievement. These two countries were chosen because they stand at the extreme ends of the spectrum in their institutional arrangements, Japan having an integrated school system and Singapore having a differentiated system. To take into account the data structure

(students nested within classes nested within schools), I employ multilevel models to gauge the effects of instructional practice on student achievement. Singapore has more than one classroom sampled within each sampled school, which makes it possible to specify three-level models, while Japan sampled one classroom per school, which allows for two-level models. A three-level model for Singapore is specified as follows:

$$Y_{ijk} = \gamma_{000} + \beta_{01} IP_{jk} + A_{ijk}\boldsymbol{\pi} + X_{jk}\boldsymbol{\beta} + W_k\boldsymbol{\gamma} + u_{00k} + r_{0jk} + \varepsilon_{ijk}$$

$$u_{00k} \sim N(0, \tau_\beta)$$

$$r_{0jk} \sim N(0, \tau_\pi)$$

$$\varepsilon_{ijk} \sim N(0, \sigma^2)$$

where, Y_{ijk} is the mathematics score of student *i* in classroom *j* for school *k*, γ_{000} is the country mean mathematics score, β_{01} is the effect of instructional practice on achievement, IP_{jk} is the instructional practice in classroom *j* in school *k*, A_{ijk} is a vector of student-level covariates, X_{jk} is a vector of class-level covariates, W_k is a vector of school-level covariates, and u_{00k}, r_{0jk}, and ε_{ijk} are the school-, class- and student-level random effects. All continuous variables are country-mean centered and standardized.

TIMSS uses a two-stage stratified clustered sampling design, which requires an appropriate use of sampling weights when running multilevel models. The multilevel models are estimated with HLM 6.0 software using weights that are manually calculated for each level (Rabe-Hesketh and Skrondal 2006). For three-level models, in case of Singapore, the total student weight is decomposed into the inverse of the probability of selection for a school, for a class within a school given that the associated school was selected, and for a student given that the student's school and class were selected. For two-level models, in case of Japan, class weight is used for the second level and student weight is used for the first level. Using HLM software also allowed me to take into account all the five plausible values of mathematics achievement for the analyses (Rutkowski et al. 2010).

2.5 Results

Cross-national comparison of the prevalence and within-country variation of classroom instructional practices in East Asian countries

Let's begin by examining the instructional attributes in East Asian countries compared to those in other countries. Is classroom instruction in East Asian countries more procedural and teacher-centered compared to that in other countries? Are classrooms in East Asian countries more homogeneous within each country in terms of instructional practices compared to those in other countries? Figure 2.1 addresses these questions by plotting the country average and coefficient of variation of instructional practices in the 50 countries (See Table A2.4 for the values of each country).

With respect to complex instruction, four of the East Asian countries (the exception being Korea) showed the lowest scores among the 50 countries. Less distinctively, East Asian countries scored lower in student-centered instruction. This result is consistent with the traditional notion that East Asian classrooms are less student-centered and less complex. But more importantly, classrooms within East Asian countries vary considerably, as indicated by the large coefficients of variation in complex, procedural, and student-centered instruction. The result counters the general assumption of homogeneity within East Asian countries and cautions us to refrain from generalizing an average or typical pattern to the whole country. On the other hand, classrooms in East Asian countries did consistently employ teacher-centered instruction. Four of the East Asian countries (the exception being Singapore) not only ranked highest in the teacher-centeredness scale but also had small within-country variations compared to other countries in the sample.

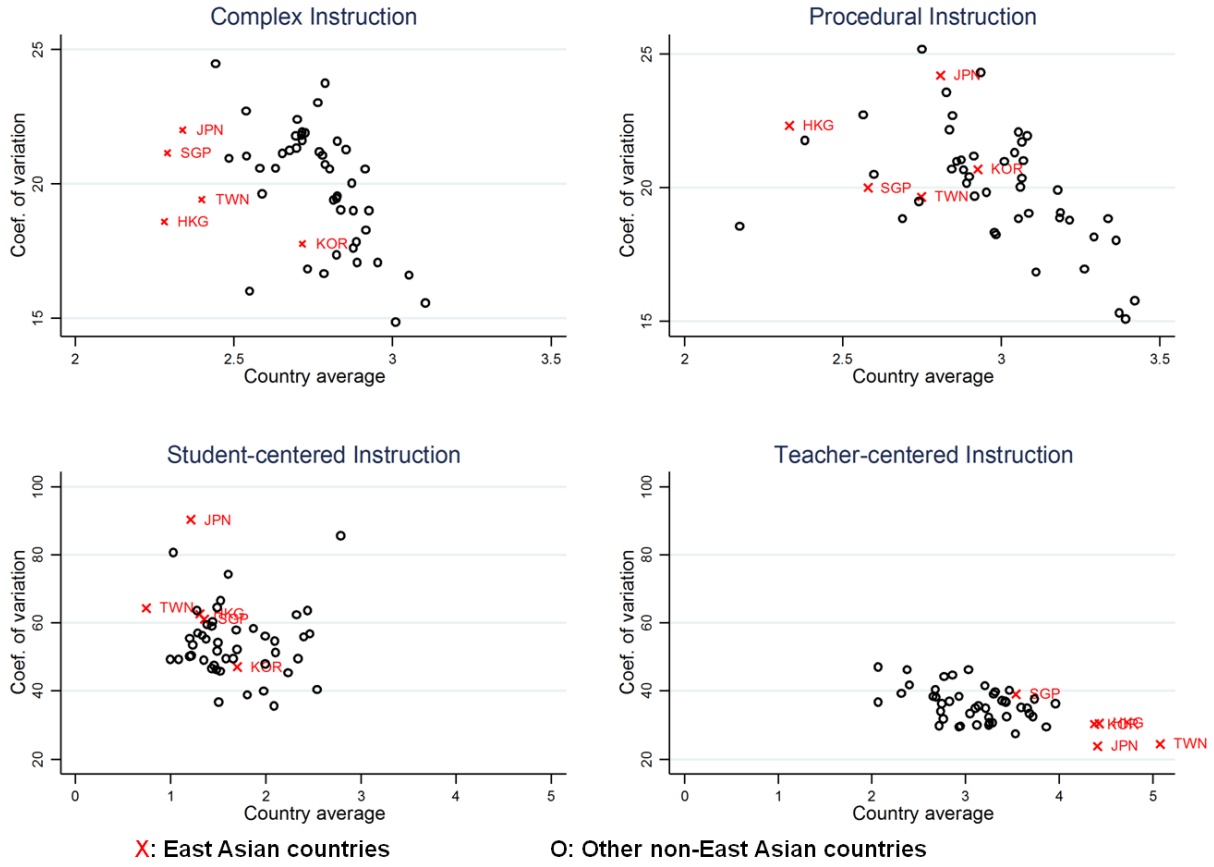


Figure 2.1 The relationship between country average and coefficient of variation of instructional practices

Within-country association of instructional organizations and student achievement

Next, I examine whether and how the varied instructional practices explain student achievement within the country. The effects of instructional practices on student achievement are examined in two East Asian countries, Japan and Singapore. As a preliminary analysis, the variance in student math scores is decomposed into student, classroom, and school-level variances in Japan and Singapore, along with the other East Asian countries (Table A2.5). Note that the total variation in math scores that stem from between classrooms or school levels was smallest in Japan (19%) and largest in Singapore (77%) among the five East Asian countries. That is, 81% of the variance in math scores is attributable to between-student sources within classrooms in Japan, and 23% in Singapore. Different patterns of apportioning reflect the different educational systems of the two countries, namely, the comprehensive system in Japan and the stratified system in Singapore. This signals possible sorting between schools or classrooms, especially in Singapore, and further necessitates considering the diversity in school, classroom, and student characteristics when examining the association between classroom instructional organization and student achievement.

Tables 2.2 and 2.3 present the estimated effects of instructional practices on student achievement in Japan and Singapore. Table 2.2 shows that in Japan, classrooms with more complex instruction tend to have higher class average mathematics scores (model 1). Other classroom instructional practices are not significantly associated with student achievement (models 4, 7, and 10). Models that only include classroom instructional practices (not shown) explain only 2% to 4% of student achievement variance at the classroom level. This may be partly due to the already small achievement variation across classrooms in Japan. As shown in Table 2.3, in Singapore, the effects of complex, student-centered, and teacher-centered instruction on student achievement in models with no covariate controls are significantly positive (models 1, 7, and 10). Supplementary analyses that only include classroom instructional practices in the models show that classroom instruction variables accounted for 4% to 6% of achievement variance at the school level and 2% to 5% of variance at the classroom level.

Table 2.2 Relationship between instructional practices and student achievement in Japan

	Complex			Procedural			Student-centered			Teacher-centered		
	model1	**model2**	**model3**	**model4**	**model5**	**model6**	**model7**	**model8**	**model9**	**model10**	**model11**	**model12**
Instructional practice												
Complex instruction	7.20**	2.16	3.29									
Procedural instruction				-0.92	0.10	-0.13						
Student-centered instruction							-2.57	-3.08	0.85			
Teacher-centered instruction										4.71	4.79+	1.57
Student Background												
Male	5.13	5.63+	6.38*	5.24	5.68+	6.45*	5.20	5.62+	6.47*	5.23	5.68+	6.44*
Foreign born parent	-22.06+	-21.87+	-22.68+	-22.18+	-21.96+	-22.80+	-22.21+	-22.02+	-22.77+	-22.22+	-22.02+	-22.83+
Language at home	39.17+	40.35+	40.98*	39.21+	40.36+	41.01*	39.27+	40.51+	40.97*	39.20+	40.37+	41.00*
Books at home	13.52***	12.50***	12.47***	13.52***	12.50***	12.46***	13.52***	12.50***	12.46***	13.56***	12.50***	12.46***
Parent education	35.34***	33.06***	33.07***	35.56***	33.06***	33.07***	35.57***	33.06***	33.07***	35.54***	33.06***	33.07***
Teacher Background												
Male teacher	-17.36**	-8.28	-14.02**	-15.96**	-7.53	-12.92**	-15.89**	-7.47	-13.00**	-15.04*	-6.66	-12.58**
Teacher's years of experience	0.93	0.08	-0.18	-0.29	-0.31	-0.76	-0.26	-0.37	-0.75	-0.18	-0.14	-0.70
Teacher's Degree in math or math education	10.54	11.49	13.07	9.83	11.44	12.73	11.44	12.99	12.26	12.48	13.73	13.50
Class Characteristics												
Class average number of books at home		11.08**	4.90+		6.63*	2.20		10.82**	5.01+		11.47**	5.23+
Class proportion of parents with college degree		9.31**	4.20		11.11**	5.00+		9.79**	5.21		9.34**	5.12
Class size		6.45*	1.93		9.83**	5.17		7.03*	2.07		6.07*	2.14
Different ability of students as limiting instruction		-2.45	1.65		-1.95	2.42		-1.97	2.59		-3.53	1.90
Uninterested students as limiting instruction		-8.92	-4.68		-10.37	-7.07		-12.42	-6.50		-13.49+	-8.16
School Characteristics												
School size			0.89			1.17			1.09			0.88
Student from affluent family>50%			0.33			0.36			0.44			0.53
School resource			-7.97			-8.03			-8.12+			-8.29+
School climate			23.16**			23.99**			24.59**			23.32**
Public school in large city			-0.14			-1.07			-1.43			-1.00
Private school			68.70***			66.26***			66.50***			65.29***
Intercept	512.12***	508.01***	504.57***	511.48***	507.51***	504.29***	510.00***	506.23***	504.76***	508.50***	505.51***	503.92***
Random Effect												
Variance explained(classroom)	36.3%	59.4%	77.5%	32.8%	59.1%	76.7%	33.4%	59.9%	76.7%	34.4%	60.6%	76.9%
Variance explained(student)	8.2%	8.2%	8.1%	8.2%	8.2%	8.1%	8.2%	8.1%	8.1%	8.2%	8.1%	8.1%
N(classroom)	140	140	140	140	140	140	140	140	140	140	140	140
N(student)	2831	2831	2831	2831	2831	2831	2831	2831	2831	2831	2831	2831

+p<0.1, *p<0.05, **p<0.01, ***p<0.001

Table 2.3 Relationship between instructional practices and student achievement in Singapore

	Complex			Procedural			Student-centered			Teacher-centered		
	model1	**model2**	**model3**	**model4**	**model5**	**model6**	**model7**	**model8**	**model9**	**model10**	**model11**	**model12**
Instructional practice												
Complex instruction	17.60**	6.79+	5.90+									
Procedural instruction				7.40	1.56	3.47						
Student-centered instruction							17.44***	10.84**	7.34*			
Teacher-centered instruction										13.94**	7.07*	4.40
Student Background												
Male	4.44+	4.70+	4.53+	4.41+	4.69+	4.49+	4.41+	4.66+	4.52+	4.42+	4.67+	4.53+
Foreign born parent	6.97**	6.95**	7.03**	6.97**	6.94**	7.01**	6.96**	6.95**	7.02**	6.98**	6.95**	7.02**
Language at home	1.65	0.90	0.68	1.63	0.89	0.67	1.66	0.88	0.67	1.67	0.92	0.71
Books at home	0.63	-0.18	-0.17	0.64	-0.18	-0.17	0.65	-0.18	-0.17	0.66	-0.18	-0.17
Parent education	-0.40	-1.35	-1.30	-0.42	-1.35	-1.29	-0.36	-1.34	-1.30	-0.45	-1.35	-1.31
Teacher Background												
Male teacher	-17.87	-10.18	-15.34*	-17.37	-9.96	-15.16*	-17.05	-9.82	-14.92*	-19.36+	-10.84	-15.69*
Teacher's years of experience	3.00	1.30	1.78	1.85	0.68	1.53	2.24	1.26	1.72	-0.75	-0.23	0.71
Teacher's Degree in math or math education	-6.94	2.79	1.85	-2.76	4.74	3.08	-4.45	2.93	2.08	-4.42	3.64	2.45
Class Characteristics												
Class average number of books at home		46.36***	35.58***		47.60***	35.95***		46.53***	36.21***		48.02***	37.11***
Class proportion of parents with college degree		15.77***	9.31*		15.68***	8.84*		16.67***	10.11*		14.51**	8.92*
Class size		14.05***	15.73***		13.57***	15.10***		14.37***	15.46***		13.02***	15.13***
Different ability of students as limiting instruction		-14.55+	-13.86+		-16.71+	-15.90*		-15.88+	-15.00+		-16.22+	-15.64+
Uninterested students as limiting instruction		-23.97*	-23.18**		-23.44*	-23.85**		-19.19*	-20.04*		-23.12*	-23.06**
School Characteristics												
School size			-13.23***			-13.39***			-11.85**			-3.02
Student from affluent family>50%			-1.68			-0.95			-0.47			-1.58
School resource			-0.58			-1.71			-4.36			-0.27
School climate			-1.18			0.18			1.49			-13.45***
School performance strata			29.89***			30.51***			27.87***			28.92***
Intercept	591.11***	594.36***	598.23***	587.61***	593.17***	598.50***	588.83***	593.46***	600.07***	590.09***	594.28***	600.48***
Random effects												
Variance explained(school)	-1.6%	65.3%	92.4%	-11.7%	67.8%	93.5%	-1.1%	64.3%	91.1%	0.1%	70.6%	93.8%
Variance explained(classroom)	9.7%	59.5%	57.4%	6.8%	57.8%	56.5%	8.6%	60.8%	57.9%	5.9%	57.8%	56.5%
Variance explained(student)	0.6%	0.5%	0.5%	0.6%	0.6%	0.5%	0.6%	0.6%	0.6%	0.6%	0.6%	0.5%

+p<0.1, *p<0.05, **p<0.01, ***p<0.001

Recognizing the diversity of schools and classrooms

The final explanatory question asks whether and how the association between classroom instructional organization and student achievement changes when diversity in schools and classrooms is considered. Models in Tables 2.2 and 2.3 that include student, classroom and school background variables address this question. By comparing the coefficients of instructional practice variables before and after including classroom and school organization variables, we can see how much of the effects of instructional practice are actually due to the nonrandom placement of students to classrooms or schools that employ different types of instruction. As illustrated in Table 2.2, in Japan, the effect of complex instruction is no longer significant when class average student background variables are introduced to the model (model 2). The small yet positive and significant relationship between class average parent education and complex instruction may have confounded the effect (See Table A2.2). The final models accounted for about 77% of achievement variance at the classroom level and 8% of variance at the student level.

Table 2.3 shows that, in Singapore, the effects of complex and teacher-centered instruction are attenuated (model 2 and model 11) when classroom characteristics are included in the models. Change in the coefficients is likely due to the positive correlation between classroom averages of student socioeconomic background and both complex and teacher-centered instruction (See Table A2.3). The full models show that the effects of complex instruction and student-centered instruction remain significant after including all school characteristics variables (models 3 and 9). The final models explain over 55% of achievement variance at the classroom level and over 90% of variance at the school level. However, less than 1% of the student-level variance in achievement was accounted for by the final models. This is not surprising given the tiered school system, where there is very little within-classroom achievement variation to begin with. In both Japan and Singapore, the results demonstrate that wealthier students are selecting into classrooms that adopt certain types of instructional approaches or that teachers are differentiating their instruction based on the class average student background. Students' differential exposure to different types of classroom practice is further assisted by educational institutions in Singapore; the school ranks

Not surprisingly, students' family socioeconomic background is highly associated with mathematics achievement in all three countries. Male students and students that use the test language in their homes tend to score higher in mathematics in Japan. Students with foreign-born parents have higher mathematics scores in Singapore and lower scores in Japan. Classrooms where teachers reported uninterested students as limiting their instruction tend to score lower in mathematics in Singapore. In Japan, classrooms with male teachers tend to have lower average mathematics scores, and classrooms in schools with better climates and private schools tend to have higher average mathematics scores. In Singapore, larger classrooms and classrooms in schools with higher performance strata score higher in mathematics on average.

2.6 Discussion

To gain a more holistic understanding of classroom instructional practices in East Asian countries, this paper examined the level and dispersion of complex, procedural, student-centered, and teacher-centered instruction in these countries, and the effectiveness of such practices within the countries. The results complement previous studies on East Asian classrooms by attending to the within-country variation and effects of the four dimensions of instructional practice. By doing so, the study reassesses the long-held stereotypes about classroom practices in East Asian countries and presents empirically grounded evidence that may be useful for further research as well as policy discussion. Overall, there is much more to learn from the East Asian countries if we take into account the classroom practices at both the between- and within-country levels and consider varying organization of schooling, which may confound the link between instructional practice and achievement.

This study has three main findings. First, while classrooms in East Asian countries are indeed among the most teacher-centered, less student-centered and least complex on average, the extent to which classrooms within these countries employ certain practices varies considerably compared to the other 45 countries in the sample. The fact that there is great variability in instructional practices among classrooms demands caution in generalizing about East Asian classrooms based on national average features and assuming causal linkage between national features of East Asian classrooms and their high performance.

This finding also provides an interesting contrast with prior studies that report that countries with national control over curricula, which includes most East Asian countries, have less variation across classrooms in terms of instructional content (Stevenson and Baker 1991; Schmidt and Prawat 2006). The central control of curriculum and the consistency of classroom instructional topics across classrooms created by the centralization are often identified as key factors that produce high performance in East Asian countries. But as my findings show, although teachers may be similar in *what* they teach, *how* they teach these topics differs vastly across classrooms within these countries. Literature on Japanese elementary schools reports that Japanese teachers are trained to design classroom activities in a way that allows students to naturally discover the contents listed in the national curriculum (Lewis and Tsuchida 1997). Whether the same explanation applies in secondary schools could be found out empirically. Moreover, how the variability in classroom instructional organization in East Asian countries might explain their overall high achievement could be further unpacked in future studies.

Learner-centered pedagogy is increasingly becoming a global norm (Bromley, Meyer, and Ramirez 2011), and governments in East Asian countries try to promote more student-centered and active learning approaches in classrooms. But still, East Asian classrooms consistently employ more teacher-centered practices and variably have less student-centered practices. Why we see this broad variability in classroom instructional organization despite the tight control of curriculum content and standardized teacher education is a question that may guide future research. The variability may reflect the multiple models of "good teaching" present in the system. And some schools or classrooms may be faster at incorporating the new models, while others adhere to the traditional mode of teaching. How teachers adopt, negotiate, or resist the new models of instructional practice can be empirically explored through appreciation of the heterogeneity within the field.

Secondly, by examining the within-country association between classroom instructional organization and student achievement in two East Asian countries, Japan and Singapore, this study finds that complex or student-centered instruction is positively associated with student achievement within these countries. The directions of these associations within the countries contrast with those at the between-country level, where classrooms in East Asian countries are overall high in teacher-centeredness and low in complexity compared to those in other countries and rank the highest in international assessments. The result indicates that association patterns of instructional practices and student achievement at the between-country and within-country levels are distinct, reflecting different causal mechanisms that warrant further investigation in future research.

The third finding shows that the associations between classroom practices and student achievement change after adjusting for school and classroom characteristics. In both Japan and Singapore, the positive association between complex or student-centered instruction and achievement decreased in size or lost its statistical significance after adjusting for class average student socioeconomic background. This could be a result of sorting, as suggested by many U.S.-based studies. Studies have reported that teachers of students from low socioeconomic backgrounds tend to spend more time on procedural learning, while teachers of affluent students tend to present more complex material and emphasize problem-solving skills (Oakes 1985; Lucas and Gamoran 2002; Bodovski and Farkas 2007; Maaz et al. 2008; Desimone and Long 2010). The data used in this study do not allow us to specify to what extent students are differentially sorted into classrooms with different pedagogical approaches, how teachers adjust their instructional practices based on the ability or background of the student body, and how much students actually benefit from the instructional practices. Further studies that use qualitative approaches could examine what instructional approaches are sought by parents and whether certain practices are deemed desirable for particular students within the East Asian context. Quantitative studies that use quasi-experimental methods could also help unpack causal relationships and possible sorting through parental choice or teachers' selective adjustment.

Another point related to the third finding is that differential exposure to certain types of instructional practices is conditioned by institutional structures as evidenced by the different patterns in Japan and Singapore. This is consistent with findings of U.S.-based studies, different types of instructional modes being differentially provided through institutional structures such as curriculum tracks or ability grouping (Barr, Dreeben, and Wiratchai 1983; Oakes 1985; Raudenbush, Rowan, and Cheong 1993; Zohar, Degani, and Vaaknin 2001). Evidence of institutional differentiation is especially pronounced in Singapore, as classrooms in higher-ranking schools tend to employ more complex tasks and teacher-centered instruction.

Institutional sorting is less prominent in Japan, perhaps due to the egalitarian and comprehensive lower secondary educational system. Supplementary analyses of Japanese data suggest that private schools enroll students from wealthier families and tend to have less student-centered instruction compared to public schools. This is consistent with previous Japanese studies that found that middle-class families send their children to private schools in pursuit of academically rigorous instruction, avoiding the relaxed curriculum promoted by the government (Kariya and Rosenbaum 1999). Increased stratification after governments' decentralization efforts has been reported in other East Asian countries as well (Bjork 2006; Wu and Zhang 2010; Kariya 2011). Future studies that use longitudinal data from before and after decentralization and include classrooms and schools of various types could examine whether and how much the interplay between the institutional structure

of schooling, classroom instructional practices, and achievement increased opportunity gaps and intensified inequality after the decentralization policies were implemented.

Often, classroom practices prevalent in East Asian countries are celebrated as best practices to emulate, criticized for their rigidity, or simply regarded as non-importable due to the peculiarity of East Asian culture. But as this study shows, East Asian countries face educational challenges similar to those observed in the U.S. and elsewhere. Once we acknowledge the commonality, a variety of research questions arise, which may guide future research. For example, this study finds that East Asian governments' efforts to promote certain types of teaching practices have little impact on how teachers teach their classes, which is a *decoupling* of policy and classrooms. Why do we see this decoupling even under the condition of strong central control of the curriculum? What accounts for the broad variation in teacher classroom practices when teachers teach more or less similar curriculum content? How might cultural factors play a role in the policy implementation process? This article also reports that access to certain types of classroom instructional practices is unequally distributed across students by socioeconomic status or previous achievement in East Asian countries, as is the case in other developed nations. What explains the link across class average socioeconomic status, instructional practices, and student achievement in Japan, where the system is comprehensive and no institutional sorting is in place, at least in junior secondary level public schools? To what extent do teachers tailor the level of their instruction to their perception of students' prior learning experience and learning needs? What role do parents play? In Singapore, where schools are ranked by student achievement, how do teachers differentiate their teaching by students' ability? How do cultural beliefs about teaching and learning affect this process? These questions will allow us to better understand our own problems and assumptions about educational processes and theories and could lead comparative education research in more fruitful directions.

Appendix A2

Table A2.1 Description of variables

Variables	Description
Student math achievement	Five plausible values of the math IRT scores (bsmmat01-05)
Instructional practice	
Complex instruction	Average of four items: teachers' reported frequency on asking students to explain the reasoning behind and idea (bt4masea), relate the learning topic to their daily lives (bt4masdl), work on problems with no immediately obvious method of solution (bt4masws), and decide on their own procedures for solving complex problems (bt4mascp). Recoded the original item from 1 "never" to 4 "every or almost every lesson"
Procedural instruction	Average of three items: teachers' reported frequency on asking students to practice adding, subtracting, multiplying, and dividing without using a calculator (bt4maspc), memorize formulas and procedures (bt4masmf), and apply facts, concepts and procedures to solve routine problems (bt4masac). Recoded the original item from 1 "never" to 4 "every or almost every lesson"
Student-centered instruction	Teachers' report on the percentage of time students spend on working on own (bt4mptoo). Percentage is divided by 10.
Teacher-centered instruction	Teachers' report on the percentage of time students spend on listening to lecture-style presentation (bt4mptls) and listening to the teacher re-teach and clarify content/procedures (bt4mptrt). Percentage is divided by 10.
Student Background	
Male	Coded 1 if male, and 0 if female (bs4gsex)
Foreign born parent	Coded 1 if at least one parent is foreign born, and 0 otherwise (bsdgborn)
Language at home	Coded 1 if student speak language of test at home, and 0 otherwise (bs4golan)
Books at home	Used bs4gbook. 5=1 to 10, 18=11 to 25, 64 = 26 to 100 , 151 = 101 to 200, 300 = over 200
Parent education	Coded 1if at least one parent completed post-secondary degree and 0 otherwise (bsdgedup)
Teacher Background	
Male teacher	Coded 1 if male, and 0 if female (bt4gsex)
Teacher's years of experience	Teachers' report on their years of experience (bt4gtaut)
Teacher's Degree in math or math education	Coded 1 if majored on math or math education and 0 otherwise (bt4mpsma and bt4mpsem)
Class Characteristics	
Class average number of books at home	Class average of students' books at home
Class proportion of parents with college degree	Class average of parent education
Class size	Teachers' report on class size (bt4mstud)
Different ability of students as limiting instruction	Teachers' report on the extent to which different academic ability of students limit how they teach the class. Coded 1 if "a lot" and 0 otherwise (bt4mli01)
Uninterested students as limiting instruction	Teachers' report on the extent to which uninterested students limit how they teach the class. Coded 1 if "a lot" and 0 otherwise (bt4mli04)
School Characteristics	
School size	Principals' report on school enrollment (bc4gtenr)
Student from affluent family>50%	Principals' report on students' economic backgrounds. Coded 1 if more than 50% of students are from economically affluent homes, and 0 otherwise (bc4gsbea).
School resource	Categorical variable derived by TIMSS using principal survey items. Coded 1 if school resource for math instruction is "high", and 0 if "low" or "medium" (bcdsrmi).
School climate	Categorical variable derived by TIMSS using principal survey items. Coded 1 if principal's perception of school climate is "high", and 0 if "low" or "medium" (bcdgppsc)
School type (Japan)	School types are categorized to public schools in large or very large cities, public schools in small cities or non-city area(reference category), and private schools(idstrate).
School rank (Singapore)	School rank based on students' performance (idstrati).

Note. All continuous variables are standardized to the country mean (for Japan and Singapore) when included as predictors in the multilevel models

Table A2.2 Correlation of Variables, for Japanese classes (n=140)

		Instructional Practice					Class Characteristics						School Characteristics					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Class average math score[a]	1.00																
2	Instructional Practice: Complex	0.17*	1.00															
3	Instructional Practice: Procedural	-0.06	0.27**	1.00														
4	Instructional Practice: Student-centered	0.01	0.13	0.07	1.00													
5	Instructional Practice: Teacher-centered	0.07	-0.16	0.01	-0.42***	1.00												
6	Class: Average number of books at home	0.60***	0.15	-0.10	-0.02	-0.08	1.00											
7	Class: Prop. of parents with college degree	0.52***	0.23**	-0.06	-0.05	0.05	0.47***	1.00										
8	Class: Size	0.30***	0.00	0.04	0.17*	0.20*	0.07	0.00	1.00									
9	Class: Different ability limits instruction	-0.02	0.06	-0.08	-0.05	0.19*	-0.04	-0.01	0.09	1.00								
10	Class: Uninterested students limits instruction	-0.11	-0.11	-0.09	-0.18*	0.22**	-0.10	-0.01	0.06	0.32***	1.00							
11	School: Size	0.21*	0.11	0.17*	0.09	0.24**	0.11	0.19*	0.54***	0.13	0.08	1.00						
12	School: More than 50% from affluent family	0.21*	0.08	-0.03	-0.04	-0.06	0.38***	0.32***	-0.04	-0.15	-0.07	0.10	1.00					
13	School: High resource	0.06	-0.05	0.08	0.00	0.04	0.13	0.04	-0.08	-0.10	-0.13	-0.11	0.08	1.00				
14	School: High climate	0.38***	0.10	0.01	-0.16	0.11	0.19*	0.14	0.14	-0.04	-0.09	0.05	0.07	0.19*	1.00			
15	School: Public school in large city	-0.03	0.00	-0.04	0.24**	0.04	-0.05	0.16	0.12	0.24**	-0.03	0.16	0.01	0.01	-0.02	1.00		
16	School: Public school in small city	-0.26**	0.00	0.04	-0.18*	-0.08	-0.15	-0.35***	-0.21*	-0.16	0.06	-0.21*	-0.13	-0.09	-0.12	-0.89***	1.00	
17	School: Private school	0.63***	0.00	-0.02	-0.12	0.10	0.44***	0.42***	0.21*	-0.15	-0.07	0.11	0.26**	0.19*	0.30***	-0.16	-0.31***	1.00

NOTE.—a. Average of the five class average math scores based on the five plausible values were used for simplicity.
*p<0.05, **p<0.01, ***p<0.001

Table A2.3 Correlation of Variables, for Singaporean classes (n=284)

		Instructional Practice					Class Characteristics					School Characteristics					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	Class average math score[a]	1.00															
2	Instructional Practice: Complex	0.23***	1.00														
3	Instructional Practice: Procedural	0.09	0.34***	1.00													
4	Instructional Practice: Student-centered	0.21***	0.05	-0.06	1.00												
5	Instructional Practice: Teacher-centered	0.14*	0.00	0.10	-0.44***	1.00											
6	Class: Average number of books at home	0.71***	0.19**	0.10	0.13*	0.04	1.00										
7	Class: Proportion of parents with college degree	0.54***	0.12*	0.09	0.04	0.15*	0.59***	1.00									
8	Class: Size	0.13*	-0.01	0.08	-0.01	0.09	-0.07	-0.10	1.00								
9	Class: Different ability of students limits instruction	-0.30***	-0.11	0.09	-0.11	-0.02	-0.20***	-0.22***	-0.10	1.00							
10	Class: Uninterested students limit instruction	-0.31***	-0.06	0.11	-0.18**	0.00	-0.23***	-0.23***	-0.02	0.39***	1.00						
11	School: Size	0.29***	0.05	-0.04	-0.06	0.13*	0.34***	0.37***	0.13*	-0.10	-0.14*	1.00					
12	School: More than 50% from affluent family	0.26***	0.09	0.06	0.03	0.09	0.25***	0.28***	-0.01	-0.08	-0.07	0.21***	1.00				
13	School: High resource	-0.05	-0.11	-0.04	0.07	0.09	-0.03	-0.03	-0.01	-0.01	-0.03	-0.01	-0.02	1.00			
14	School: High climate	0.31***	0.06	-0.05	-0.05	0.19**	0.31***	0.32***	0.04	-0.16**	-0.11	0.28***	0.20***	0.05	1.00		
15	School: Performance rank	0.63***	0.13*	-0.01	0.12	0.21***	0.64***	0.60***	-0.02	-0.17**	-0.17**	0.57***	0.39***	0.02	0.45***	1.00	

NOTE.—a. Average of the five class average math scores based on the five plausible values were used for simplicity.
*p<0.05, **p<0.01, ***p<0.001

55

Table A2.4 Means and coefficient of variations of instructional practices by country

Country	Complex Instruction		Procedural Instruction		Student-centered Instruction		Teacher-centered Instruction	
	Mean	Coef. of Var.	Mean	Coef. of Var.	Mean	Coef. of Var.	Mean	Coef. of Var.
ARM	2.51	21.38	2.58	36.75	1.58	49.52	3.44	32.50
AUS	2.44	24.47	2.56	22.71	2.40	55.74	2.65	38.35
BGR	2.88	19.02	3.37	15.32	1.69	52.18	2.86	44.50
BHR	2.79	23.74	3.05	18.83	1.20	50.27	3.72	32.42
BIH	2.84	19.02	2.87	21.02	1.50	54.25	3.96	36.27
BWA	2.78	21.05	3.01	20.98	2.10	54.62	2.32	39.14
COL	3.10	15.56	2.98	18.32	1.65	49.53	2.93	38.40
CYP	3.05	16.58	3.11	16.85	1.21	50.38	2.68	40.41
CZE	2.95	17.06	2.69	18.84	2.09	35.71	2.93	29.43
DZA	2.82	19.46	2.90	20.42	1.53	66.56	3.30	38.95
EGY	2.83	19.54	2.83	22.17	1.43	46.59	3.66	34.91
ENG	2.59	19.64	2.60	20.50	2.32	62.43	2.83	37.04
GEO	2.82	17.36	3.18	19.93	1.48	46.15	3.26	30.73
GHA	2.71	21.93	3.36	18.05	1.46	47.53	2.74	36.37
HKG	2.28	18.60	2.33	22.33	1.30	62.56	4.42	30.54
HUN	3.01	14.87	2.98	18.23	2.23	45.23	2.07	47.08
IDN	2.54	22.70	2.86	20.99	1.48	51.78	3.04	33.30
IRN	2.87	20.03	2.91	19.71	1.38	59.35	3.14	35.65
ISR	2.71	21.82	2.75	25.18	1.87	58.26	2.77	44.11
ITA	2.91	20.55	2.83	23.58	1.23	53.52	3.53	27.35
JOR	2.88	17.62	3.29	18.16	1.50	36.70	3.25	29.97
JPN	2.34	22.00	2.81	24.21	1.21	90.26	4.41	23.71
KOR	2.72	17.76	2.93	20.70	1.70	46.98	4.37	30.20
KWT	2.71	21.61	2.93	24.32	1.37	55.05	3.73	37.61
LBN	2.89	17.08	2.91	21.18	1.03	80.72	2.73	34.06
LTU	2.78	16.64	3.21	18.79	2.54	40.34	2.07	36.80
MAR	2.80	20.55	2.88	20.66	1.49	64.63	3.21	41.48
MLT	2.58	20.57	2.74	19.49	1.52	45.83	3.05	33.45
MNG	2.70	22.37	2.84	22.69	2.34	49.50	2.38	46.22
MYS	2.63	20.56	3.08	21.93	1.29	56.89	3.42	37.09
NOR	2.48	20.93	2.17	18.55	2.47	56.65	3.25	32.34
OMN	2.91	18.25	3.19	19.08	1.44	60.41	3.31	39.78
PSE	2.81	19.41	3.18	18.86	1.43	59.06	3.29	30.62
QAT	2.69	21.80	3.04	21.33	1.33	56.22	3.47	40.09
ROU	2.92	19.02	3.26	16.96	1.35	49.10	2.76	31.73
RUS	2.55	16.00	3.42	15.78	1.97	39.85	2.72	29.83
SAU	2.72	21.89	3.06	21.72	1.09	49.37	3.68	33.32
SCT	2.54	21.01	2.84	20.69	2.44	63.58	2.95	29.58
SGP	2.29	21.15	2.58	20.00	1.35	61.08	3.53	39.07
SLV	2.83	21.59	3.06	20.36	1.99	47.91	2.69	38.02
SRB	2.77	21.20	3.09	19.04	1.99	56.06	3.39	37.20
SVN	2.79	20.70	2.95	19.85	2.11	51.32	3.22	35.03
SWE	2.65	21.12	2.38	21.78	2.78	85.62	2.40	41.80
SYR	2.76	23.00	3.34	18.83	1.00	49.34	3.86	29.39
THA	2.85	21.28	3.07	21.00	1.20	55.43	3.59	35.17
TUN	2.67	21.23	3.05	22.10	1.61	74.11	3.03	46.07
TUR	2.89	17.85	3.06	20.03	1.28	63.68	3.43	36.78
TWN	2.40	19.42	2.75	19.67	0.74	64.32	5.07	24.50
UKR	2.73	16.82	3.39	15.11	1.81	38.75	3.12	30.03
USA	2.70	21.32	2.89	20.18	1.68	57.94	3.10	34.92
Total	2.73	19.92	2.94	20.47	1.64	55.18	3.21	35.47

Table A2.5 Decomposition of math achievement variance in the five East Asian countries

Variance	SGP	JPN	KOR	TWN	HKG
Between schools	1853.06	1360.97	2294.06	2588.31	5997.75
Between classrooms	4860.03	-	-	-	-
Within classrooms	2020.89	5747.24	5899.83	8088.63	3190.69
Variance at school level	21%	19%	28%	24%	65%
Variance at class level	56%	-	-	-	-
Variance at student level	23%	81%	72%	76%	35%

Chapter 3 Can standards-based policy tools raise complex instruction and student achievement? A cross-national examination

Abstract

Debate over Common Core State Standards (CCSS) in the U.S. context includes renewed claims that central policy can advance complex instruction and, in turn, boost achievement. Many countries already hold long experience with three elements of central policy – national curricular standards, instructional guides, and regional or national testing – but little is known about whether such regimes are associated with complex instruction or predictive of higher student performance. This paper examines these relationships across 48 nations and 9,154 classrooms that participated in the 2007 Trends in International Mathematics and Science Study. Nations that pressed uniform instructional guides tended to host classrooms with less complex instruction; those with greater emphasis on national testing manifest lower achievement. Most importantly, the positive effects of complex instruction on student achievement were stronger when national curriculum standards and more specified instructional guides were in place. Even when the degree of complex instruction appears to be exogenous to contemporary policy, the presence of common curricular standards and pedagogical guides appears to moderate the effectiveness of complex instruction.

3.1 Introduction

Many governments deploy central policy tools to encourage complex instruction, aiming to equip students with the skills necessary to compete in a post-industrial global economy (Davies and Guppy 1997; Mundy 2005). Debate over the Common Core State Standards (CCSS) offers one contemporary example from the U.S. context, following earlier federal efforts that encouraged didactic pedagogy. Yet other nations have long arranged national curricular standards, teaching guides, and standardized testing to encourage complex instructional practices. Or, teaching traditions in some societies may sustain complex instruction exogenous to the rise of centralized curricular and assessment policies. This paper – drawing on data from 48 nations – examines the association between centralized policies and classroom complexity, then estimates how each may shape student achievement.

The key idea behind the standards-based policy is that setting clear and rigorous learning standards will assist in effective implementation of complex practice in classrooms and thereby boost student achievement. In the U.S., moving past the former standard-based accountability policy, No Child Left Behind (NCLB) and toward CCSS, there is much discussion about whether and how the new policy will achieve desired results as it unfolds. It is debated whether a national standard alone can achieve the desired goal. Other controversies involve how other policy tools, such as standardized assessments or specified instructional guides, work in conjunction with the national standard.

The U.S. is not the only nation to debate the relationship between assertive centralized standards and the level of complexity or routinization in classrooms. For example, Germany has introduced national educational standards in 2003 and 2004 (Ertl 2006), and countries such as Australia (Klenowski and Wyatt-Smith 2010) and Norway (Møller and Skedsmo 2013) are developing national curriculum and achievement standards. Debates within these countries revolve around the appropriate combination of the three key policy components—national curricular standards, instructional guides, and test-based control—and how each of the elements affect complex teaching and learning. For example, there are diverging arguments about whether specified instructional guides will assist successful implementation of the standards or constrain complex instruction in classrooms. Another discussion is around whether supplementing the national standards with test-based accountability will incentivize schools and teachers to spur higher-order learning or demoralize actors at stake to game the system in exchange for student learning.

These controversies on the success or failure of the policy are based on the presence or absence of and the relative emphasis placed upon each of the policy instruments; national standards, instructional guides, and test-based control. Currently, research that can inform the discussion is limited because studies in the U.S. setting have not been able to study the independent impact of these policy components. But nations have long implemented varying mixes of centralized policies hoping to shape classrooms that efficiently deliver 'known knowledge' or to advance a wider range of cognitive skills. We can look retrospectively, taking advantage of natural variation among nations, to better understand these empirical relationships.

Figure 3.1 illustrates our basic conceptual framework. Next, we turn to a pair of interrelated literatures: what's known about central curricular policy and classroom complexity, and how policy intensity may affect student achievement. This leads to our research questions, emphasizing that

how classroom organization shapes learning may be conditioned by central policy. Our argument takes into account that the complexity of classroom instruction may be historically exogenous to the policy context.

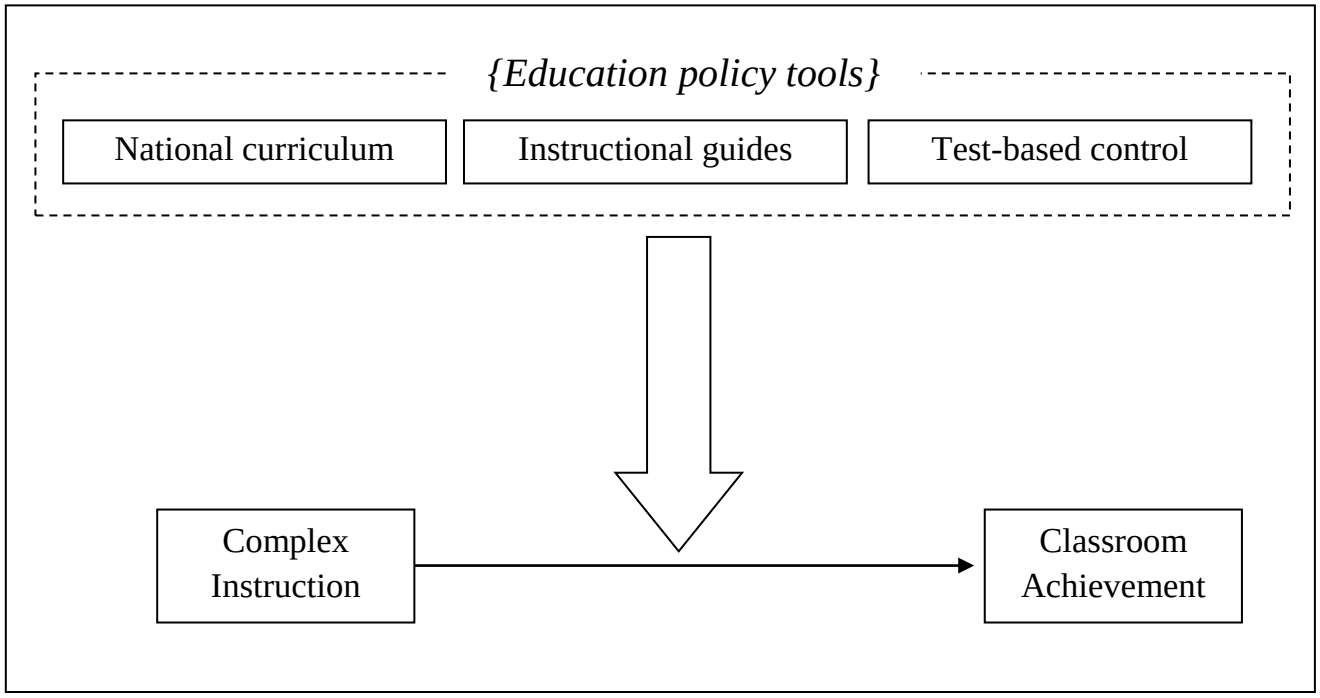


Figure 3.1 Moderating effects of the three policy tools on complex classroom instruction

3.2 Literature review

The effects of standards-based policy on classroom complex instruction

We have learned much over the past decade on how standards-based accountability can shape the academic knowledge covered in classrooms. But less evidence exists on how centralized policies shape the social organization of classrooms, whether moving toward routinization or greater complexity of instructional techniques. Some small scale studies show that the policies lead to more emphasis on didactic pedagogy (Diamond and Spillane 2004; Sandholtz, Ogawa, and Scribner 2004; Booher-Jennings 2005). Au's (2007) metasynthesis of 49 qualitative studies of the curricular impact of high-states testing suggests that when accountability policy influenced classroom instruction, it led to more didactic as opposed to complex approaches.

On the other hand, a few scholars have found that high-stakes accountability has resulted in more teacher-reported use of conceptual, problem-solving approaches in classrooms (Hamilton et al. 2003; Desimone 2013). International classroom studies also found that specified mandated curriculum material in centralized countries did not necessarily constrain teachers' autonomy and provided opportunities for teacher learning and collaboration (Stigler and Hiebert 1999; Wang and Paine

2003). Still others found no change in instruction (Wong et al. 2003), or the relative emphasis on complex or procedural instruction depended on teacher's skill and experience (Achinstein, Ogawa, and Speiglman 2004).

Multi-state studies in the U.S. have utilized the variation among state policies under NCLB to relate different policy attributes to teachers' instructional choices (Clune 1998; Desimone 2002; Desimone, Smith, Hayes et al. 2005; Polikoff 2012). States with tighter test-based accountability tend to have teachers that reportedly align content covered in their classrooms (Polikoff 2012). States with clear and specific instructional guidance and strong rewards and sanction system tend to improve procedural knowledge, but bring no change in problem solving or conceptual understanding (Desimone, Smith, Hayes et al. 2005). It could be speculated from this finding that specificity of instructional guides and the degree of test-based control may affect the level of complex instruction in classrooms. However, variability of policy variables may be limited in these studies because state-level policies are couched under a national educational regime. A cross-national comparison would allow a fuller scope of policy variation, thereby better informing the discussion around CCSS.

Several cross-national studies examined the effect of central educational policy on the coherence of teacher classroom practices. Several features of national curricular policies, such as centralization of curriculum control (Stevenson and Baker 1991) or specificity of instructional guides (Schmidt and Prawat 2006), have been shown to relate to coherence of instructional content taught in classrooms. These studies mainly focused on policy effects on content, and not pedagogy. But Schmidt and Prawat's study importantly distinguished the centralization of curricular standards and specificity of instructional guides as two discrete policy tools, which may vary independently and affect the social organization of classroom practices. The finding indicates that it is the specificity of curriculum and not the where the final authority resides (national or state or local level) that affect classroom practice. In a similar vein, this study examines three aspects of centralized policy efforts and their impact on the effective implementation of complex instruction in classrooms.

The effects of standards-based policy on student achievement

Research in the U.S. context tends to show that centralized curricular standards and higher stakes testing can raise achievement, at least learning that's gauged on standardized tests (Hout and Elliott 2011). This work leaves out how the social organization of classrooms may moderate or mediate the influence of centralized policy tools. And findings remain a bit inconsistent with regard to policy effects on achievement (Dee and Jacob 2011; Grodsky, Warren, and Kalogrides 2009; Lee and Reeves 2012). It's almost impossible to take into account a variety of possible confounding factors that co-occur with change (or cross-state variability) in policy regimes.

Several cross-state studies exploit interstate variation in accountability policy to explain state achievement differences. Studies diverge in whether it is the existence of a basic accountability system (Hanushek and Raymond 2004) or the strength the state accountability system (Carnoy and Loeb 2002) that explain the achievement differences. Cross-national studies used the broader variation in education systems and academic achievement to associate educational policy tools and international test scores. Studies report that countries with external exit exams (Bishop 1997; Jürges, Schneider, and Büchel 2005) and standardized examination (Fuchs and Woessmann 2008) show higher achievement.

Previous research informs us on the direct effects of the standards-based policy on classroom practice or student achievement. But the research on achievement effect of education policy treat classroom instruction as a black box and the studies on instructional effect of policies rarely take achievement into account. There is little research that investigates the interrelationship between policy attributes, classroom practice, and achievement. But as current educational policies, such as CCSS, are aiming to advance student learning through complex and conceptual instructional strategies, it is crucial to understand how different aspects of the policy interrelate with classroom practice and achievement.

This study fills the gap in the literature and contributes to the discussion on possible effects of educational policies taking form at the national level. First, by taking a cross-national approach, this study uses the natural variation in educational policy tools to assess the impact of the policy tools. Given that the inferences that we intend to make involve national level characteristics, such as the effect of a common national standard or national assessment system, a cross-national comparison would be an appropriate level of inquiry. Second, while previous cross-national studies mainly attended to national assessment systems, this study includes the presence of national curriculum and specificity of instructional guides in the analyses. If the presence or importance of each policy tools vary from country to country, we can find how each of the policy tools influences complex teaching and learning in classrooms. Third, the study unpacks the instructional effect of policies by attending to the moderating effect of education policy on the relationship between complex instruction and achievement. Because the current educational policies aim for advancing student learning through complex instruction, it is important to understand whether and how the policy tools foster effective implementation of complex instruction.

3.3 Research questions and analytic strategy

In order to inform the discussion on the effects of standards-based reforms unfolding in many developed countries, this paper examines how the key aspects of the policy interrelate with complex teaching and classroom achievement. More specifically, this paper asks how the three aspects of standards-based policies—the presence of national curricular standards, the specificity of instructional guides, and the intensity of test-based control—directly and indirectly affect complex instructional practice and achievement in classrooms. The specific research questions are as follows:

1. How does each component of the standards-based policy affect the level of complex instructional practices?

2. How does each component of the standards-based policy affect mathematics achievement?

3. How does each component of the standards-based policy moderate the association between complex instructional practice and achievement?

This study uses the Trends in International Mathematics and Science Study (TIMSS) 2007 data set to address the questions. Using the TIMSS data allowed for examining a large number of classrooms across countries with varying arrangements of the three policy components. By addressing the first two questions, this study extends previous U.S.-based studies on standards-based accountability

policy effects by using an internationally comparative framework. Then, this study examines the moderating effects of policy aspects on the relationship between complex instruction and achievement using large scale cross-national data for the first time.

3.4 Methods

Data

The TIMSS data assessed mathematics and science achievement of the eighth graders in secondary schools and collected extensive information on students' family and school experiences (student questionnaire), various school and classroom characteristics (teacher and school questionnaire), and various aspects of each countries' educational system. Based on a two-stage stratified cluster design, schools were selected with probability proportional to size. Then, one or two intact classrooms for the eighth grade were chosen randomly within each selected school.

As one of the key associations this study examines is the relationship between complex instructional practices and classroom achievement, it is essential to identify the learning units that link teachers exclusively to the classroom or group of students they taught. In cases where the sampled classroom was taught by one teacher, the classroom was treated as the learning unit. But there were cases where the sampled classrooms were taught by more than one teacher. In the majority of such cases, students in the sampled classroom were grouped and exposed to different teachers. There were also a few cases in which all students in the sample classroom were exposed to multiple teachers. For the former cases, I treat the teacher as the learning unit and link each teacher to the group of students she taught. As it is hard to discern each teacher's effect on the classroom achievement for the latter cases, I dropped such cases, which amounted to 124 sampled classrooms. I use the term "classroom" to indicate the operationalized learning unit.

For the analysis, I use the surveys completed by students' mathematics teachers, as well as student mathematics assessment and background questionnaire data aggregated to the class level. Policy measures are derived from the curriculum survey. Two countries, in which key policy and classroom practice measures were not collected, are dropped in the analysis.[30] After deleting cases with missing information, I ended up with a total of 9,154 classrooms across 48 countries.

Measures

Class achievement

The dependent variable, classroom average mathematics achievement, is derived by aggregating students' achievement scores to the classroom level. The TIMSS 2007 mathematics assessments cover the appropriate subject matter domains within mathematics and the cognitive domains that correspond to students' thinking processes when engaging with the mathematics contents (Mullis et al., 2008; Olson, Martin, & Mullis, 2008). The cognitive domains ranges from basic skills such as

[30] All TIMSS curriculum policy measures were missing for El Salvador. Question on teacher's educational degree was not administered in Hungary. As a teacher's educational degree is an important measure for teacher quality and is conventionally controlled when analyzing effect of instructional practice (cite), I decided to drop Hungary from the analysis. A sensitivity analysis was conducted (not reported) by running the models without the teacher degree measure and including Hungary. The results were largely the same.

knowing facts and procedures to more higher-order thinking skills such as dealing with unfamiliar problems and complex contexts (Mullis et al., 2005). Therefore, the TIMSS assessment could be regarded as a comprehensive measure of students' mathematics achievement.

TIMSS conceptualizes student mathematics achievement as a latent construct measured with uncertainty and provides five separate imputed scores. Any analysis that involves student achievement should be replicated five times, and the results should be aggregated to obtain the final estimate as well as the imputation error (Foy & Olson, 2009). Thus, I aggregate each of the five plausible values of students' mathematics scores from the student level to class level and use all five class average math achievement values for the analysis in this paper. See Appendix A3.1 for complete variable definitions and descriptions.

Complex instruction variable

In forming the key independent variable at the classroom level, I derive a ratio of two instructional practice variables: complex instruction and procedural instruction. Items that asked how often teachers ask students to relate the lessons to their daily lives, solve problems with no obvious solution, explain their answers, and decide on their own procedures in solving complex problems were categorized as complex instruction (alpha=0.664). In the literature, such practices are also termed higher-order instruction, authentic pedagogy, reform-oriented math instruction or conceptual instruction (e.g., Carpenter et al. 1989; Raudenbush, Rowan, and Cheong 1993; Newmann, Marks, and Gamoran 1996; Desimone, Smith, Hayes et al. 2005). Items that asked how often teachers ask students to solve routine problems, memorize formulas, and practice computation were categorized as procedural instruction (alpha=0.526). In indicating such practices, scholars use different terms, including basic skills instruction and traditional instruction (Shouse 2001; Hamilton and Martinez 2007).

The ratio of the two variables measures the relative emphasis teachers place on complex versus procedural instruction in their math classrooms. Prior research has also used ratio measures to capture the teachers' relative emphasis on instructional approaches (Desimone, Smith, Baker et al. 2005; Minor et al. 2015). I choose to use this measure rather than a single measure of complex instruction, because educational policies in the U.S. and other developed countries concern has been to bring in more complex instruction in place of the currently overemphasized didactic and procedural approaches. This paper examines whether national curriculum, instructional guides, and test-based control encourage teachers to place more weight on complex instruction relative to procedural instruction.

Class-level control variables

I control for the class average socioeconomic background by including the class average number of books in students' homes. Teachers' years of teaching experience and teachers' educational degree in math or math education are included as proxies of teachers' content knowledge and pedagogical content knowledge. Since class size is known to affect teachers' work, overall classroom environment and student achievement, I control for teachers' report of class size. Local and school contexts, including instructional resources, collegiality among teachers and school climate supportive of teachers work, are also important factors that condition teaching and learning in classrooms. To control for such factors, I include teachers' perception of sufficiency of material resources, teacher collaboration, and school climate in the models.

Policy tools

National curriculum standards, instructional guides, and test-based control are the three key policy instruments used across countries to increase complex teaching and learning in classrooms. I use the TIMSS 8th Grade Mathematics Curriculum Questionnaire completed by National Research Coordinators to derive the policy measures. For the first measure, national curriculum, I use the item that indicates whether a country has a common curriculum at the national level. The second measure, specificity of instructional guide, is derived from items that asked whether textbook, instructional guide, and recommended activities are made available, and whether instructional methods and materials are prescribed in the curriculum. Countries are grouped according to the degree of specificity of instructional guide: 0 (low), 1 (medium), and 2 (high). The third measure, degree of test-based control, is derived from items that asked whether a country has a high-stakes assessment, whether student test results are used to evaluate curriculum implementation, and whether percentage of students reaching defined goals are prescribed in the curriculum. Countries are distinguished into three groups according to the degree of test-based control: 0 (low), 1 (medium), and 2 (high).

Country control variables

At the country level, I control for the level of economic development of each country by including the GDP per capita.

Analytic approach

To take into account the data structure that classrooms are nested within schools that are nested within countries, I employ a three-level hierarchical linear model (HLM) technique with the pooled data of 48 countries. In the class-level equation, the class average mathematics score for class i in school j in country k is predicted as follows:

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(complex\ instruction)_{ijk} + \sum_{p=2}^{8} \pi_{pjk}(class - level\ control)_{pijk} + \varepsilon_{ijk}$$

where π_{0jk} represents the average class math scores of school *j* in country *k*, and ε_{ijk} is a class-level residual. π_{1jk} is the slope of complex instruction affecting class average math score in school *j* in country *k*. The effects of other class-level control variables, are represented through π_{2jk} to π_{8jk}. All continuous variables are centered about the grand mean and standardized to have a mean of 0 and a standard deviation of 1.

In the school-level equations, the coefficients of the class-level equation become dependent variables. The school-level equations are:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$

$$\pi_{1jk} = \beta_{10k}$$

$$\pi_{pjk} = \beta_{p0k}$$

where β_{00k} represents the average class math score in country *k*, and r_{0jk} is a school-level residual. Class-level slopes are set not to vary at the school level. This places an assumption that the effects of class-level covariates, π_{1jk} to π_{pjk}, are set constant within country *k*.

Lastly, school-level coefficients become dependent variables in country-level equations as follows:

$$\beta_{00k} = \gamma_{000} + \sum_{q=1}^{3} \gamma_{00q}(policy\ tools)_{qk} + \gamma_{004}(ln\ GDP)_k + u_{00k}$$

$$\beta_{10k} = \gamma_{100} + \sum_{q=1}^{3} \gamma_{10q}(policy\ tools)_{qk} + u_{10k}$$

$$\beta_{p0k} = \gamma_{p00}$$

In order to examine how policy tool variables at the country level shape the relationship between classroom complex instruction and achievement, I model the slope of complex instruction, β_{10k}, as a function of the three country-level policy tool variables. γ_{100} indicates the slope of complex instruction in a country without a national curriculum, with a low level of specificity of instructional guides, and with a low level of test-based controls (i.e., when all three policy variables equal zero). Representing the impact of a policy tool variable on the complex instruction slope, γ_{10q} , indicates the extent to which a policy tool variable moderates the relationship between complex instruction and class math achievement. For the national curriculum variable, it indicates how much the complex instruction slope differs in a country with a national curriculum from that in a country without a national curriculum controlling for the other two policy tool variables. For the instructional guide variable, it indicates the difference in slope of complex instruction between countries with a unit difference in specificity of instructional guides. The test-based control model follows the same scheme of interpretation.

Each country's average class math score β_{00k} is also modeled to vary across countries as a function of the policy tools and the national logarithm of GDP per capita, which is grand mean centered and standardized. γ_{000} indicates the average class math score in a country without a national curriculum, with a low level of specificity of instructional guides, with a low level of test-based controls, and with a grand mean of the GDP measure. γ_{00q} indicates the change in the average class math score with a unit increase of a certain policy variable controlling for the other two policy variables. Specific meanings attached to γ_{000} and γ_{00q} in the models will be described in the result section.

3.5 Results

Presence of cross-national differences in the arrangement of policy tools

Table 3.1 presents the policy arrangements of national curriculum, instructional guides, and test-based control, as well as country means and standard deviations of complex instruction and class math scores in all 48 countries. There are substantial variations in ways countries arrange the three

policy tools. Having a national curriculum does not necessarily entail more instructional guide or stronger test-based control. For example, Sweden and England have national curriculums but have low levels of instructional guides, and Czech Republic, Georgia, and Taiwan have national curriculums but have low levels of test-based controls. On the other hand, countries like Hong Kong, Turkey, and Singapore have high levels of instructional guides.

Table 3.1 Policy tools, complex instruction, and mathematics score across 48 countries

Country	# of classes (unweighted)	Complex instruction		Class math score		Policy tools		
		Mean	SD	Mean	SD	National curriculum	Instructional guides	Test-based control
Norway	217	1.15	0.25	468.89	26.62	1	0	1
Armenia	237	1.13	0.50	499.71	45.78	1	2	1
Czech Republic	189	1.13	0.28	501.21	41.92	1	0	0
Sweden	368	1.12	0.26	477.53	51.62	1	0	1
Italy	279	1.08	0.28	475.81	44.24	1	0	1
Colombia	113	1.06	0.25	378.15	59.24	0	0	1
England	198	1.04	0.26	496.16	80.45	1	0	1
Bosnia and Herzegovina	142	1.03	0.28	457.90	38.76	0	0	1
Lebanon	135	1.03	0.26	448.73	51.70	1	2	0
Egypt	190	1.03	0.28	397.94	53.59	1	0	1
Algeria	98	1.02	0.30	387.94	22.35	1	1	0
Hong Kong	107	1.02	0.27	554.02	80.46	1	2	1
Israel	164	1.02	0.29	454.87	78.33	1	0	1
Cyprus	204	1.01	0.22	463.63	28.29	1	0	0
Iran, Islamic Rep. of	191	1.01	0.23	398.60	50.41	0	1	1
Morocco	76	0.99	0.24	382.13	37.20	1	1	1
Mongolia	116	0.98	0.25	437.51	47.03	1	1	1
Australia	208	0.97	0.24	490.30	64.79	0	0	1
Malta	215	0.97	0.23	477.23	80.56	1	0	1
Slovenia	677	0.97	0.25	491.62	59.87	1	0	2
Botswana	121	0.96	0.26	369.70	43.88	1	2	1
Korea	211	0.96	0.23	596.12	56.10	1	1	1
Turkey	125	0.96	0.21	432.89	71.57	1	2	1
United States	410	0.96	0.25	511.12	61.27	0	1	0
Oman	144	0.95	0.23	375.52	47.17	1	0	1
Thailand	147	0.95	0.21	431.91	64.49	1	0	1
Kuwait	86	0.94	0.24	358.49	33.32	1	1	2
Georgia	161	0.93	0.24	417.67	61.29	1	0	0
Bahrain	159	0.92	0.24	401.14	37.06	0	1	1
Qatar	214	0.92	0.26	309.36	55.16	1	2	1
Serbia	202	0.92	0.31	483.71	43.20	1	1	1
Palestinian Nat'l Auth.	127	0.91	0.21	369.52	49.07	1	1	1
Romania	228	0.91	0.20	456.41	61.84	1	1	0
Tunisia	125	0.91	0.27	418.71	27.54	1	2	1
Scotland	262	0.91	0.21	475.64	70.72	1	0	2
Chinese Taipei	134	0.90	0.20	590.64	55.90	1	2	0
Indonesia	131	0.90	0.22	394.66	62.59	0	1	1
Japan	127	0.90	0.28	564.62	39.40	1	2	0
Saudi Arabia	130	0.90	0.25	330.95	36.09	1	1	1
Singapore	312	0.90	0.21	587.32	83.43	1	2	1
Jordan	171	0.88	0.19	431.85	59.72	1	2	1
Lithuania	231	0.88	0.20	498.21	45.66	1	1	1
Malaysia	160	0.88	0.20	474.52	64.81	1	1	1
Bulgaria	214	0.87	0.20	469.58	67.79	1	0	0
Ghana	140	0.84	0.20	309.01	60.73	1	2	1
Syrian Arab Republic	125	0.84	0.21	396.53	53.33	1	0	2
Ukraine	180	0.83	0.17	457.79	50.02	1	2	1
Russian Federation	253	0.78	0.23	506.15	55.43	1	0	1

Note. Countries are sorted by country average complex instruction index

Correlations between policy tools, complex instruction, and math achievement

Before testing the moderating effects of the three policy tools on complex instruction slopes, I examine the bivariate relationships among the key variables through correlation analyses (Table 3.2). Figures 3.2 to 3.4 assist in detecting any systematic relationship between the three policy tool variables and complex instruction. Table 3.2 shows that having national curriculums or stronger test-based controls are not systematically associated with countries' average level of complex instruction (Figures 3.2 and 3.4). On the other hand, the correlation between instructional guide and complex instruction is negative and marginally significant ($r = -0.28$; $p = 0.05$). As can be seen in Figure 3.3, countries with more prescribed instructional guides tend to have lower levels of complex instruction. However, the patterns are not always consistent. For example, there are countries such as Armenia, Lebanon, and Hong Kong that have highly prescribed instructional guides but also have higher level of average complex instruction. Interestingly, countries with higher test-based control tend to have lower achievement, although the correlation is not significant. Countries with high GDP level tend to have more complex instruction in classrooms ($r = 0.28$; $p = 0.06$), as well as higher math achievement ($r=0.43$; $p<0.01$). Again, the non-significant correlation among the three policy tool variables shows that these policy tools are independent of each other and that countries adopt and arrange the three policy tools in different ways.

Table 3.2 Correlation matrix for country-level variables

		1	2	3	4	5	6
1	Mean math score	1.00					
2	Mean complex instruction	0.11	1.00				
3	National curriculum	0.10	-0.10	1.00			
4	Instructional guide	0.01	-0.28*	0.14	1.00		
5	Test-based control	-0.23	-0.12	0.01	-0.14	1.00	
6	GDP(log)	0.43***	0.28*	0.03	-0.15	0.02	1.00

Note. Correlations are based on 48 countries.
***p<0.01, **p<0.05, *p<0.10

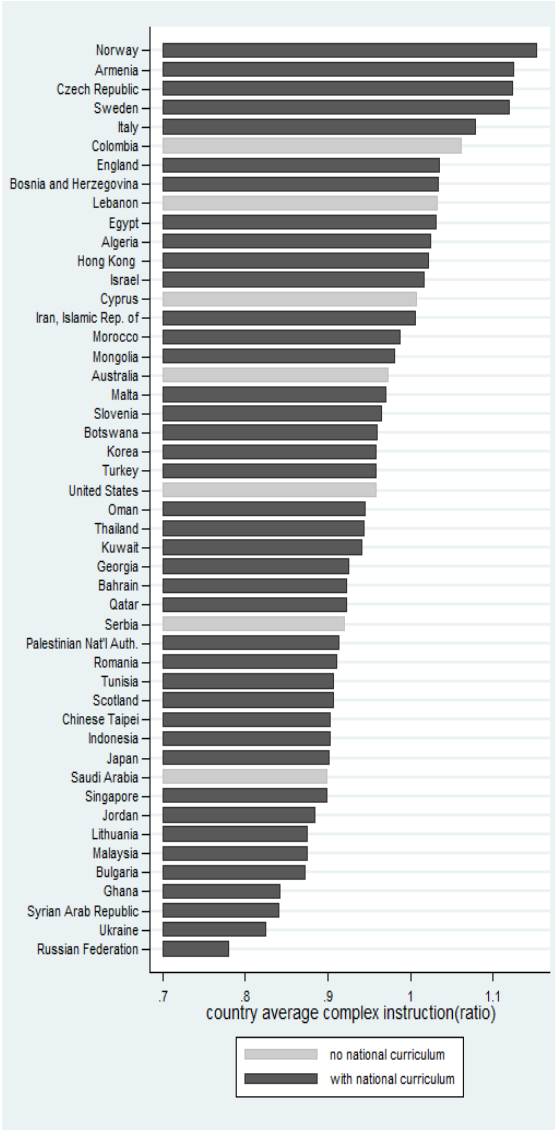


Figure 3.2 Country average complex instruction and national curriculum

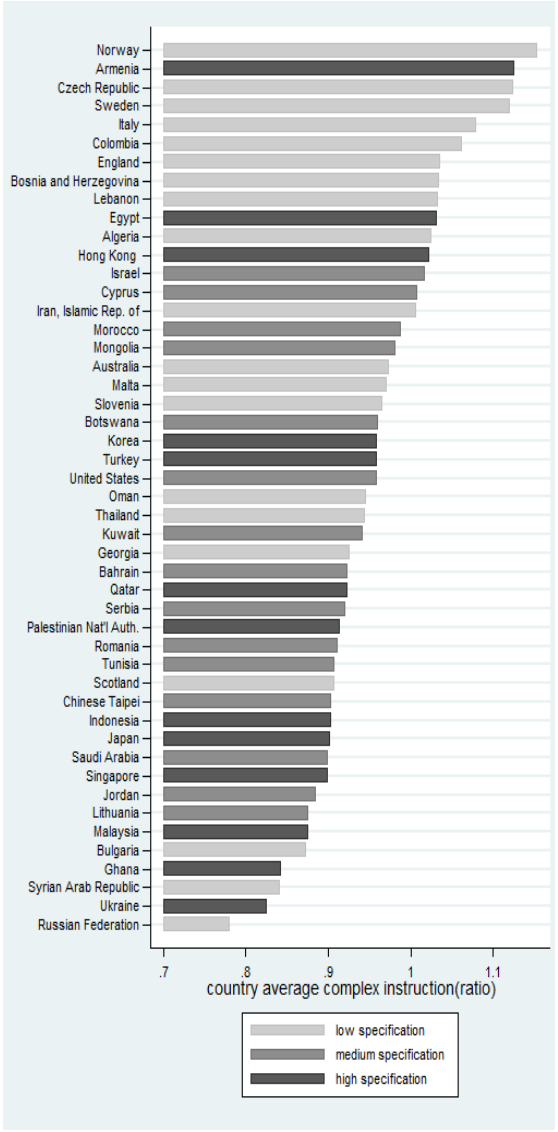


Figure 3.3 Country average complex instruction and degree of specification of instructional guides

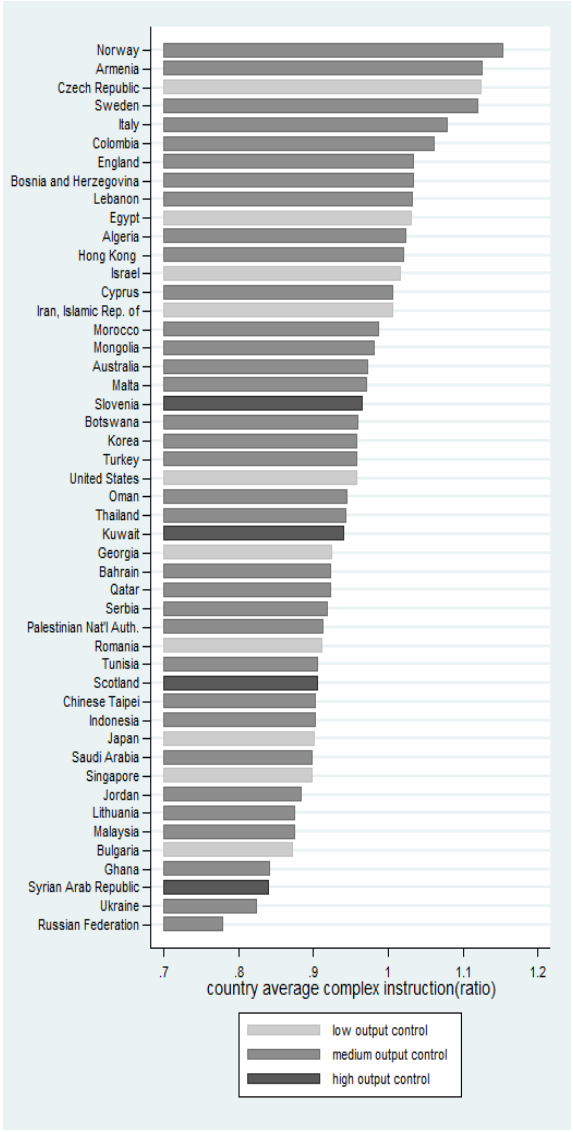


Figure 3.4 Country average complex instruction and degree of test-based control

The moderating effect of policy tools on the association between complex instruction and academic achievement

Table 3.3 presents the results of the three-level multilevel models as specified in the method section. Because the focus of this paper is on the impact of policy tools on the slope of complex instruction at the country-level, coefficients of classroom-level control variables are not shown. In the national curriculum model, the estimated slope (-3.6) of complex instruction indicates that complex instruction is not significantly associated with classroom achievement in countries with no national curriculum. The interaction term of national curriculum and complex instruction is significantly positive. The coefficient 6.31 indicates that in countries that have national curriculum the complex instruction slope is estimated as 2.71 (-3.6 +6.31), which is significantly positive (Figure 3.5).

Because the variable of instructional guide is scaled as 0/1/2, the slope of complex instruction indicates the association between complex instruction and achievement in counties that have a low level (0) of instructional guides. The coefficient of 2.61 means that in countries with a medium level of instructional guides the slope of complex instruction is 2.18 (-0.43+2.61). The slope in countries with high level of instructional guides is 4.79 (-0.43+2.61×2), and is significant at p=0.05 (Figure 3.5). In other words, the effect of complex instruction tends to be more positive in countries with more specified instructional guides.

Table 3.3 Effects of educational policy tools on the slope of complex instruction

	National Curriculum		**Instructional Guide**		**Test-based Control**		**National Curriculum, Instructional Guide, and Test-based Control**	
Class-level Equation								
Intercept	432.11***	(14.66)	436.77***	(8.59)	463.47***	(12.30)	449.12***	(19.31)
Complex instruction	-3.6	(2.76)	-0.43	(1.64)	0.09	(2.72)	-7.15*	(3.73)
	(Effects of other class-level control variables are not shown)							
Country-level Equation								
Effects on the slope of complex instruction								
National Curriculum	6.31**	(3.01)					5.44*	(3.04)
Instructional Guide			2.61*	(1.44)			2.5*	(1.29)
Output Control					1.95	(2.69)	2.5	(2.55)
Effects on the intercept								
National Curriculum	12.63	(16.26)					11.6	(14.63)
Instructional Guide			7.17	(12.26)			4.37	(12.04)
Output Control					-23.53*	(12.71)	-22.7*	(12.96)
GDP(log)	15.89*	(9.22)	16.88	(10.27)	16.26*	(9.00)	16.68*	(9.95)

***p<0.01, **p<0.05, *p<0.10

In the test-based control model, the moderating effect of test-based control on complex instruction slope, estimated as 1.95, was positive but not statistically significant. This means that the effect of complex instruction on achievement is not significantly different among countries with different levels of test-based control. However, it is noteworthy that countries with more test-based control tend to have significantly lower overall math achievement.

The last column of Table 3.3 presents a model with all policy tool variables and interaction terms. The interaction term of national curriculum and complex instruction has decreased in size (from 6.31 to 5.44) and significance. Overall, the direct and moderating effects of each of the policy tools remain significant even after controlling for the other policy aspects.

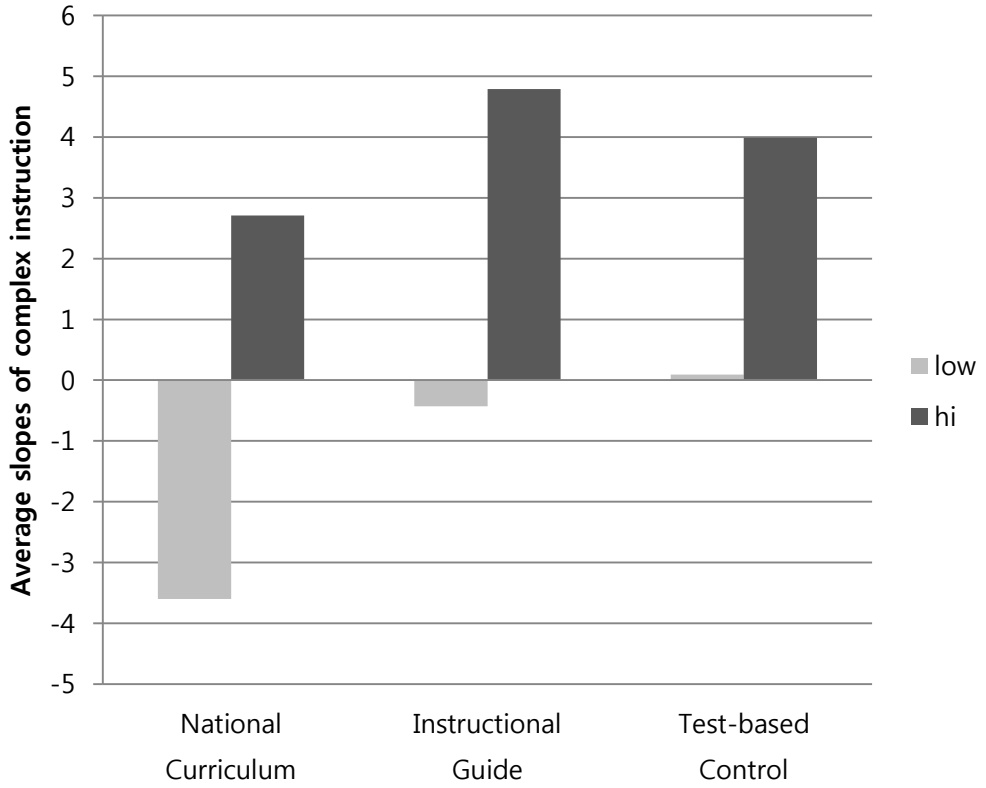


Figure 3.5 The average slopes of complex instruction among countries with different levels of policy tools

Table 3.4 presents variance components associated with the multilevel models to show the extent to which each policy tool variable explains country variances in the class math score intercept and complex instruction slope. Model 1 shows the between-country variances in intercepts and slopes when only control variables are included. Model 2 adds the national curriculum variable and the interaction between national curriculum and complex instruction. Including the variables reduces the slope variance by 8.8 percent. Models 3 and 4 show that adding instructional guide and test-based

control variables reduces the country variances in complex instruction slopes by 7.69 and 1.98 percents, respectively. Regarding the between-country variance for math scores, adding the national curriculum or instructional guide variables made little difference to the between-country variance in math score intercepts, while adding test-based control variables reduced the between-country variance of math scores by 4.7 percent.

Table 3.4 Variances of intercept and complex instruction slope estimated from the multilevel models

	Model 1 Only control variables	Model 2 National Curriculum	Model 3 Instructional Guide	Model 4 Test-based Control
Intercept Variance (Country)	3203.5	3184.3	3171.6	3053.4
Intercept Variance Explained (%)		0.60	1.00	4.69
Slope Variance (Country)	56.3	51.3	51.9	55.1
Slope Variance Explained (%)		8.79	7.69	1.98

3.6 Discussion

Moving past the former standard-based accountability policy, NCLB and toward CCSS, there is much discussion about whether and how the new policy will advance complex instruction and student achievement. The theory of action CCSS posits that improved teaching and learning will result from setting common high-quality standards, providing a system of supports to help teachers teach to the standards, and establishing accountability mechanisms through assessments. One of the key questions begging to be answered is how the three key components of CCSS—national standards, instructional guides, and test-based control—interrelate with complex teaching and classroom achievement. This study uses the natural cross-national variations in the three policy tools to inform this question. More specifically, this study examines how national curriculum, instructional guides, and test-based controls moderate the effect of complex instruction on classroom achievement.

The findings suggest an emerging story about how the three key components of standard-based reform policy differentially affect the association between complex instruction and classroom achievement. While national curriculum and test-based control were not directly associated with the level of complex instruction in classrooms, countries with more specified instructional guides were less likely to have complex instruction. And countries with stronger test-based control tend to have overall lower achievement. Most interestingly, the effects of complex instruction on student achievement tend to be stronger when national curriculum and more specified instructional guides were in place.

Lending explanation to previous cross-state research that associates specificity of state policies and student achievement domains such as procedural and complex knowledge (Desimone, Smith, Hayes et al. 2005), the this study shows that in countries with highly specified instructional guides, teachers are less likely to employ complex instruction. Consistent with the view that regard teaching as improvisational (Sawyer 2004), teachers in countries with less specified instructional guides may have used professional autonomy to experiment with different instructional approaches, including

complex instruction. It may be that highly prescriptive instructional guides are more effective in advancing procedural instruction. Still, there are countries such as Armenia, Lebanon, and Hong Kong that have highly specified instructional guides and have high level of complex instruction, Countries such as Russia, Syria, and Bulgaria have less specified instructional guides and yet have low level of complex instruction. Taking a closer look at these cases can help us understand the causal links between specificity of instructional guides and complex instruction in classrooms.

In contrast to previous research, the result shows a negative association between the degree of test-based control and overall achievement. Previous cross-national research reports positive effects of high-stakes testing on student achievement (Bishop 1997; Woessmann 2003; Jürges, Schneider, and Büchel 2005). This may be due to the difference in how test-based policy was measured. Previous studies have used indicators such as presence of external exit exams (Bishop 1997; Jürges, Schneider, and Büchel 2005), while what this study measures captures not only the presence of high-stakes tests, but the amount of test-based control pressure. This measure is closer to the accountability index used in U.S. cross-state studies, which reported positive effect of accountability pressure on achievement. The different direction of association may be a result of using a fuller range of test-based control in this study. It may be that test-based control increases achievement when moderately exercised, but has an adverse effect on achievement when used excessively. If this is the case, we would need to know about the mechanism through which test-based accountability pressure inhibits student achievement. Another fruitful next step would be to examine what might be the right level of test-based control that can incentivize people without placing unnecessary burden.

The most novel finding of this paper is the moderating effects of national curriculum and specified instructional guides on the slope of complex instruction. This may be reflecting the educative role curriculum materials can play in teachers' learning (Ball and Cohen 1996). Policy implementation scholars posited that policies are more likely to be successfully implemented when communicated through concrete, content-specific, and instructionally usable instruments, than when ambiguously stated (Ball and Cohen 1996; Cohen, Moffitt, and Goldin 2007). The moderating effects may also be in line with classroom studies that found that specified mandated curriculum material in centralized countries not necessarily limited teachers' autonomy, but provided shared grounds for discussing teaching with other teachers (Stigler and Hiebert 1999; Wang and Paine 2003).

The findings should be interpreted in the context of several limitation of this study. First, the study relies on teachers' report on their teaching practice, which may be less desirable than observations measures. Another validity concern may have to do with using teachers' self-reported practice. But as research on surveys shows, anonymous surveys are less prone to social desirability bias than interviews or focus groups (Aquilino 1998; Desimone and Floch 2004; Fowler Jr 2009), and teachers' self-reports of their instructional practice are highly correlated with classroom observations and teacher logs (Smithson and Porter 1994; Mullens and Kasprzyk 1999).

Secondly, while this study focused on only the three key policy tools most prevalent in standard-based reforms and examined how the tools moderate complex instruction and achievement, it is possible that the cross-national variation in the extent to which complex instruction affects class achievement can be explained by country-level factors other than the three policy variables. The ways policy tools, complex instruction, and achievement are interrelated may be conditioned by many other contextual factors, including educational system factors such as educational tracking, teacher recruitment and training, teacher professional community, and other country-specific

structural, cultural, and social differences. Country fixed-effect models were run to reduce omitted variable bias. The supplementary runs yielded consistent results.

Finally, it should be noted that findings from this study are associational and not necessarily causal. Without controls for prior achievement it is not possible to assert that complex instruction caused student achievement. Similarly, the moderating effects of policy tools can be deemed causal only when the relationship between complex instruction and achievement are captured over time. Although it is intriguing that complex instruction seems to be more effective in countries with national curriculum and specified instructional guides, this does not mean that the positive results could be replicated by adopting these policy tools in any country. However, findings of this study can serve as useful working hypotheses in future studies, where causal links between curriculum material, teachers' use of curriculum material and student learning can be tested

The findings show that the three components of standard-based policies are arranged in various ways, and that each of the policy tools differentially affects the level and the effectiveness of complex instruction. This implies that having a national curriculum, highly specified instructional guides, and strong test-based control in place may not lead to the desired goal. And, the discussion about CCSS should be more about the appropriate emphasis placed on the policy components. Cross-national studies with data on the content covered in national curriculum, instructional guides, and assessments can not only enrich the interpretation of this study's findings but also deepen our understanding of the theories of actions behind the policy tools.

Certainly, findings from this cross-national study should be supplemented by further examination to better understand why we see these relationships between nation-level action and teaching and learning in classrooms. Qualitative research that includes interviews and observations on classroom practices is warranted. Interviews may include questions on how teachers plan their instruction and, in particular, whether and how national curriculum and instructional guides are consulted when planning for more complex or procedural instructional approaches. It is important to know how the level of specificity guide or constrain teachers' autonomy.

This study has mostly focused on the relationship between policy tools and complex instruction in the light of educational excellence. A natural next step would be to examine how policy tools and complex instruction relate to equity in education. An unequal distribution of procedural and complex instruction among students with different socioeconomic background and ability has been largely researched in the United States. Several international studies have also attended to this issue by linking national curriculum policies with the within-country variation in classroom practices. Furthering these lines of research, future studies could examine whether and how policy tools affect complex instruction and narrow achievement gaps between and within classrooms.

Appendix A3

Table A3.1 Description of variables

Variables	Description
	Classroom level variables
Class mathematics achievement	Five plausible values of the student mathematics IRT scores (*bsmmat01-05*) are aggregated to the classroom level. Sampling weight TOTWGT is used to derive the class average scores.
Complex instruction	A ratio of two instructional task variables: complex and procedural tasks. This measure represents the relative emphasis teachers place on complex tasks over procedural tasks. Complex task is an average of four items: teachers' reported frequency on asking students to explain the reasoning behind and idea (*bt4masea*), relate the learning topic to their daily lives (*bt4masdl*), work on problems with no immediately obvious method of solution (*bt4masws*), and decide on their own procedures for solving complex problems (*bt4mascp*). Recoded the original item from 1 "never" to 4 "every or almost every lesson" (alpha=0.664) Procedural task is an average of three items: teachers' reported frequency on asking students to practice adding, subtracting, multiplying, and dividing without using a calculator (*bt4maspc*), memorize formulas and procedures (*bt4masmf*), and apply facts, concepts and procedures to solve routine problems (*bt4masac*). Recoded the original item from 1 "never" to 4 "every or almost every lesson" (alpha=0.526)
Classroom characteristics	
Class average number of books at home	Class average of students' books at home. Student's books at home (*bs4gbook*) was used. 5=1 to 10, 18=11 to 25, 64 = 26 to 100 , 151 = 101 to 200, 300 = over 200
Class size	Teachers' report on class size (*bt4mstud*)
Teacher Background	
Teacher's years of experience	Teachers' report on their years of experience (*bt4gtaut*)
Degree in math or math education	Coded 1 if majored on math or math education and 0 otherwise (*bt4mpsm*a and *bt4mpsem*)
Teacher perceived school context	
Instructional resources in the school	Four items that asked teachers about material shortage on textbooks (*bt4mli09*), instructional equipment (*bt4mli10*), equipment for demonstration (*bt4mli11*), and physical facilities (*bt4mli12*) were coded as 1 if no shortage at all and 0 otherwise. Sum of the four recoded variables were used.
Teacher collaboration in the school	Frequency of collaboration among mathematics teachers (*btdmcoll*). Coded 1 if collaborate more than 2 or 3 times per month and 0 if never or almost never collaborate
High school climate	TIMSS derived school climate variable (*btdmtpsc*) is used. Coded 1 if climate is reported as high and 0 if medium or low.
	Country-level variables
Educational policy tools	
National curriculum	Coded 1 if the country has a national curriculum, and 0 otherwise (*buhmncc*).
Specificity of instructional guides	Items that asked whether the mathematics curriculum was made available in textbooks (*buhmcmat)*, instructional guides (*buhmcmai)*, and instructional activities (*buhmcmas)*, and items that asked whether instructional materials (*buhmmpma)* and instructional methods (*buhmmpppm*) are prescribed in the math curriculum were used. Coded 2 if all five are present, 1 if four are present, and 0 for the rest.
Test-based control	Items that asked whether a country has a high-stakes assessment (b*uhmaae*), whether student test results are used to evaluate curriculum implementation (*buhmieas*), and whether percentage of students reaching defined goals are prescribed in the curriculum (*buhmmprg*) were used. Coded 2 if all three are present, coded 1 if two are present, and coded 0 if only one item applies
Country control variable	
GDP per capita (log)	Retrieved from the World Bank's 2007 World Development Indicators data (World Bank 2007).

References

Achinstein, Betty, Rodney T. Ogawa, and Anna Speiglman. 2004. "Are We Creating Separate and Unequal Tracks of Teachers? The Effects of State Policy, Local Conditions, and Teacher Characteristics on New Teacher Socialization." *American Educational Research Journal* 41 (3):557-603.

Alexander, Robin J. 2000. *Culture and Pedagogy: International Comparisons in Primary Education*. Oxford: Blackwell Publishers.

An, Shuhua, Gerald Kulm, and Zhonghe Wu. 2004. "The Pedagogical Content Knowledge of Middle School, Mathematics Teachers in China and the U.S." *Journal of Mathematics Teacher Education* 7:145-72.

Anderson, Lorin W., Doris W. Ryan, and Bernard J. Shapiro. 1989. *The IEA Classroom Environment Study*. Oxford: Pergamon.

Anderson-Levitt, Kathryn M. 2002. "Teacher Culture as National and Transnational: A Response to Teachers' Work." *Educational Researcher* 31 (3):19-21.

Anderson-Levitt, Kathryn M. 2002. *Teaching Cultures: Knowledge for Teaching First Grade in France and the United States*. Cresskill, NJ: Hampton.

Anderson-Levitt, Kathryn M., ed. 2003. *Local Meanings, Global Schooling: Anthropology and World Culture Theory*. New York: Palgrave Macmillan.

Anderson-Levitt, Kathryn M., and Boubaca B. Diallo. 2003. "Teaching by the Book in Guinea." In *Local Meanings, Global Schooling: Anthropology and World Culture Theory*, edited by K. M. Anderson-Levitt. New York: Macmillan.

Andrews, Paul. 2007. "Negotiating Meaning in Cross-National Studies of Mathematics Teaching: Kissing Frogs to Find Princes." *Comparative Education* 43 (4):489-509.

Andrews, Paul. 2013. "Finnish Mathematics Teaching from a Reform Perspective: A Video-Based Case-Study Analysis." *Comparative Education Review* 57 (2):189-211.

Aquilino, William S. 1998. "Effects of Interview Mode on Measuring Depression in Younger Adults." *Journal of Official Statistics* 14 (1):15-29.

Astiz, M. Fernanda, Alexander W. Wiseman, and David P. Baker. 2002. "Slouching towards Decentralization: Consequences of Globalization for Curricular Control in National Education Systems." *Comparative Education Review* 46 (1):66-88.

Au, Wayne. 2007. "High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis." *Educational Researcher* 36 (5):258-267.

Baker, David, Helmut Köhler, and Manfred Stock. 2007. "Socialist Ideology and the Contraction of Higher Education: Institutional Consequences of State Manpower and Education Planning in the Former East Germany." *Comparative Education Review* 51 (3):353-77.

Baker, David P. 1997. "Surviving TIMSS: Or, Everything You Blissfully Forgot about International Comparisons." *Phi Delta Kappan* 78 (4):295-300.

Baker, David P. 2009. "The Invisible Hand of World Education Culture." In *Handbook of Education Policy Research*, edited by Gary Sykes, Barbara Schneider and David N. Plank. New York: Routledge.

Baker, David P., and Gerald K. LeTendre. 2000. "Comparative Sociology of Classroom Processes, School Organization, and Achievement." In *Handbook of the Sociology of Education*, edited by Maureen T. Hallinan. New York: Kluwer Academic/Plenum Publishers.

Baker, David P., and Gerald K. LeTendre. 2005. *National Differences and Global Similarities*. Stanford, CA: Stanford University Press.

Ball, Deborah Loewenberg, and David K. Cohen. 1996. "Reform by The Book: What is-or might be-The Role of Curriculum Materials in Teacher Learning and Instructional Reform." *Educational Researcher* 25 (9):6-8.

Ball, Stephen J. 1998. "Big Policies/Small World: An Introduction to International Perspectives in Education Policy." *Comparative Education* 34 (2):119-130.

Baroody, A. J., and A. Benson. 2001. "Early Number Instruction." *Teaching Children Mathematics* 8 (154-158).

Barr, Rebecca, Robert Dreeben, and Nonglak Wiratchai. 1983. *How Schools Work*. Chicago: The University of Chicago Press.

Bartlett, Lesley. 2003. "World Culture or Transnational Project? Competing Educational Projects in Brazil." In *Local Meanings, Global Schooling: Anthropology and World Culture Theory*, edited by K. M. Anderson-Levitt. New York: Macmillan.

Benavot, Aaron, and Cecilia Braslavsky, eds. 2007. *School Knowledge in Comparative and Historical Perspective: Changing Curricula in Primary and Secondary Education*. Hong Kong: Comparative Education Research Center, University of Hong Kong / Springer.

Benavot, Aaron, Yun-Kyung Cha, David H. Kamens, John W. Meyer, and Suk-Ying Wong. 1991. "Knowledge for the Masses: World Models and National Curricula, 1920-1986." *American Sociological Review* 56 (1):85-100.

Bishop, John H. 1997. "The Effect of National Standards and Curriculum-Based Exams on Achievement." *The American Economic Review* 87 (2):260-264.

Bjork, Christopher, ed. 2006. *Educational Decentralization: Asian Experiences and Conceptual Contributions*. Dordrecht: Springer.

Bjork, Christopher. 2009. "Local Implementation of Japan's Integrated Studies Reform: A Preliminary Analysis of Efforts to Decentralise the Curriculum." *Comparative Education* 45 (1):23-44.

Bodovski, Katerina, and Gearge Farkas. 2007. "Do Instructional Practices Contribute to Inequality in Achievement? The Case of Mathematics Instruction in Kindergarten." *Journal of Early Childhood Research* 5 (3):301-322.

Boli, John, Francisco O. Ramirez, and John W. Meyer. 1985. "Explaining the Origins and Expansion of Mass Education." *Comparative Education Review* 29 (2):145-70.

Booher-Jennings, Jennifer. 2005. "Below the Bubble: "Educational Triage" and the Texas Accountability System." *American Educational Research Journal* 42 (2):230-268.

Bossert, Steven T. 1979. *Tasks and Social Relationships in Classrooms: A Study of Instructional Organization and its Consequences*. London: Cambridge University Press.

Bromley, Patricia, John W. Meyer, and Francisco O. Ramirez. 2011. "Student-Centeredness in Social Science Textbooks, 1970-2008: A Cross-National Study." *Social Forces* 90 (2):547-70.

Bromley, Patricia, John W. Meyer, and Francisco O. Ramirez. 2011. "The Worldwide Spread of Environmental Discourse in Social Studies, History, and Civics Textbooks, 1970-2008." *Comparative Education Review* 55 (4):517-45.

Buchmann, Claudia. 2011. "Frontiers in Comparative and International Sociology of Education: American Distinctiveness and Global Diversity." In *Frontiers in Sociology of Education*, edited by M. T. Hallinan. New York: Springer.

Burstein, L., L. M. McDonnell, J. Van Winkle, T. Ormseth, J. Mirocha, and G. Guitton. 1995. *Validating National Curriculum Indicators*. Santa Monica, CA: RAND.

Carney, Stephen. 2009. "Negotiating Policy in an Age of Globalization: Exploring Educational 'Policyscapes' in Denmark, Nepal, and China." *Comparative Education Review* 53 (1):63-88.

Carney, Stephen, Jeremy Rappleye, and Iveta Silova. 2012. "Between Faith and Science: World Culture Theory and Comparative Education." *Comparative Education Review* 56 (3):366-93.

Carnoy, Martin, and Susanna Loeb. 2002. "Does External Accountability Affect Student Outcomes? A Cross-State Analysis." *Educational Evaluation and Policy Analysis* 24 (4):305-331.

Carpenter, T., E. Fennema, P. Peterson, C. Chiang, and M. Loef. 1989. "Using Knowledge of Children's Mathematics Thinking in Classroom Teaching: An Experimental Study." *American Educational Research Journal* 26 (4).

Carpenter, Thomas P., Elizabeth Fennema, Penelope L. Peterson, Chi-Pang Chiang, and Megan Loef. 1989. "Using Knowledge of Children's Mathematics Thinking in Classroom Teaching: an Experimental Study." *American Educational Research Journal* 26 (4):499-531.

Cha, Yun-Kyung. 2007. "The Spread of English Language Instruction in the Primary School." In *School Knowledge in Comparative and Historical Perspective: Changing Curricula in Primary and Secondary Education*, edited by Aaron Benavot and Cecilia Braslavsky. Hong Kong: Springer.

Clarke, David, and Li Hua Xu. 2008. "Distinguishing between Mathematics Classrooms in Australia, China, Japan, Korea and the USA through the Lens of the Distribution of Responsibility for Knowledge Generation: Public Oral Interactivity and Mathematical Orality." *ZDM Mathematics Education* 40 (6):963-72.

Clune, William. 1998. "Toward a theory of systemic reform: The case of nine NSF statewide systemic initiatives." *National Institute for Science Education Research Monograph* 16.

Cohen, David K. 1990. "A Revolution in One Classroom: The Case of Mrs. Oublier." *Educational Evaluation and Policy Analysis* 12 (3):311-29.

Cohen, David K., and Carol A. Barnes. 1993. "Pedagogy and Policy." In *Teaching for Understanding: Challenges for Policy and Practice*, edited by David K. Cohen, Milbrey W. McLaughlin and Joan E. Talbert. San Francisco: Jossey-Bass.

Cohen, David K., Susan L. Moffitt, and Simona Goldin. 2007. "Policy and Practice: the Dilemma." *American Journal of Education* 113 (4):515-548.

Cohen, David K., and James P. Spillane. 1992. "Policy and Practice: The Relations between Governance and Instruction." *Review of Research in Education* 18:3-49.

Common Core State Standards Initiative. 2010. "Common Core State Standards for Mathematics." http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf.

Cummings, William K. 1999. "The InstitutionS of Education: Compare, Compare, Compare!" *Comparative Education Review* 43 (4):413-37.

Dahlin, Bo, and David A. Watkins. 2000. "The Role of Repetition in the Process of Memorising and Understanding: A Comparison of the Views of German and Chinese Secondary School Students in Hong Kong." *British Journal of Educational Psychology* 70 (1):65-84.

Davies, Scott, and Neil Guppy. 1997. "Globalization and Educational Reforms in Anglo-American Democracies." *Comparative Education Review* 41 (4):436-459.

DeCoker, Gary, and Christopher Bjork, eds. 2013. *Japanese Education in an Era of Globalization: Culture, Politics, and Equity*: Teachers College Press.

Dee, Thomas S., and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and Management* 30 (3):418-446.

Desimone, Laura. 2002. "How Can Comprehensive School Reform Models be Successfully Implemented?" *Review of Educational Research* 72 (3):433-479.

Desimone, Laura M. 2013. "Teacher and Administrator Responses to Standard-Based Reform."

Teachers College Record 115 (8).

Desimone, Laura M., and Kerstin Carlson Le Floch. 2004. "Are We Asking the Right Questions? Using Cognitive Interviews to Improve Surveys in Education Research." *Educational Evaluation and Policy Analysis* 26 (1):1-22.

Desimone, Laura M., and K. LeFloch. 2004. "Are We Asking the Right Questions? Using Cognitive Interviews to Improve Surveys in Education Research." *Educational Evaluation and Policy Analysis* 26:1-22.

Desimone, Laura M., and Daniel Long. 2010. "Teacher Effects and the Achievement Gap: Do Teacher and Teaching Quality Influence the Achievement Gap between Black and White and High- and Low-SES Students in the Early Grades?" *Teachers College Record* 112 (12):3024-3073.

Desimone, Laura M., Thomas Smith, David Baker, and Koji Ueno. 2005. "Assessing Barriers to the Reform of U.S. Mathematics Instruction from an International Perspective." *American Educational Research Journal* 42 (3):510-535.

Desimone, Laura M., Thomas M. Smith, Susan A. Hayes, and David Frisvold. 2005. "Beyond Accountability and Average Mathematics Scores: Relating State Education Policy Attributes to Cognitive Achievement Domains." *Educational Measurement: Issues and Practice* 24 (4):5-18.

Diamond, John, and James Spillane. 2004. "High-Stakes Accountability in Urban Elementary Schools: Challenging or Reproducing Inequality?" *The Teachers College Record* 106 (6):1145-1176.

Doyle, Walter. 1981. "Research on Classroom Contexts." *Journal of Teacher Education* 32 (6):3-6.

Doyle, Walter. 1988. "Work in Mathematics Classes: The Context of Students' Thinking During Instruction." *Educational Psychologist* 23 (2):167-80.

Ertl, Hubert. 2006. "Educational Standards and the Changing Discourse on Education: the Reception and Consequences of the PISA Study in Germany." *Oxford Review of Education* 32 (5):619-634.

Fan, Lianghuo, and Gurcharn S. Kaeley. 2000. "The Influence of Textbooks on Teaching Strategies: An Empirical Study." *Mid-Western Educational Researcher* 13 (4):2-9.

Fang, Yanping, and S. Gopinathan. 2009. "Teachers and Teaching in Eastern and Western Schools: A Critical Review of Cross-Cultural Comparative Studies." In *International Handbook of Research on Teachers and Teaching*, edited by Lawrence J. Saha and A. Gary Dworkin. New York, NY: Springer.

Feenstra, Robert C. 1998. "Integration of Trade and Disintegration of Production in the Global Economy." *The Journal of Economic Perspectives* 12 (4):31-50.

Fiala, Robert. 2007. "Educational Ideology and the School Curriculum." In *School Knowledge in*

Comparative and Historical Perspective: Changing Curricula in Primary and Secondary Education, edited by Aaron Benavot and Cecilia Braslavsky. Hong Kong: Springer.

Fowler Jr, Floyd J. 2009. *Survey Research Methods*. Thousand Oaks, CA: Sage publications.

Foy, P., and J. F. Olson. 2009. TIMSS 2007 User Guide for the International Database. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Frank, David John, Ann Hironaka, and Evan Schofer. 2000. "The Nation-State and the Natural Environment over the Twentieth Century." *American Sociological Review* 65 (1):96-116.

Frank, David John, Suk-Ying Wong, John W. Meyer, and Francisco O. Ramirez. 2000. "What Counts as History: A Cross-National and Longitudinal Study of University Curricula." *Comparative Education Review* 44 (1):29-53.

Fuchs, Thomas, and Ludger Woessmann. 2008. "What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data " In *The Economics of Education and Training*, edited by Christian Dustmann, Bernd Fitzenberger and Stephen Machin. Springer Science & Business Media.

Fukuzawa, Rebecca Erwin. 1996. "The Path to Adulthood According to Japanese Middle Schools." In *Teaching and Learning in Japan*, edited by Thomas P. Rohlen and Gerald K. LeTendre. Cambridge: Cambridge University Press.

Fukuzawa, Rebecca Erwin, and Gerald K. LeTendre. 2001. *Intense Years: How Japanese Adolescents Balance School, Family, and Friends*. New York: Routledge Falmer.

Fuller, Bruce. 2008. "Overview: Liberal Learning in Centralizing States." In *Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability*, edited by Bruce Fuller, Melissa K. Henne and Emily Hannum. Howard House, Bingley, UK: Emerald Group Publishing Limited.

Givvin, Karen Bogard, James Hiebert, Jennifer K. Jacobs, Hilary Hollingsworth, and Ronald Gallimore. 2005. "Are There National Patterns of Teaching? Evidence from the TIMSS 1999 Video Study." *Comparative Education Review* 49 (3):311-43.

Grodsky, Eric, John Robert Warren, and Demetra Kalogrides. 2009. "State High School Exit Examinations and NAEP Long-Term Trends in Reading and Mathematics, 1971-2004." *Educational Policy* 23 (4):589-614.

Grossman, Pam, and Morva McDonald. 2008. "Back to the Future: Directions for Research in Teaching and Teacher Education." *American Educational Research Journal* 45 (1):184-205.

Hamilton, Laura S., and Jose Felipe Martinez. 2007. "What Can TIMSS Surveys Tell Us About Mathematics Reforms in the United States During the 1990s?" In *Lessons Learned: What International Assessments Tell Us About Mathematics Achievement*, edited by Tom Loveless. Washington, DC: Brookings Institutions.

Hamilton, Laura S., Daniel F. McCaffrey, Brian M. Stecher, Stephen P. Klein, Abby Robyn, and

Delia Bugliari. 2003. "Studying Large-Scale Reforms of Instructional Practice: An Example from Mathematics and Science." *Educational Evaluation and Policy Analysis* 25 (1):1-29.

Hanushek, Eric A, and Margaret E. Raymond. 2004. Does School Accountability Lead to Improved Student Performance? : National Bureau of Economic Research.

Hargreaves, Andy. 2003. *Teaching in the Knowledge Society: Education in the Age of Insecurity*. New York: Teachers College Press.

Hayashi, Akiko, and Joseph Tobin. 2014. "The Power of Implicit Teaching Practices: Continuities and Discontinuities in Pedagogical Approaches of Deaf and Hearing Preschools in Japan." *Comparative Education Review* 58 (1):24-46.

Hiebert, James, Ronald Gallimore, Helen Garnier, Karen Bogard Givvin, Hilary Hollingsworth, Jennifer K. Jacobs, Angel Miu-Ying Chui, Diana Wearne, Margaret Smith, Nicole Kersting, Alfred Manaster, Ellen Tseng, Wallace Etterbeek, Carl Manaster, Patrick Gonzales, and James W. Stigler. 2003. *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study*. Washington DC: National Center for Education Statistics.

Hiebert, James, and Diana Wearne. 1993. "Instructional Tasks, Classroom Discourse, and Students' Learning in Second-Grade Arithmetic." *American Educational Research Journal* 30 (2).

Ho, Irene T. 2001. "Are Chinese Teachers Authoritarian?" In *Teaching the Chinese Learner: Psychological and Pedagogical Perspectives*, edited by David A. Watkins and John B. Biggs. Hong Kong/ Melbourne: Comparative Education Research Centre, The University of Hong Kong/Australian Council for Educational Research.

House, J. Daniel. 2002. "Instructional Practices and Mathematics Achievement of Adolescent Students in Chinese Taipei: Results from the TIMSS 1999 Assessment." *Child Study Journal* 32 (3):157-78.

House, J. Daniel. 2009. "Elementary-school Mathematics Instruction and Achievement of Fourth-Grade Students in Japan: Findings from the TIMSS 2007 Assessment." *Education* 130 (2):301-307.

House, J. Daniel, and James A. Telese. 2008. "Relationships Between Student and Instructional Factors and Algebra Achievement of Students in the United States and Japan: An Analysis of TIMSS 2003 Data." *Educational Research and Evaluation: An International Journal on Theory and Practice* 14 (1):101-12.

Hout, Michael, and Stuart W. Elliott, eds. 2011. *Incentives and Test-Based Accountability in Education*: National Academic Press.

Huang, Rongjin, and Frederick Koon Shing Leung. 2004. "Cracking the Paradox of Chinese Learners: Looking into the Mathematics Classrooms in Hong Kong and Shanghai." In *How Chinese Learn Mathematics: Perspectives from Insiders*, edited by Fan Lianghuo, Wong Ngai-Ying, Cai Jinfa and Li Shiqi. Singapore: World Scientific Publishing.

Irwin, Douglas A. 1996. "The United States in a New Global Economy? A Century's Perspective." *The American Economic Review* 86 (2):41-46.

Jürges, Hendrik, Kerstin Schneider, and Felix Büchel. 2005. "The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany." *Journal of the European Economic Association* 3 (5):1134-1155.

Kamens, David H., John W. Meyer, and Aaron Benavot. 1996. "Worldwide Patterns in Academic Secondary Education Curricula." *Comparative Education Review* 40 (2):116-38.

Kariya, Takehiko. 2011. "Japanese Solutions to the Equity and Efficiency Dilemma? Secondary Schools, Inequity and the Arrival of 'Universal' Higher Education." *Oxford Review of Education* 37 (2):241-66.

Kariya, Takehiko, and James E. Rosenbaum. 1999. "Bright Flight: Unintended Consequences of Detracking Policy in Japan." *American Journal of Education* 107 (3):210-30.

Kember, David. 1996. "The Intention to Both Memorise and Understand: Another Approach to Learning?" *Higher Education* 31 (3):341-54.

Kilpatrick, Jeremy, Jane Swafford, and Bradford Findell. 2001. *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: National Academy Press.

Klenowski, Valentina, and Claire Wyatt-Smith. 2010. "Standards, Teacher Judgment and Moderation in Contexts of National Curriculum and Assessment Reform." *Assessment Matters* 2:107-131.

Lee, Jaekyung, and Todd Reeves. 2012. "Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990–2009 Reading and Math Achievement Gaps and Trends." *Educational Evaluation and Policy Analysis* 34 (2):209-231.

LeTendre, Gerald K., ed. 1999. *Competitor or Ally? Japan's Role in American Educational Debates*. New York: Falmer Press.

LeTendre, Gerald K. 1999. "The Problem of Japan: Qualitative Studies and International Educational Comparisons." *Educational Researcher* 28 (2):38-45.

LeTendre, Gerald K., David P. Baker, Motoko Akiba, Brian Goesling, and Alex Wiseman. 2001. "Teachers' Work: Institutional Isomorphism and Cultural Variation in the U.S., Germany, and Japan." *Educational Researcher* 30 (6):3-15.

Leung, Frederick Koon Shing. 1995. "The Mathematics Classroom in Beijing, Hong Kong and London." *Educational Studies in Mathematics* 29 (4):297-325.

Leung, Frederick Koon Shing. 2001. "In Search of an East Asian Identity in Mathematics Education." *Educational Studies in Mathematics* 47:35-51.

Leung, Frederick Koon Shing. 2002. "Behind the High Achievement of East Asian Students." *Educational Research and Evaluation* 8 (1):87-108.

Leung, Frederick Koon Shing. 2006. "Mathematics Education in the East Asia and the West: Does Culture Matter?" In *Mathematics Education in Different Cultural Traditions-A Comparative Study of East Asia and the West. The 13th ICMI study*, edited by Frederick Koon-Shing Leung, Klaus-D. Graf and Francis J. Lopez-Real, 21-46. New York: Springer.

Levin, Benjamin. 2001. *Reforming Education: From Origins to Outcomes*. New York: Falmer.

Lewis, Catherine. 1995. *Educating Hearts and Minds: Reflections on Japanese Preschool and Elementary Education*. Cambridge: Cambridge University Press.

Lewis, Catherine, Rebecca Perry, and Aki Murata. 2006. "How Should Research Contribute to Instructional Improvement? The Case of Lesson Study." *Educational Researcher* 35 (3):3-14.

Lewis, Catherine, and Ineko Tsuchida. 1997. "Planned Educational Change in Japan: The Case of Elementary Science Instruction." *Journal of Education Policy* 12 (5):313-31.

Li, Jin. 2002. "A Cultural Model of Learning: Chinese "Heart and Mind for Wanting to Learn"." *Journal of Cross-Cultural Psychology* 33 (3):248-69.

Lianghuo, Fan, Wong Ngai-Ying, Cai Jinfa, and Li Shiqi, eds. 2004. *How Chinese Learn Mathematics: Perspectives from the Insiders*, *Series on Mathematics Education*: World Scientific.

Lucas, Samuel R, and Adam Gamoran. 2002. "Tracking and the Achievement Gap." In *Bridging the Achievement Gap* edited by John E. Chubb and Tom Loveless. Brooking Institution Press.

Ma, Liping. 1999. *Knowing and Teaching Elementary Mathematics: Teachers' Understanding of Fundamental Mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Maaz, Kai, Ulrich Trautwein, Oliver Lüdtke, and Jürgen Baumert. 2008. "Educational Transitions and Differential Learning Environments: How Explicit Between-School Tracking Contributes to Social Inequality in Educational Outcomes." *Child Development Perspectives* 2 (2):99-106.

Marton, Ference, Gloria Dall'Alba, and Lai Kun Tse. 1996. "Memorizing and Understanding: The Keys to the Paradox?" In *The Chinese Learner: Cultural, Psychological and Contextual Influences*, edited by David A. Watkins and John B. Biggs. Hong Kong: Comparative Education Research Centre, The University of Hong Kong.

McNeely, Connie L. 1995. "Prescribing National Education Policies: The Role of International Organizations." *Comparative Education Review* 39 (4):483-507.

Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103 (1):144-81.

Meyer, John W., Patricia Bromley, and Francisco O. Ramirez. 2010. "Human Rights in Social Science Textbooks : Cross-national Analyses, 1970-2008." *Sociology of Education* 83 (2):111-34.

Meyer, John W., David H. Kamens, Aaron Benavot, Yun-Kyung Cha, and Suk-Ying Wong. 1992.

School Knowledge for the Masses: World Models and National Primary Curricular Categories in the Twentieth Century. Washington, DC: The Falmer Press.

Meyer, John W., and Francisco O. Ramirez. 2000. "The World Institutionalization of Education." In *Discourse Formation in Comparative Education*, edited by J. Schriewer, 111-32. New York: Peter Lang Publishers.

Meyer, John W., Francisco O. Ramirez, Richard Rubinson, and John Boli-Bennett. 1977. "The World Educational Revolution, 1950-1970." *Sociology of Education* 50 (4):242-58.

Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83 (2):340-63.

Minor, Elizabeth Covay, Laura M. Desimone, Kristie J. R. Phillips, and Kailey Spencer. 2015. "A New Look at the Opportunity-to-Learn Gap across Race and Income." *American Journal of Education* 121 (2):241-269.

Mitter, Wolfgang. 2003. "A Decade of Transformation: Educational Policies in Central and Eastern Europe." In *Comparative Education: Continuing Traditions, New Challenges, and New Paradigms*, edited by Mark Bray. Dordrecht: Kluwer Academic Publishers.

Mohammadpour, Ebrahim. 2013. "A Three-Level Multilevel Analysis of Singaporean Eighth-Graders Science Achievement." *Learning and Individual Differences* 26:212-20.

Mok, Ida Ah Chee. 2006. "Shedding Light on the East Asian Learner Paradox: Reconstructing Student-Centredness in a Shanghai Classroom." *Asia Pacific Journal of Education* 26 (2):131-142.

Møller, Jorunn, and Guri Skedsmo. 2013. "Norway: Centralisation and Decentralisation as Twin Reform Strategies." In *Transnational Influences on Values and Practices in Nordic Educational Leadership: Is There a Nordic Model?*, edited by Lejf Moos. Netherlands: Springer Science & Business Media.

Mullens, John E., and Daniel Kasprzyk. 1999. Validating Item Responses on Self-Report Teacher Surveys. Washington, DC: US Department of Education.

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

Mullis, Ina V. S., Michael O. Martin, Eugenia J. Gonzalez, Kelvin D. Gregory, Robert A. Garden, Kathleen M. O'Connor, Steven J. Chrostowski, and Teresa A. Smith. 2000. TIMSS 1999 International Mathematics Report. Boston: International Study Center, Boston College, Lynch School of Education.

Mullis, Ian V.S., Michael O. Martin, Graham J. Ruddock, Christine Y. O'Sullivan, Alka Arora, and Ebru Erberber. 2005. TIMSS 2007 Assessment Frameworks. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mundy, Karen. 2005. "Globalization and Educational Change: New Policy Worlds." In *International Handbook of Educational Policy*, edited by Nina Bascia, Alister Cumming, Amanda Datnow, Kenneth Leithwood and David Livingstone, 3-17. Netherlands: Springer.

National Council of Teachers of Mathematics. 2000. *Principles and Standards for School Mathematics*. Vol. 1: National Council of Teachers of Mathematics

Newmann, Fred M., Helen M. Marks, and Adam Gamoran. 1996. "Authentic Pedagogy and Student Performance." *American Journal of Education* 104 (4):280-312.

Norris, N., R. Aspland, B. MacDonald, J. Schostak, and B. Zamoriski. 1996. An Independent Evaluation of Comprehensive Curriculum Reform in Finland. Helsinki: National Board of Education.

Oakes, Jeannie. 1985. *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.

Olson, John F., Michael O. Martin, and Ian V. S. Mullis. 2008. *TIMSS 2007 Technical Report*: IEA TIMSS & PIRLS.

Paik, Susan J. 2004. "Korean and US Families, Schools, and Learning." *International Journal of Educational Research* 41:71-90.

Paine, Lynn W. 1990. "The Teacher as Virtuoso: A Chinese Model for Teaching." *Teachers College Record* 92:49-81.

Park, Hyunjoon. 2013. *Re-evaluating Education in Japan and Korea : Demystifying Stereotypes*. New York: Routledge.

Pennington, Matthew. 1999. "Asia Takes a Crash Course in Educational Reform." *UNESCO Courier-English Edition* 52 (7):17-20.

Polikoff, Morgan S. 2012. "Instructional Alignment under No Child Left Behind." *American Journal of Education* 118.

Pong, Suet-ling, and Aaron Pallas. 2001. "Class Size and Eighth-Grade Math Achievement in the United States and Abroad." *Educational Evaluation and Policy Analysis* 23 (3):251-73.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2006. "Multilevel Modeling of Complex Survey Data." *Journal of the Royal Statistical Society* 169 (4):805-27.

Ramirez, Francisco O. 2003. "The Global Model and National Legacies." In *Local Meanings, Global Schooling: Anthropology and World Culture Theory*, edited by Kathryn Anderson-Levitt. New York: Palgrave Macmillan.

Ramirez, Francisco O. 2012. "The World Society Perspective: Concepts, Assumptions, and Strategies." *Comparative Education* 48 (4):423-39.

Ramirez, Francisco O., and John Boli. 1987. "The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization." *Sociology of Education* 60 (1):2-17.

Raudenbush, Stephen W. 1989. "The Analysis of Longitudinal, Multilevel Data." *International Journal of Educational Research* 13 (7):721-40.

Raudenbush, S. W., and A. S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Raudenbush, Stephen W., Brian Rowan, and Yuk Fai Cheong. 1993. "Higher Order Instructional Goals in Secondary Schools: Class, Teachers, and School Influences." *American Educational Research Journal* 30 (3):523-53.

Rice, Jennifer King. 1999. "The Impact of Class Size on Instructional Strategies and the Use of Time in High School Mathematics and Science Courses." *Educational Evaluation and Policy Analysis* 21 (2):215-29.

Rohlen, Thomas P. 1983. *Japan's High Schools*. Berkeley: University of California Press.

Rohlen, Thomas P., and Gerald K. LeTendre. 1996. *Teaching and Learning in Japan*. Cambridge: Cambridge University Press.

Romberg, Thomas A. 1988. "One Point of View: NCTM's Curriculum and Evaluation Standards: What They Are and Why They Are Needed." *Arithmetic Teacher* 35 (9):2-3.

Rowan, Brian. 1990. "Commitment and Control: Alternative Strategies for the Organizational Design of Schools." *Review of Research in Education* 16:353-89.

Rutkowski, Leslie, Eugenio Gonzalez, Marc Joncas, and Matthias von Davier. 2010. "International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting." *Educational Researcher* 39 (2):142-51.

Sahlberg, Pasi. 2006. "Education Reform for Raising Economic Competitiveness." *Journal of Educational Change* 7:259-87.

Sahlberg, Pasi. 2010. "Educational Change in Finland." In *Second International Handbook of Educational Change*, edited by Andy Hargreaves, Ann Lieberman, Michael Fullan and David Hopkins. New York: Springer.

Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Cambridge: Harvard University Press.

Sandholtz, Judith, Rodney Ogawa, and Samantha Scribner. 2004. "Standards Gaps: Unintended Consequences of Local Standards-Based Reform." *The Teachers College Record* 106 (6):1177-1202.

Sawyer, R. Keith. 2004. "Creative Teaching: Collaborative Discussion ss Disciplined Improvisation." *Educational Researcher* 33 (2):12-20.

Schaub, Maryellen, and David P. Baker. 1991. "Solving the Math Problem: Exploring Mathematics Achievement in Japanese and American Middle Grades." *American Journal of Education* 99 (4):623-42.

Schmidt, William H., Doris Jorde, Leland S. Cogan, Emilie Barrier, Ignacio Gonzalo, Urs Moser, Katsuhiko Shimizu, Toshio Sawada, Gilbert A Valverde, Curtis McKnight, Richard S Prawat, David E Wiley, Senta A Raizen, Edward D Britton, and Richard G Wolfe. 2002. *Characterizing Pedagogical Flow: An Investigation of Mathematics and Science Teaching in Six Countries*. Netherlands: Springer.

Schmidt, William H., Doris Jorde, Leland S. Cogan, Emilie Barrier, Ignacio Gonzalo, Urs Moser, Katsuhiko Shimizu, Toshio Sawada, Gilbert A. Valverde, Curtis McKnight, Richard S. Prawata, David E. Wiley, Senta A. Raizen, Edward D. Britton, and Richard G. Wolfe. 1996. *Characterizing Pedagogical Flow: An Investigation of Mathematics and Science Teaching in Six Countries*. Netherlands: Kluwer Academic Publishers.

Schmidt, William H., Curtis C. McKnight, Leland S. Cogan, Pamela M. Jakwerth, and Richard T. Houang. 1999. *Facing the Consequences : Using TIMSS for a Closer Look at U.S. Mathematics and Science Education*: Kluwer Academic Publishers.

Schmidt, William H., Curtis C. McKnight, Richard T. Houang, HsingChi Wang, David E. Wiley, Leland S. Cogan, and Richard G. Wolfe. 2001. *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*. San Francisco, CA: The Jossey-Bass Education Series.

Schmidt, William H., and Richard S. Prawat. 2006. "Curriculum Coherence and National Control of Education: Issue or Non-Issue." *Journal of Curriculum Studies* 38 (6):641-58.

Schneiberg, Marc, and Elisapeth S. Clemens. 2006. "The Typical Tools for the Job: Research Strategies in Institutional Analysis." *Sociological Theory* 24 (3):195-227.

Scott, W. Richard, and Gerald F. Davis. 2007. *Organizations and Organizing: Rational, Natural, and Open System Perspectives*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Shen, Ce. 2005. "How American Middle Schools Differ from Schools of Five Asian Countries: Based on Cross-national Data from TIMSS 1999." *Educational Research and Evaluation* 11 (2):179-99.

Shouse, Roger. 2001. "The Impact of Traditional and Reform-Style Practices on Student Mathematics Achievement." In *Great Curriculum Debate: How Should We Teach Reading and Math?*, edited by Tom Loveless. Washington, DC: Brookings Institution Press.

Silova, Iveta. 2009. "Varieties of Educational Transformation: The Post-socialist States of Central/Souteastern Europe and the Former Soviet Union " In *Second International Handbook of Comparative Education*, edited by Robert Cowen and Andreas M. Kazamias. New York: Springer.

Simola, Hannu. 2005. "The Finnish Miracle of PISA: Historical and Sociological Remarks on Teaching and Teacher Education." *Comparative Education* 41 (4):455-70.

Smithson, John L., and Andrew C. Porter. 1994. Measuring Classroom Practice: Lessons Learned from Efforts to Describe the Enacted Curriculum--The Reform Up Close Study. In *CPRE Research Report Series*.

Spillane, James P. 2007. "The Making and Effects of Education Policy." In *The State of Education Policy Research*, edited by Susan H. Fuhrman, David K. Cohen and Fritz Mosher. New Jersey: Lawrence Erlbaum Associates, Inc.

Stein, Mary Kay, Barbara W. Grover, and Marjorie Henningsen. 1996. "Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms." *American Educational Research Journal* 33 (2):455-88.

Stevenson, David Lee, and David P. Baker. 1991. "State Control of the Curriculum and Classroom Instruction." *Sociology of Education* 64 (1):1-10.

Stevenson, Harold W., and James W. Stigler. 1992. *The Learning Gap: Why Our Schools are Failing and What We Can Learn from Japanese and Chinese Education*. New York: Summit Books.

Stigler, James W., and Ruth Baranes. 1988. "Culture and Mathematics Learning." *Review of Research in Education* 15:253-306.

Stigler, James W., and James Hiebert. 1999. *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*. New York: The Free Press.

Straubhaar, Rolf. 2014. "The Multiple Influences on Nonformal Instructional Practices in Rural Mozambique: Exploring the Limits of World Culture Theory." *Comparative Education Review* 58 (2):215-40.

Suarez, David. 2006. "The Institutionalization of Human Rights Education." In *The Impact of Comparative Education Research on Neoinstitutional Theory*, edited by David P. Baker and Alexander W. Wiseman. Oxford, UK: Elsevier Science.

Tabulawa, R. 2003. "International Aid Agencies, Learner-centered Pedagogy and Political Democratisation: A Critique." *Comparative Education* 39 (1):7-26.

Tobin, Joseph, Yeh Hsueh, and Mayumi Karasawa. 2009. *Preschool in Three Cultures Revisited: China, Japan and the United States*. Chicago: The University of Chicago Press.

Tobin, Joseph, David Y. H. Wu, and Dana H. Davidson. 1989. *Preschool in Three Cultures: Japan, China, and the United States*. New Haven and London: Yale University Press.

Tobin, Joseph Jay, Yeh Hsueh, and Mayumi Karasawa. 2009. *Preschool in Three Cultures Revisited: China, Japan and the United States*. Chicago: The University of Chicago Press.

Tobin, Joseph Jay, David Y. H. Wu, and Dana H. Davidson. 1987. "Class Size and Student/Teacher Ratios in the Japanese Preschool." *Comparative Education Review* 31 (4):533-49.

Tobin, Joseph Jay, David Y. H. Wu, and Dana H. Davidson. 1989. *Preschool in Three Cultures: Japan,*

China, and the United States. New Haven and London: Yale University Press.

Torney-Purta, Judith, and Jo-Ann Amadeo. 2013. "International Large-Scale Assessments: Challenges in Reporting and Potentials for Secondary Analysis." *Research in Comparative and International Education* 8 (3):248-58.

Tweed, Roger G., and Darrin R. Lehman. 2002. "Learning Considered within a Cultural Context: Confucian and Socratic Approaches." *American Psychologist* 57 (2):89-99.

Union of International Associations. 2000. "The Yearbook of International Organizations Online." Brill. http://ybio.brillonline.nl/system/files/pdf/v5/2001/2_1_1.pdf.

Vavrus, Frances, and Lesley Bartlett. 2012. "Comparative Pedagogies and Epistemological Diversity: Social and Material Contexts of Teaching in Tanzania." *Comparative Education Review* 56 (4):634-58.

Waldow, Florian, Keita Takayama, and Youl-Kwan Sung. 2014. "Rethinking the Pattern of External Policy Referencing: Media Discourses over the 'Asian Tigers" PISA Success in Australia, Germany and South Korea." *Comparative Education* 50 (3):302-321.

Wallerstein, Immanuel. 1974. *The Modern World System*. New York: Academic Press.

Wang, Jian, and Lynn W. Paine. 2003. "Learning to Teach with Mandated Curriculum and Public Examination of Teaching as Contexts." *Teaching and Teacher Education* 19:75-94.

Watkins, David A., and John Burville Biggs. 1996. *The Chinese Learner: Cultural, Psychological and Contextual Influences*. Hong Kong & Melbourne, Australia: Comparative Education Research Centre & Australian Council for Educational Research.

Watkins, David A., Murari Regmi, and Estela Astilla. 1991. "The Asian-Learner-as-a-Rote-Learner Stereotype: Myth or Reality?" *Educational Psychology* 11 (1):21-34.

Westbury, Ian, and Chao-Sheng. Hsu. 1996. "Structures of Curriculum Governance and Classroom Practice in Mathematics." *Educational Evaluation and Policy Analysis* 18 (2):123-39.

White, Merry. 1987. *The Japanese Educational Challenge: A Commitment to Children*. New York: Free Press.

Wiseman, Alexander W. 2008. "A Culture of (In)Equality? A Cross-National Study of Gender Parity and Gender Segregation in National School Systems." *Research in Comparative and International Education* 3 (2):179-201.

Woessmann, Ludger. 2003. "Schooling Resources, Educational Institutions and Student Performance: The International Evidence." *Oxford Bulletin of Economics and Statistics* 65 (2):117-170.

World Bank. 1999. "World Development Indicators." http://data.worldbank.org/data-catalog /world-development-indicators.

World Bank. 2007. "World Development Indicators." http://databank.worldbank.org/data/reports.aspx?Code=NY.GDP.PCAP.

Wu, Xiaogang, and Zhouni Zhang. 2010. "Changes in Educational Inequality in China, 1990-2005: Evidence from the Population Census Data." In *Research in the Sociology of Education: Globalization, Changing Demographics, and Educational Challenges in East Asia*, edited by Emily Hannum, Hyunjoon Park and Yuko Goto Butler. New York: Emerald.

Yang, Hua. 1999. "Work Roles and Norms for Teachers in Japan and the United States." In *Competitor or Ally? Japan's Role in American Educational Debates*, edited by Gerald K. LeTendre. New York: Falmer Press.

Zhou, Ning, Shui-Fong Lam, and Kam Chi Chan. 2012. "The Chinese Classroom Paradox: A Cross-Cultural Comparison of Teacher Controlling Behaviors." *Journal of Educational Psychology* 104 (4):1162-1174.

Zohar, Anat, Adi Degani, and Einav Vaaknin. 2001. "Teachers' Beliefs about Low-achieving Students and Higher Order Thinking." *Teaching and Teacher Education* 17 (4):469-85.