Influence of International Studies of Student Achievement on Mathematics Teaching and Learning

Vilma Mesa, University of Michigan-Ann Arbor, USA
Pedro Gómez, Universidad de Granada, Spain and Universidad de Los Andes, Colombia
UiHock Cheah, South East Asian Ministers of Education Organisation,
Regional Centre for Education in Science and Mathematics, Penang, Malaysia

Abstract

In this chapter, we present findings regarding the ways in which the results of international studies of student achievement have influenced the teaching and learning of mathematics in the classroom. We put forward a model of curriculum composed of four levels (global, intended, implemented, and attained) and four dimensions (conceptual, cognitive, formative, and social). This model allows us to describe the differences between two major international studies of student achievement—the Trends in the International Mathematics and Sciences Study (TIMSS) and the Programme for International Student Assessment (PISA)—and to situate the influences of these studies on classroom practice. Our search revealed that the question of how these studies have directly affected practice has not been systematically addressed. Although we found that some influences of the international studies on classroom practice do exist—for example, in the language used in public documents, in the localization of curriculum design, and in the impact of using imported textbooks—research on these influences has been conducted mostly in isolation, without any coherent plan. We use our curriculum model to propose a research agenda on three major areas: the impact of the notion of competency and the use of the studies' frameworks; curriculum control, design, and management; teacher preparation and development and textbook use.

Keywords: curriculum, TIMSS, PISA, international studies, curriculum

Our charge for this chapter is to present an account of efforts that have been made to take advantage of the information that international studies of student achievement offer to alter practice: the day to day of teaching and learning mathematics in classrooms. This chapter is complementary to Dossey and Wu's chapter D7, which speaks about the impact that these studies have had at the policy level.

Comparative education in mathematics is an old enterprise (e.g., Cairn, 1935; Young, 1900, pp. 3-4). Only after the 1960s did the efforts to investigate how students from different countries perform in mathematics become more systematic and collaborative, involving a larger number of nations and educational systems (Bottani, 2006; Husén, 1967; Robitaille & Travers, 1992; Schmidt & McKnight, 1995; Travers & Westbury, 1989). Although there are many arguments for and against participating in studies that contrast student attainment across countries (see, e.g., Bracey, 1998; Freudenthal, 1975; Husén, 1983; Keitel & Kilpatrick, 1998; Kilpatrick, 1971; Robitaille & Travers, 1992), there is an anticipation that the community will benefit from these studies, from the "research findings, the methods used in research, and [their] theoretical constructs" (Ferrini-Mundy & Schmidt, 2005, p. 169). Indeed, it is believed that these studies

> *can* serve as valuable sources of data and information against which educators in a given country *can* compare and contrast the curriculum, the teaching practices, and the outcomes attained by students in their own system. The possible impact of alternative curricular offerings, teaching strategies, administrative arrangements, and the like *can* be estimated efficiently by examining their implementation in other jurisdictions, even when the countries are quite dissimilar culturally or economically. … Achievement comparisons *can* also provide indications about what is possible … what *can* be accomplished. They *should* serve as a spur and incentive for improvement. (Robitaille & Travers, 1992, p. 707, emphasis added)

As the emphasized words suggest, these are hypothetical expectations that researchers have formulated about comparative studies. This chapter presents our efforts in assessing the extent to which these studies have indeed influenced practice, directly or indirectly; whether alternative "curricular offerings, teaching strategies, administrative arrangements, and the like" have indeed occurred, and if they have, whether they have resulted in the changes or real effects on learning and attitudes toward mathematics of students and on the teaching of school mathematics in the participating countries and elsewhere. We discuss the ways in which these studies have been "catalysts" for research that informs certain levels of practice (Ferrini-Mundy & Schmidt, 2005) and the ways in which they have been sensitive to cultural variation (Clarke, 2003).

We organized the chapter into three sections. We start by presenting a conceptualization of curriculum that allows us to organize our findings regarding ways in which the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International

Student Assessment (PISA)[1] have exerted influences at different levels. This section is followed by illustrative examples of these influences, which leads to a final section in which we propose research ideas that would move our community toward a better understanding of the actual impact that these studies can have on classroom processes.

Conceptualizations of Curriculum

An interesting feature of international collaboration is the need to clarify terms and concepts to make the work transparent. This was the case in writing this chapter. As our writing progressed, it became increasingly clear we were using different definitions for *curriculum*. From the Latin *currere*—to run—the word can refer to the sequences of courses that a student can take, the topics that are covered in a given grade, or the content, skills, competencies, and habits of mind that a person needs to acquire through schooling in order participate successfully in the society. The classical distinction between intended, implemented, and attained curricula (Travers & Westbury, 1989) was useful to describe how either notion of curriculum is transformed, but it did not differentiate other aspects that play significant roles in defining a curriculum. Thus, in this section, we propose a definition that will help us situate the influences that we found. We depart from a definition of curriculum encompassing only content or competencies; we define curriculum as a teaching and learning plan that can be described at different levels and that has different dimensions. We start with a description of how curriculum is understood by the studies that are the central to this chapter; namely, TIMSS and PISA.

Curriculum in the International Studies

The IEA studies have conceptualized curriculum as a tripartite model consisting of the intended, implemented, and attained curriculum (see Figure 1). The intended curriculum corresponds to the goals for learning mathematics that students are expected to attain, goals that may be established by national organizations (bureaus) or central ministries, states, departments, districts, or schools and explained through course outlines, official syllabi, and textbooks. Some systems differentiate goals according to types of students (e.g., Gymnasium, Hauptschule, and Realschule in Germany), others produce the textbooks that all children should use (e.g., Cyprus), and others use achievement tests and diagnostic assessments that children take at different stages of their schooling (e.g., South Korea) to define learning expectations. The implemented curriculum

[1] We refer the reader to the chapter by Dossey and Wu in this volume, which provides details on the history of these two studies. A third group of studies, those conducted by the UNESCO, focus on Latin America and Africa. These comparative regional studies of student attainment are part of the "Education for All" initiative, which seeks to have all the world's primary-aged children enrolled in school by 2015. The studies collect information on third- and sixth-grade students' attainment in mathematics, science, and reading. In addition, there are questionnaires for students, teachers (content coverage and pedagogy), principals, and parents. The Sub-Saharan study also tested teachers' knowledge of the content on which the students were tested. There are very few reports (Bonnet, 2008; V. Lee & Zuze, 2010; M. Saito, 2010; Valdés et al., 2008), and they are mostly descriptive, although some seek to interpret variability using the contextual variables collected. The Centre for Innovation in Mathematics Teaching at the University of Exeter has conducted two longitudinal studies of mathematical achievement for Innovation in Mathematics Teaching in the United Kingdom, the Kassel Project (http://www.cimt.plymouth.ac.uk/projects/kassel/default.htm), and the International Project on Mathematical Attainment (http://www.cimt.plymouth.ac.uk/projects/ipma/default.htm). The data have been given to heads of the departments of participating schools, but there is no information on their effects in the classroom. For examples of reports in Singapore, see Kaur and Yap (2009) and Kaur, Koay, and Yap (2009).

corresponds to what is actually taught to students in classrooms by teachers; it refers to the interpretations of the intended curriculum made by teachers, who are directly responsible for helping students learn. The attained curriculum corresponds to what students demonstrate that they have learned (via standardized tests, including the international studies tests, and classroom assessment) as a consequence of the teaching received.

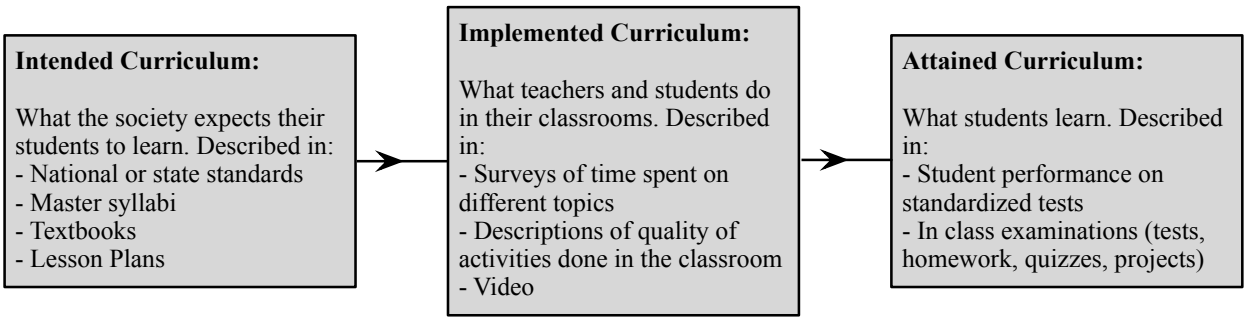

Figure 1. Different versions of the curriculum; adapted from Travers and Westbury (1989).

This view of curriculum presupposes different degrees of expertise in curriculum design at each level and assumes that teachers and students have little agency in designing the curriculum, accepting and agreeing with the information from the previous level. In particular, teachers are expected to use what is given to them (goals for society, goals for schooling, textbooks, official syllabi) to make decisions about what is best for the students they have. This view of curriculum suggests that the studies are expected to exert a major influence at the policy level; that is, that the results will be used to shape intentions of the whole system that in turn will influence what will happen in the classrooms and with students, as those intentions get transformed into action.[2]

The TIMSS mathematics framework developed for the 1995 study consists of three aspects: *content domains* (numbers; measurement; geometry; proportionality; functions, relations and equations; data representations, probability, and statistics; elementary analysis; and validation and structure), *processes or performance expectations* (knowing, using routine procedures, investigating and problem solving, mathematical reasoning, and communicating), and the *affective outcomes or perspectives* of school mathematics and science (attitudes, careers, participation, increasing interest, and habits of mind). The participating countries agreed upon the content domains, whereas the performance category was "aligned to the [U.S. National Assessment of Educational Progress NAEP]'s concepts of mathematical abilities and mathematical power, both with roots in the [National Council of Teachers of Mathematics] NCTM standards" (Mullis, 1999, p. 15). Note that this definition encompasses more than content, although content is a main component of this framework.

[2] We acknowledge that this model excludes the possibility that after implementation, changes can be made to the intention of the curriculum. This local view of curriculum transformation is not necessarily accounted for in this model.

PISA departs from a content-focused approach. Indeed, the first paragraph of the PISA assessment framework states a contrasting view:

> The aim of the OECD/PISA assessment is to develop indicators of the extent to which the educational systems in participating countries have prepared 15-year-olds to play constructive roles as citizens in society. *Rather than being limited to the curriculum content students have learned,* the assessments focus on determining if students can use what they have learned in the situations they are likely to encounter in their daily lives. (OECD, 2003, p. 24, emphasis added)

The PISA mathematics framework uses three components to describe a domain to be assessed in relation to the problems that students are expected to solve: (a) the situations or contexts in which the problems are located, (b) the mathematical content that has to be used to solve the problems, organized by certain overarching ideas, and, most importantly, (c) the competencies that have to be activated in order to use mathematics to solve real world problems. *Content* is organized into four "overarching ideas: quantity, space and shape, change and relationships, and uncertainty" (pp. 35–37). Mathematical competence is described in terms of eight specific *competencies*: "(a) thinking and reasoning, (b) argumentation, (c) communication, (d) modeling, (e) problem posing and solving, (f) representation, (g) using symbolic, formal and technical language and operations, and (h) use of aids and tools" (pp. 40–41). The cognitive activities encompassed by these competencies have been structured in three *competency clusters*: "(a) reproduction, (b) connection, and (c) reflection" (pp. 41–49).

The PISA framework introduces the idea of *mathematical literacy* in order to emphasize a functional view of school mathematics: tools that should enable students to make well-founded judgments and be useful in students' lives as citizens. Thus, conceptually, PISA seeks to assess the extent to which schools have indeed prepared students for participation in the society, whereas TIMSS assesses the extent to which students show proficiency with particular mathematical content at specific points in their school life. We take these differences into account by we situating the two studies as attending to two different aspects of what we will propose as curriculum.

Curriculum in This Chapter

So far, we have identified three levels for the curriculum: the intended, implemented, and attained. Within the intended level, we distinguish several sublevels: The first is the education system that is particular to each individual country and in turn can be differentiated by national or federal, regional and state, or district mandates or guidelines; second, at the school level, we include the plans schools use, perhaps adapted or adopted directly from federal, national, or regional mandates or guidelines; and third is the classroom level which refers to the plans that teachers create for teaching particular lessons, using their institutional, district, state, or national guidelines or other sources. In contrast to these multiple sublevels, the implemented curriculum manifests mainly in the classroom,[3] whereas the attained curriculum manifests mainly through

[3] Other versions of the curriculum (e.g., hidden, null) manifest also at the school level. We do, however, concentrate only on the curriculum within the classroom, as that was the task for this chapter.

students' performance, both through class assessments and through standardized tests when available. These levels are represented in Figure 2. We illustrate the flow of the suggested influences by a gray arrow in the background.

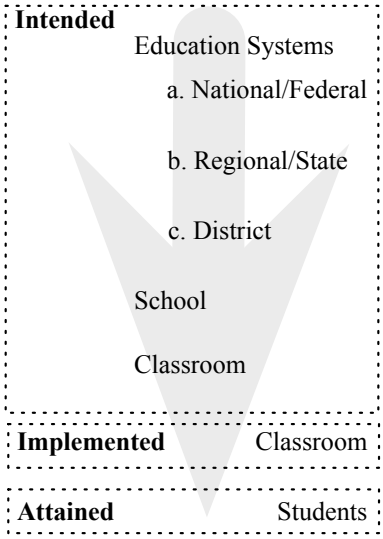

Figure 2. Levels of the curriculum considered in this chapter.

Concurrently with these levels, the curriculum is composed of four dimensions: conceptual, cognitive, formative, and social (Rico, 1997), each of which deals with four fundamental and interrelated questions: "what is knowledge, what is learning, what is teaching, and what is useful knowledge" (p. 386).[4] The *conceptual* dimension refers to content and topics that are specific to a given discipline; it defines those elements particular to a discipline (e.g., mathematics, the sciences) that are a synthesis of historical and cultural traditions; this dimension is informed by epistemology and the history of mathematics and defines larger cultural aims. The *cognitive* dimension refers to learning and the learner, and deals with understanding what learning is, how it happens, and how do different people learn; it also has particular manifestations depending on a given discipline; it is directly informed by learning theories and defines specific expectations, development, and learning aims. The *formative* dimension refers to teaching and the teacher; it deals with aspects such as what teaching is and what mathematics teaching is; it specifies practices that are believed to be useful for teaching (e.g., planning, differentiating instruction) and it provides the basis for generating programs for future and practicing teachers; it is informed by pedagogical theories and defines formative aims. The *social* dimension refers to the value that the society places on the utility and usefulness of the mathematical knowledge; it deals with questions such as

> Which criteria determine the mathematical capacity of an individual, which instruments are used to judge the mathematical capacity of an individual? What social mechanisms support that judgment? How and with what criteria are

[4] Here we note that other conceptualizations (e.g., Beyer & Liston, 1996; Stark & Lattuca, 1997) can be embedded within this definition.

> teachers' capacity and curriculum materials judged? [And] which criteria are used to assess the effectiveness of a curriculum? (p. 385)

This dimension is informed by sociology and other disciplines, and it defines social aims.

Hence, curriculum can be conceived as involving levels—from the national educational system through the classroom—and four dimensions, as described above. At a given level, each dimension of curriculum acquires a specific meaning. At the classroom level, which is of particular interest for this chapter, the conceptual dimension of the curriculum refers to the mathematics topics that configure the content of a given grade or teaching unit; the learning goals of such grade or teaching unit are the expression of the cognitive dimension of the curriculum; the formative dimension of the curriculum refers to the teaching methodology set up for the grade or teaching unit; finally, the assessment instruments and criteria selected for the grade or teaching unit configure the social dimension of the curriculum at the classroom level.

For the purposes of this chapter, we include in our model two additional elements. First, there is a global level, which in the abstract refers to the possibility of having a curriculum that transcends individual systems and that could operate, in fact, as a global curriculum—a curriculum that is common to many education systems. Although not curricula themselves, the frameworks of the international studies can be seen as part of a global level because they represent the agreements across several education systems and nations towards a common set of content and learning expectations that will be used to assess students (Clarke, 2003). Second, in the attained level we focus on and distinguish between students' performance as assessed via national standardized tests or via tests produced by the international studies.

Because each of these dimensions manifest at different levels—at the level of classroom practice, at the administrative level of a particular school, and at the larger level of an educational system—we have combined these definitions with the tripartite version of the curriculum to generate a matrix that allows us to situate the different manifestations of curriculum, and we have added several levels to better represent the different influences that we were able to identify (Figure 3).

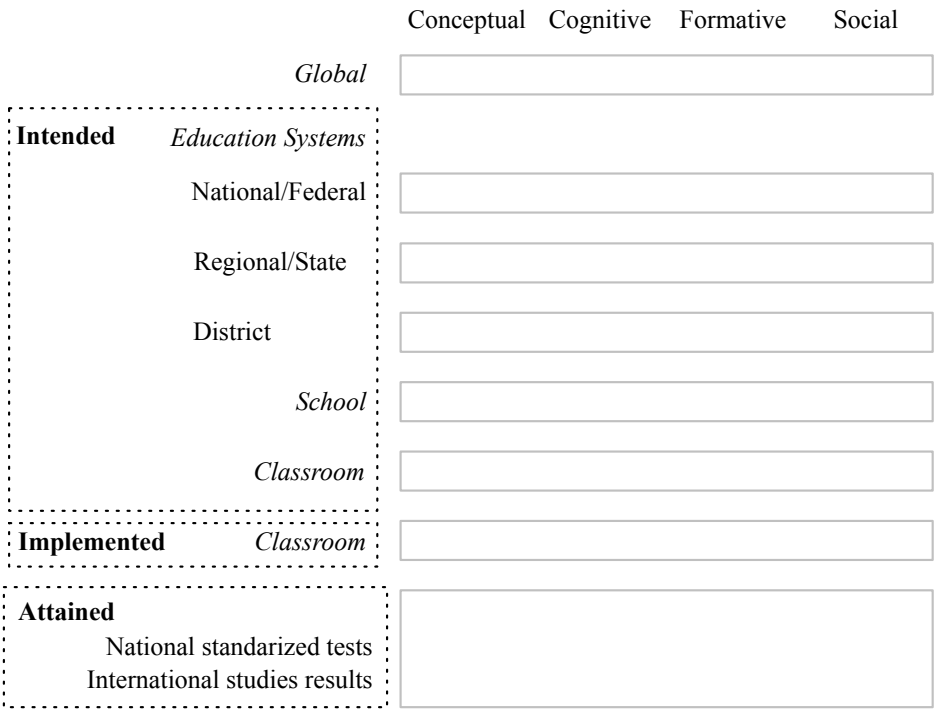

Figure 3. Levels and dimensions of the curriculum as understood in this chapter.

Figure 3 allows us to situate different documents and aspects of the international comparison studies. In spite of their differences in emphasis, we situate the frameworks of TIMSS and PISA at the global level; and because both TIMSS and PISA are concerned with what students have learned and suggest content (TIMSS) and competencies (PISA) that are considered relevant, these studies relate to the conceptual and cognitive dimensions only. The NCTM (2000) *Principles and Standards for School Mathematics* and official syllabi in individual countries with centralized curricula can fall also into the conceptual and cognitive dimensions. The NCTM (1991, 1995) also published professional teaching standards and assessment standards. These documents explicitly describe what is knowledge, what is learning, what is teaching, and what is useful knowledge, thus spanning all the dimensions of the curriculum at the national level. The Common Curriculum Standards, recently released in the United States, as result of an agreement among 48 governors,[5] are an attempt at a definition of content and learning outcomes at a national level, and it would be situated at the conceptual and cognitive dimensions of the national level. Some countries (e.g., Spain) include specific mandatory norms for each of these dimensions at the national level, whereas others (e.g., Colombia) only give guidelines, suggestions, at the cognitive level (learning expectations). In Asian countries, the ministries of education are empowered to oversee the development of their national curriculum, which all schools follow. The national curricula tend to attend primarily to content, but over the last 10 years more emphasis has been placed on mathematical processes. In some countries, there are national teachers' guides that include ideas about how students can learn or how topics should be

[5] To date, 40 states plus the District of Columbia have adopted these standards.

taught (e.g., Cyprus, South Korea). In some countries there might be information about these areas, but these appear o\in teachers' editions of student textbooks and might be totally unregulated (e.g., the United States). With this definition as a backdrop, we turn now to the task for this chapter.

Influences of Comparative Studies

To address the question driving this chapter, we searched several sources to acquire reports, documents, and articles related to the topic. We approached key informants in academia and in ministries or bureaus of education in several countries seeking information about possible studies that would speak about the impact of these international comparative studies at the classroom level. These key informants were able to identify researchers and sources from which we sought further information. Library searches provided dissertations, conference presentations, reports, books, and journal articles.

The majority of the studies that we found, including the most recent ones, were primary and secondary analyses of the results of the international studies. Studies in this group mined the richness of the data sets in order to establish connections between the variables collected at country, school, classroom, teacher, and student level with the scores obtained. Because the international comparison studies have been increasing in their sophistication, more powerful analyses have been conducted. These studies did not document changing practice or study possible changes to practice; rather, they sought to understand and explain the sources of differences in scores between and within countries. For this reason, we did not include these studies. We also excluded studies that mentioned students' performance (either high or low) to justify attending to a specific issue of educational interest, but we kept studies in which substantive elements of the studies (the frameworks, the test items, or the findings) were used.

We discuss here the only study that was very close in nature to the charge that we had for this chapter, and which anticipated our findings: There is very little documented evidence that the studies have reached classrooms in a significant way. The study, by Saracho (2006), describes in detail some of the policies taken by countries to improve their PISA results as consequence of their participation in PISA 2000 and 2003. The study was commissioned by the OECD in Latin America and conducted by the Mexican foundation IDEA [Implementation, Design, Evaluation and Analysis of Public Policies]. Its sources of information were documents, secondary sources, and news, together with phone interviews and questionnaires answered by specialists, public officers, and university professors in the countries involved. The 14 countries analyzed (Austria, Belgium, Brazil, England, France, Germany, Ireland, Luxembourg, Mexico, the Netherlands, Norway, Poland, Spain, and Sweden) were selected according to the availability of information, the "size" of their reactions to the PISA 2000 results, and their relevance to the Mexican case. Even though the report focused on policy reactions, it included issues related to the influence of PISA in the classroom and provided an overview of how each country reacted to the publication of the 2000 and 2003 results. Even though Saracho was interested in the impact of PISA in schools and classrooms, he states explicitly that he did not find any evidence of this impact. Nevertheless, he documented some reactions that could, indirectly, influence the teaching and

learning of mathematics in the classroom, and where pertinent, we included these findings in our review.

A second group of documents used elements of the international comparison of achievement studies—their frameworks, the released items, or the videos—to induce some change and to test the impact of those changes. Included in this group are studies that capitalized on the logic of "data-driven" analyses, by which changes at the local level can be initiated and evaluated using information that is of interest to the participants. We included these studies in the review even though they do not assess the impact of the changes.

A third group of studies reported actions taken looking at what successful countries do and consequently adopting salient elements believed to be directly associated with that success. The three most prominent cases are the shift of the locus of control for curriculum development from higher to lower levels, the incorporation of Japanese lesson study, and the adoption of Singapore mathematics textbooks. We included these studies, and in the case of the textbooks, we included available information about their impact in the classroom.

Thus, what we offer is an inventory of projects, activities, and programs whose impetus can be traced to either PISA or TIMSS and that may have a direct or indirect connection to classroom practice. We believe that the inventory is not comprehensive, as many reports may exist that are unavailable. The majority of the initiatives have not been formally studied to assess their impact: How are they used? By how many people? What do teachers and students perceive about their effectiveness? And how are they related to student and teacher performance? Nevertheless, we believe that they merit consideration for their potential to generate substantive research in the future. Because both PISA and TIMSS have tests on science, when available we included works conducted in science education because they suggest possible uses that may have been given in mathematics, although we have not been successful in finding documentation of these uses.

We provide selected examples, classified by the main categories of influences that we were able to document. Within each category, we include details of how it was exemplified in individual countries as a consequence of either PISA or TIMSS. The readers will notice that most of the examples related to PISA focus on two regions, Europe and Latin America, and that most of the examples related to TIMSS focus on South East Asia and the United States, which reflect the regional impact of these studies. Dossey and Wu (this section) note that in terms of reactions to the studies, some countries (usually the high-performing ones) assume a congratulatory approach, and in general may not worry about making specific changes in their own systems; other countries assume an indifferent position, taking the results as yet another indicator of performance, with little interest in making changes; a third group of countries use their students' performance as a justification for engaging in activities that would alter practice. These last countries provide the examples of influences that we report here.

Our conceptualization of curriculum allows us to better understand and situate the influences that we have identified in our search. We start with the model presented in Figure 3 and use numbered arrows to depict those influences (Figure 4). We have identified two sets influences by looking at where they originate: at the global level (Influences 1 and 2) or at the attained level of the international studies (Influences 3 to 7). Each influence starts at a particular dimension of these levels and ends in more than one level and dimension, illustrating that a specific aspect of the international studies can influence different features of a country's mathematics curriculum. The first set of influences start at a global level, with the frameworks proposed by the international studies. The two influences that can be tracked down to each level of the intended curriculum are as follows:

1. competencies and standards in official documents; and
2. the adaptation and adoption of the frameworks and the items used in these studies.

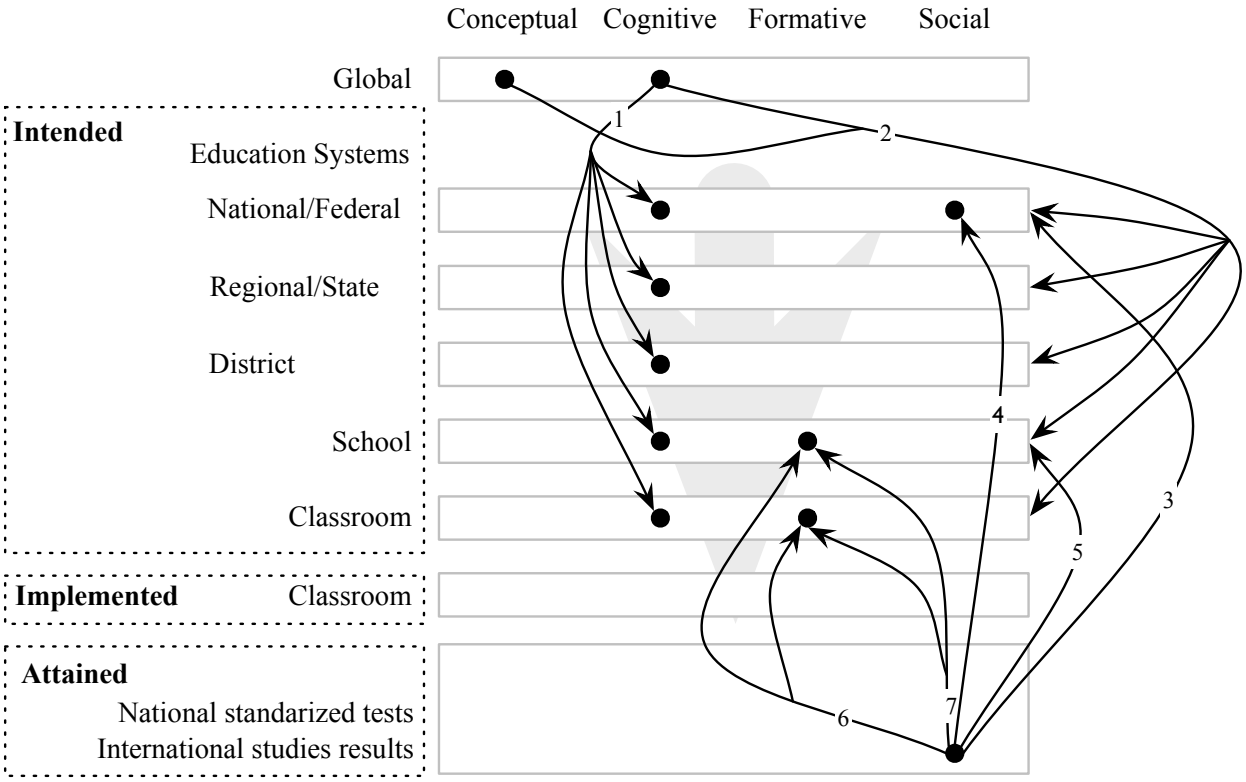

Figure 4. Influences of international studies on different aspects of the curriculum.

Influence 1 starts at the cognitive dimension of the global level; it represents how the frameworks that have been used by the international studies, concretely PISA's, have had an effect on the cognitive dimension at all the other levels of the intended curriculum via the notion of *competencies*. In Spain, for example, a number of documents have been written at all the levels of the intended curriculum, targeting student learning of these competencies. Influence 2 starts at both the conceptual and cognitive dimension of the global level to show how the content and learning expectations in these studies have been used as benchmarks for standards and attainment, influencing various levels of the intended curriculum.

The second set of influences corresponds to those that we could track as consequence of the results of students' performance in the tests. The influences originate at the attained level, but they refer in particular to the results of tests generated by the international studies, rather than to the results of tests created by individual education systems), and go "up" to the intended level. We identified five different types of influences:

3. localization of curriculum design;[6]
4. national changes in assessment processes;
5. the implementation of cycles of data analysis to design, enact, and assess changes;
6. professional development and teacher education; and
7. use of textbooks from other countries.

In Figure 4, we represent Influence 3, related to localization of curriculum design, with Arrow 3 from the social dimension of the attained level (results of the international studies) to the intended national/federal level expressing the influence that these results of the studies have had, in some countries, on giving more freedom to schools in designing all the dimensions of the mathematics curriculum, whereas Arrow 4 represents influences of the results of the tests on assessment processes at the intended national level. Arrow 5 represents the influence of the results of the tests on curriculum on all the dimensions of curriculum at the intended school level, whereas Arrows 6 and 7 represent influences of the results of the tests on teaching methods both at the school and classroom intended levels.

The reader will, by now, have noticed that the arrows start at strategic points from which the influences could be expected according to the conceptualization of the international studies—a global, overarching level, and with their outcomes—and that the arrows end at various places in the intended level, none of them reaching the implemented level. This is in itself an important finding for us. There have been no systematic investigations on how these studies have reached the classroom.

In the next two sections, we describe these two sets of influences. Readers will note that the sections are *uneven* in the amount of information provided. This unevenness is a consequence of the types of documentation that we were able to secure; the differences in *depth* reveal, indirectly, the amount of information available about each type of influence.

From the Global Level to the Intended Curriculum

In this section, we describe the influences we established in Figure 4 that originate in the global level of or our curriculum model and that affect the intended curriculum. We briefly describe the studies and documents that illustrated the ways in which these influences were operating.

---

6 We use "localization of curriculum design" to refer to what is sometimes known as the pedagogical aspect of school autonomy.

Competencies and standards in official documents (Influence 1). An important influence of these studies in several countries (mostly Latin American and European) has been on the language used in public and political discussions of education. The notion of *competencies* has gained currency, presumably because it is the word used in the PISA frameworks. The word *competencies* has been included in several national documents and in a European directive (Education Council, 2006), where it refers to competencies for lifelong learning. Competencies are meant to go beyond specific school content or skills:

> The assessment of student performance in selected school subjects took place with the understanding, though, that students' success in life depends on a much wider range of competencies. … A competency is more than just knowledge and skills. It involves the ability to meet complex demands, by drawing on and mobilising psychosocial resources (including skills and attitudes) in a particular context. For example, the ability to communicate effectively is a competency that may draw on an individual's knowledge of language, practical IT [information technology] skills and attitudes towards those with whom he or she is communicating. (OECD, 2003, pp. 3–4)

In reality, and perhaps unsurprisingly, a noticeable change in official rhetoric might have little impact on actual classroom practices, as a number of concrete cases suggest. In Europe, the 2006 directive for the development of competencies for lifelong learning (Education Council, 2006) recommends "that Member States develop the provision of key competences for all as part of their lifelong learning strategies, including their strategies for achieving universal literacy" (p. L 394/11). Since then, most European countries have introduced the idea of competency in their curricula, giving rise to what can be called a *competency-clash*:

> Not only do different notions of competence clash with each other, but the new educational 'gospel' of standards, competencies, and outcome-orientation is at odds with [the countries'] traditions, such as content-based curricula and input-orientation (i.e., regulating the structures, processes and conditions of teaching and learning). (Sloane & Dilger, 2005, cited by Ertl, 2006, p. 628)

Spanish authorities explained the country's results in PISA by arguing that because the Spanish curriculum was centered in content, Spanish teaching focused on teaching such content (Arias, 2006), and hence Spanish students were not being prepared for the PISA tests. One action taken to address this concern was to distribute translations of the PISA's executive summary and framework to all schools nationwide, with which the idea of competency entered the Spanish educational discourse. Rico (2011) has noted that "the PISA assessment model has been determinant for the Spanish educational system" (p. 4). In fact, the new national curriculum seeks that students improve their command of basic competencies and introduces *competency* as the basis for curriculum innovation (Blanco & Rico, 2011; Maestro, 2006; Rico & Lupiáñez, 2008, pp. 104–106, 208–213). This was also the case in the most autonomous regions in Spain (Ferreras, 2006; Graña, 2006), with many deciding to participate in future PISA tests with a representative sample or a census of their students (J. A. Gómez, 2006), making the need to prepare students on those competencies more pressing. Naturally this shift in language has also had an impact on the preparation of teachers and the design of textbooks.

The results of PISA 2000 marked a milestone in Germany because the country ranked 21st among the 31 countries that participated in the study. The reactions to these results were known as the "PISA shock. " No other country had a bigger reaction to the results (Gruber, 2006). Besides the media reaction, there was also a reaction in academic and political circles. Many articles and books were written on the results and on possible strategies to improve them; teacher preparation and development courses were offered, and several educational policies were implemented. In particular, new standards were introduced for middle school based on the idea of competency, together with standardized tests assessing students' performance on those standards (Neumann, Fischer, & Kauertz, 2010).

Colombia, which has participated in TIMSS since the first study in 1995, and will take part in PISA for the first time in 2012, adopted the notion of standards in the late 1990s, along with other Latin-American countries (Palamidessi, 2006). The National Ministry of Education published in 2006 the *Basic Standards of Competencies* (Ministerio de Educación Nacional [MEN], 2006). This document introduces some *general processes* that are similar in nature to PISA's competencies and sets standards for pairs of grades, organized by types of mathematical knowledge—for instance, communicating; modeling; and formulating, treating, and solving problems (MEN, 2006, pp. 51–55). The Colombian standards are supposed to contribute to the development of competencies. This document has been distributed to most schools and teachers in the country. The ministry has organized conferences and teacher-training events with the purpose of explaining the standards and the idea of competency. But, "teachers do not use the standards" (Monica López, personal communication, March 16, 2010). Because schools are autonomous in designing and developing a curriculum based on these standards, and most public schools do not provide or require textbooks, the responsibility of curriculum design and development has fallen onto teachers, who usually produce so-called teaching guides for implementing instruction in their classrooms. A study by P. Gómez and Restrepo (2011) found that the word *competency* seldom appears in the school planning for any given grade. Furthermore, when the word appears, it is interpreted in many ways, usually differently from what the *Basic Standards of Competencies* document intended and most notably to refer to learning goals for a specific content topic and grade level, which is contrary to the original meaning of *competency*.

In contrast, the use of standards and competencies does not seem to be a trend in Asia. Singapore, Hong Kong, and Japan, three of the Asian countries with outstanding results in TIMSS and PISA, do not refer to standards or competencies in the mathematics curriculum. Singapore's mathematics curriculum is based on its problem-solving curricular framework that focuses on five key components: skills, attitudes, concept, metacognition, and processes. This framework has been used since it was first implemented in 1990. Since then, there have been some changes to the curriculum in order to keep abreast with the shift of global trends towards a knowledge-based economy. For example, in 1998, there was an increased focus placed on thinking and processes as well as a trimming of some content (Kaur, 2003). However, there was no specific reference to TIMSS or PISA. There have however been numerous secondary analyses using the TIMSS data, which indicate that the TIMSS has been used as an international benchmark to gauge the success of the mathematics curriculum in Singapore (e.g., Kaur, 2002, 2005, 2009; Kaur & Pereira-Mendoza, 2000a, 2000b).

Likewise, Hong Kong's curriculum does not directly refer to standards or competencies but rather focuses on the main aims of developing interest, communication, lifelong abilities, numeracy, spatial skills and understanding, and the acquisition of basic skills (Curriculum Development Council, 2000). In Japan, the mathematics curriculum has been undergoing a process of revision since 2005. One of the main purposes of the revision is to address weaknesses as indicated by the PISA results, such as students having difficulty writing problems that require thinking, decision making, and expressing, and lacking motivation to learn (Ministry of Education Culture Sports and Science, 2010). The new revised course of study in the Japanese curriculum aims to prepare students through the acquisition of basic knowledge and skills and the development of abilities to think and to express ideas mathematically. Although the new curriculum indicates that there were considerations made based on concerns arising from the PISA results, it was not structured according to standards and competencies.

We found one Asian country, Indonesia, which explicitly mentions standards and competencies. The Indonesian mathematics standards use competencies to map the curriculum. The curriculum lists core competencies and corresponding student outcomes that can be used to indicate the achievement of these competencies (Departemen Pendidikan Nasional, 2003). There were no clear references to the PISA frameworks in the curriculum, but because Indonesia has regularly participated in the PISA studies since 2000, the use of standards and competencies in the curriculum could perhaps serve as some evidence of the effects of PISA on the development of the mathematics curriculum.

Adaptation and adoption of frameworks and items (Influence 2). Besides the influences noted above, in terms of the language used and how it has affected curriculum standards, we have anecdotal evidence that the TIMSS frameworks have influenced the redesign of U.S. curriculum guidelines at the state level. The extent of this influence is unknown, although one dissertation (Landry, 2010) looked at the alignment between the TIMSS or PISA frameworks and the content standards in several states in relation to the high or low performance of countries and states. Landry found that, from the content point of view, most of the standards in most of the states in the United States cover a wide variety of topics that were repeated year after year—a finding that is consistent with results of the curricular analysis for TIMSS conducted by Schmidt and colleagues (Schmidt, McKnight, Valverde, Houang, & Wiley, 1996), which showed that in the United States topics tend to enter and stay longer in the curriculum than in other countries. Landry also found that high levels of alignment to the curriculum in high-performing countries and their assessments did not absolutely equate to high performance at the state level.

We have found that, in science education, researchers have used sections of the published questionnaires to create new instruments that are used in pretest-posttest designs to establish effectiveness of interventions (O. Lee, Deaktor, Enders, & Lambert, 2008; O. Lee, Deaktor, Hart, Cuevas, & Enders, 2005; Shymansky, Yore, & Anderson, 2004). However, we could not locate similar papers produced by the mathematics education community.

From Attained to Intended Curriculum

In this section, we discuss the second set of influences: documents and reports that speak about the way in which the results of the international studies—the scores that students obtained—have influenced different levels of the intended curriculum.

Localization of curriculum design (Influence 3). Most countries that perform well in PISA give local authorities the freedom to adapt curriculum (Schleicher & Shewbridge, 2008, p. 20). In PISA, school autonomy for curriculum is measured with an index of school responsibility for curriculum and assessment. This index is derived from categories that school principals classify as being the responsibility of schools (establishing student-assessment policies, choosing textbooks, and determining which courses are offered and the content of those courses; OECD, 2010, p. 68). For instance, in reading literacy, school autonomy has a statistically significant positive relationship with students' performance (OECD, 2005, p. 71). However, when these findings are controlled for student and school-level factors, the relationship between school autonomy and students' performance is weak (p. 72).

Nonetheless, these results have led countries to push for more school autonomy, assuming that such change might have a larger effect on students' learning than other aspects of schooling. In Germany, for example, the decision was taken to give more authority to the states, with the central bureau prescribing a few core ideas (rather than all the curricular content), giving schools the autonomy for finishing the curriculum (Ertl, 2006), and requiring at the same time "external and internal assessments (of schools and students)" (p. 626). Such a shift has resulted in teachers in some states becoming responsible for developing the whole curriculum for their schools, thus localizing the design of the curriculum. Although it is too early to establish the impact of such reform (Kotthoff & Pereyra, 2009), in Germany, the perception is that the process is problematic, because German teachers are not qualified to assume responsibilities as curriculum designers and developers. Similar claims about the positive impact of giving more autonomy to schools to design curriculum have been made in Australia (Liberal Party of Australia, 2010).

Spain is also moving towards localization of curriculum design. In early March 2011, two major Spanish autonomous regions (Madrid and Catalonia) decided to transfer the authority of deciding 35% of the curriculum to schools—pending approval by the autonomous region (Alcaide & Álvarez, 2011; Bassells, 2011)—as a means to improve teaching:

> The autonomy of schools for looking for better solutions for their context is being recognized recently as one of the most recurrent strategies for improving teaching; that is what can be assumed, for instance, from several OECD studies, like the PISA report. (Alcaide & Álvarez, 2011, p. 38, trans. by authors)

The movement towards localization of curriculum design is not necessarily a consequence of the studies. Colombia is an example of curriculum localization that occurred prior to its participation in the international studies. The 1994 Colombian Education Law established that "the autonomy is a consequence of the will for differentiating each educational community, paying attention to different needs and expectations; it seeks that each educational institution educate citizens that can solve the problems of their own environment" (MEN, 1994, article 77, trans. by authors).

Assessment (Influence 4). A common justification for low performance on the PISA examinations has been that students have little experience with standardized tests. According to Saracho (2006), although the influence of PISA has been limited, mainly because of the absence of a mechanism to inform schools about individual results, schools and teachers are feeling more inclined to agree about the need of accountability of results and are more willing to have standardized test results disaggregated at the school and classroom level, something that "was an alien notion until recently" (p. 27, see also Dossey & Wu, this section).

The case of Poland illustrates an indirect influence of the results of PISA on the definition of a standardized test, MATURA, which is administered at the end of the secondary school and serves both as a diagnostic tool of the performance of the education system and as an admission test for postsecondary education. The annual character of the test allows for tracking the effectiveness of changes and the identification of areas that need further change. The Polish results in PISA were useful in overcoming the public's and the schools' skepticism about the usefulness of the proposed test (Saracho, 2006).

Danish students obtained results that were close to the OECD average, which raised concerns in Denmark given its high expenditure in education. A government study named the lack of a standardized assessment culture in the country and low levels of satisfaction with the education system as possible reasons for the outcomes (OECD, 2004). In Brazil, another country whose students scored low in PISA, the government recently introduced a new system of periodic standardized assessment of students' performance (Saracho, 2006).

Some Spanish regions (e.g., Andalusia) have introduced, by decree, diagnostic assessments in schools based on competencies (Junta de Andalucía, 2007). These diagnostic assessments have no influence on students' grades and are designed and implemented by schools. Their purpose is to establish the level at which students have developed basic competencies and help schools make decisions about their curriculum. Likewise, the Department of Education of the Autonomous Region of Navarra is promoting self- and external-assessment in schools. The department anticipates that external assessments will encourage new teaching practices and that the PISA framework can be used to enrich the curriculum. In order to ensure coherence between the school assessments and the PISA tests, the department has defined, for each knowledge domain, specific guidelines for improving reading competencies at all grade levels, and published standards in language, mathematics, and science based on the PISA framework. Regarding problem-solving competency in mathematics, the guidelines proposed that teachers use problems modeled on the PISA items during instruction, a suggestion in the formative dimension at the regional level. A web page has been set up with information concerning PISA with proposals for improvement in each area (Ferreras, 2006). This is a first step in the data-driven approach to reform that we discuss next.

Data-driven approaches to reform (Influence 5). A report by the U.S. National Research Council (1999) suggested that a way to capitalize on the results of these large-scale studies of

achievement at the ground level (namely, schools and classrooms) is by using local results to initiate a process of self-reflection. This report was a direct outgrowth of the TIMSS study, and its focus was on the training of professional developers who could direct and assist schools in initiating the data-driven, self-reflection process. The process starts with teachers (schools or districts) getting acquainted with TIMSS, its design and its findings, analyzing the implications in their own contexts, and finding a particular focus of attention (e.g., student achievement or curriculum alignment). Teachers (schools or districts) can analyze data about student achievement in the TIMSS test or carry out a content analysis (guided by the TIMSS content analysis process) by which they can determine the extent to which the content is aligned with the framework that guided the design of the TIMSS test. This analysis should lead to the identification of a specific problem to work on: an area in which there is low student performance combined with an analysis of the content coverage for that area, and may result in a suggestion for a change (e.g., emphasize or de-emphasize instruction on certain topics). The change is monitored in order to determine its effectiveness, and this analysis starts the cycle again (Figure 5).

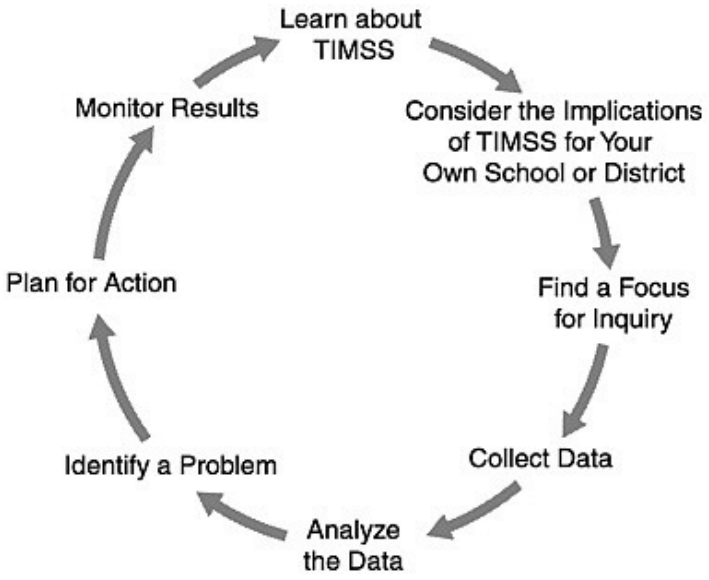

Figure 5. The data-driven inquiry process (Source: National Research Council, 1999, p. 398).

An important step in this cycle is benchmarking: testing a large sample of students to generate a baseline for later comparisons (similar to the Andalusian diagnostic tests). According to the report, several efforts emerged from this initial work (e.g., the Chicago area's First in the World Consortium, the Lake Shore school district in Michigan, and an urban school in Patterson, New Jersey). We found some documentation of these efforts in newspaper articles (e.g., Dunne, 2000) and descriptive reports (Kimmelman, et al., 1999) but little in terms of assessment of their outcomes.

We found one dissertation that studied how a group of teachers immersed in a professional learning community in one school took advantage of this cycle of data analysis to improve their practice and collected data on the impact of this process. In her dissertation, Figueroa (2008)

used information from the analysis of lessons from the TIMSS Video Study (Stigler, Gonzalez, Kawanaka, Knoll, & Serrano, 1999; Stigler & Hiebert, 1997, 1999) to generate a modified version of a TIMSS lesson plan[7] that was piloted and used by several elementary teachers in a school in Arizona. The teachers engaged in a 2-year process of learning about the process of creating the lessons, presenting them, getting feedback, and redesigning the lessons for a new application.[8] Several elements of a lesson study process were present, including time that allowed the teachers to learn the method, seek resources, plan lessons, and observe each other teaching. In this very prescribed process, the observers were trained to take notes on the percentage of students who could remember the strategies taught, were engaged in the problem-solving process, articulated the strategy, applied the strategy, were assessed for mastery/nonmastery of the learning strategy, and could relate the activities to the learning strategy. An evaluation of the impact of this 2-year initiative followed a pre-post-test design with a group of 65 fourth graders in the school. The analysis revealed a statistically significant increase in the posttest scores with respect to the pretest scores and a statistically significant association between time of test and categories of performance in the state test. The author suggests that the method may be useful in helping districts improve their students' scores on state tests. Unfortunately, this study has a problematic design that does not control for students' prior knowledge, for other students' characteristics, or for the quality of implementation of the lessons. The process relied on external funding, and there were no indications of attempts to sustain the effort.

In the United States, a large study led by William Schmidt and Joan Ferrini-Mundy from Michigan State University, Promoting Rigorous Outcomes in Mathematics and Science Education (PROM/SE) might provide information about the effectiveness of this approach for reform. PROM/SE is an 8-year grant that uses assessment of students and teachers to improve standards and content coverage, simultaneously building capacity with teachers and administrators. It involves over 300,000 students and 18,000 teachers in two states, Michigan and Ohio. The reports produced to date apply the logic model used to design TIMSS, capitalizing on many of their analytical strategies to deal with those data (Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE], 2006a, 2006b, 2008, 2009a, 2009b, 2009c). In addition, the reports illustrate vividly the kind of educational system that exists in the United States, one characterized by extreme variation that leads to substantial inequalities, not only between districts but within schools as well, variation which appears to be strongly determined by the differential access to resources (economic, cultural, and intellectual) of the community.

Teacher education and professional development (Influence 6). An example of influence on teacher education comes from Germany, as a reaction to unsatisfactory results in PISA (Ertl,

[7] The TIMSS adapted format lesson consists of nine components: "Listing math standard, competency and performance objective; Proposing a guiding mathematical question; Brainstorming strategies students will use to solve the question; Describing in detail steps teachers have to follow to implement the lesson; Modeling the strategies in class; Giving guided practice; Checking for understanding; Closing; Giving resources" (Figueroa, 2008, p. 64).

[8] The process itself consisted of 17 stages expanding on each of the 8 steps shown in Figure 5.

2006). The purpose of the SINUS[9] project is to implement better learning environments at more than 1000 schools. The project defined a strategy that was tested and later disseminated to schools (Lindner, 2008). The strategy is based on curriculum design, through teachers' meetings, implementation in the classroom, and sharing of the experience. The purpose is to develop and share a new teaching and learning model that breaks the German tradition, which is characterized by a strong emphasis on practicing rules and algorithms, little attention to competencies such as modeling, compartmentalization of subjects, teaching methods that induce students to be passive, and an inappropriate mixture of learning and assessment (Blum, 2004, p. 1). Teams of three to ten teachers meet six to eight times a year for 2 to 3 years. The meetings deal specifically with subject issues and teaching methods, and teachers are expected to produce curriculum designs that will be implemented in their classrooms. Once the learning environments are implemented, the team evaluates the experience and produces improvements to the original design (Blum, 2004). The program has had positive effects as shown in a large-scale comparison schools tested in PISA 2003 (Ostermeier, Prenzel, & Duit, 2010, p. 303). This process is similar to that of lesson study (described in the next section) and to some aspects of the data-driven approach.

Poland used the PISA results for validating its educational reform. Poland's reform included several policies related to teacher education, development, and promotion: encouraging the improvement of teachers' social and economic status, introducing transparent mechanisms for promotion, improving teacher knowledge and competencies for the classroom, and offering permanent opportunities for teacher development. The improvement of Poland's PISA results seem to give credence to the need and importance of maintaining these policies (Barber & Mourshed, 2007; Saracho, 2006, pp. 17–18).

In addition, we found at least three ways in which professional development and teacher education have benefited from the results of the international studies. First, many initiatives have capitalized on the availability of videos of mathematics classrooms from several countries collected through the TIMSS video study. Second, the Japanese lesson study has been used to engage teachers in improving mathematics lessons over time. And third, materials for teachers explaining the meaning of the changes in the curriculum and giving suggestions about using new textbooks have been produced.

Videos

The potential to generate change by observing practice has been the basis for developing professional development programs that make the analysis of video an important component (Kersting, Givvin, Sotelo, & Stigler, 2010; Roth & Givvin, 2008). It is undeniable that the availability of the TIMSS videos has been useful to many activities in which instruction is analyzed, both in programs of preparation of future teachers and in professional development, although the practice had been in place before the availability of these videos. However, there is no documentation about how these videos have been used in either of these settings.

[9] *Steigerung der Effizienz des mathematisch-naturwisschenschaftlichen Unterrichts* [Increasing the efficiency of science and mathematics instruction] (http://sinus. uni-bayreuth. de/2956/).

Most of the findings of the analysis of the TIMSS video study are descriptive, focusing on characterizing the nature of instruction in different countries, documenting what has been termed the *lesson signature* (Givvin, Hiebert, Jacobs, Hollingsworth, & Gallimore, 2005; Stigler & Hiebert, 1999), a particular way in which lessons are deployed in each country, which allows observers to understand other practices. Watching substantially different ways to organize instruction helps make visible features that are taken for granted in one's own culture. Only recently, a website presenting full lessons from different countries participating in TIMSS (Australia, Czech Republic, Hong Kong, Japan, the Netherlands, Switzerland, and the United States for mathematics, Australia, Czech Republic, Japan, the Netherlands, and the United States for science) has been made public.[10] In this website, viewers can watch and listen to a variety of English-subtitled eighth-grade lessons and download the lesson plans, a map of the class, and a one-page visual description of the lesson.[11] The lessons illustrate the many differences between countries in terms of instruction. A main purpose of this site is to offer video study readers a way to corroborate the main findings about instruction in these countries. The site also anticipates being generative of new ideas for teachers as they prepare lessons. Naturally, teachers, teacher educators, and professional developers will use these lessons in a variety of ways, and it may be informative to keep track of those uses and their connections to changes in classroom instruction.

Several practical suggestions have been derived from the analysis of the video component of TIMSS. For example, in a leadership journal targeting principals, Roth and Givvin (2008) summarize four main findings from the video study, two for mathematics and two for science, and they make recommendations for taking action. In the case of mathematics, Roth and Givvin highlight that in all countries except Japan, there was a strong emphasis on solving problems with the intention of learning procedures and that "teachers in higher-achieving countries implement making connections to problems differently from teachers in the United States" (p. 24). Regarding the second point, they indicate that the teachers in the United States tend to simplify the problem rather than having students struggle with it, make links across ideas and concepts, generalize, or conjecture. They offer suggestions for principals, in terms of having teachers participate in professional development opportunities that would increase their mathematical content knowledge, give them opportunities to observe how teachers challenge "students to think about mathematics," help them "break the pattern of simplifying problems," and reinforce the idea that "students should struggle with important mathematics" (p. 24).

Lesson study

The Japanese lesson study is a process by which teachers collectively plan lessons that then are implemented in the classroom; teachers observe and take notes, and the observations and notes are used for refining the design of the lesson. The lessons become "research lessons" which are refined over time. This strategy has a very long tradition in the Japanese educational system. Lesson study was brought to the attention of U.S. researchers in the late 1980s (Lesson Study

[10] http://timssvideo.com/videos/Mathematics

[11] There are four mathematics lessons per country, covering a wide range of topics from geometry, measurement, and algebra.

Research Group, n.d.), thus predating both TIMSS and PISA. The high performance of Japanese students in TIMSS prompted questions about Japanese instruction, and in particular about the possibility of using lesson study as a professional development strategy that could spur changes in instruction in the United States (Hiebert & Stigler, 2000). A website is maintained by Teachers College at Columbia University for archival purposes and contains many documents, research papers, and manuals about lesson study.[12] According to the site, by May 2004 nearly 2300 U.S. teachers in 32 states and 335 schools had been engaged in some form of lesson study. Most of the documentation on this site refers to understanding the method itself: What are the challenges for implementation? And what are the potential outcomes of using it in professional development (Lewis, Perry, & Hurd, 2004; Lewis & Tsuchida, 1998)? The papers highlight the need for collaboration and for school reorganization (Watanabe, 2003), and in particular highlight the need of making time for teachers to meet and study (Liptak, 2002).

In New South Wales (NSW), Australia, there was also an attempt to adapt lesson study as a professional development program since the Department of Education and Training initiated the lesson study project in 2001. Initially only three secondary schools took part, but by the end of the project in 2004, the number of schools in the project had grown to 200. The lesson study evaluation data in 2002 based on 117 teacher reports from 81 schools showed that the program had succeeded in changing teacher practices and beliefs. The teacher participants, when comparing the focus lessons to their normal practice, reported using more practical activities, concrete materials and technology, new teaching procedures and intellectually challenging mathematics as well as increased collaboration among colleagues. In a follow-up survey conducted 6 months later, 63% of the respondents ($N = 64$) reported continuing to use the lesson study model of planning, evaluating, and refining to develop further lessons (White & Southwell, 2003). Lesson study has continued to be promoted in NSW government schools across many subject discipline areas although mathematics remains the main subject of study. Examples of lessons are displayed on their website.[13]

Over the past decade, there has been a spread in the use of lesson study as a professional development program in many other countries. Some of this increase in dissemination was due in part to the efforts of the Japan International Cooperation Agency (JICA), which provides international aid to developing economies. Through the collaborative efforts of JICA and experts, lesson study was introduced to several countries: Cambodia, Colombia, Egypt, Ghana, Honduras, Indonesia, Kenya, Laos, the Philippines, and Thailand (Hattori, 2007; Inprasitha, 2007; Kimura, 2007; Koseki, 2007; López & Toro-Álvarez, 2008; Odani, 2007; Okubo, 2007; N. Saito, 2007; Shimizu, 2007; Yoshida, 2007). In addition, the Asia-Pacific Economic Cooperation (APEC) has since 2006 hosted a project in conjunction with the University of Tsukuba, Japan, and the University of Khon Kaen, Thailand, to popularize lesson study. The project grew out of the recognition that lesson study constitutes an important approach towards developing human

[12] http://www.teacherscollege.edu/lessonstudy/index.html

[13] http://www.curriculum support.education.nsw.gov.au/secondary/mathematics/prolearn/windows/public_lesson_study.htm

resources, especially the expertise of teachers in mathematics classroom instruction. Five cycles of conferences had been held as of 2011, and reports of the activities related to lesson study in the APEC countries can be found on the conference website.[14]

Much of the available literature documents the challenges in adapting the lesson study approach. Its use in Japan was very much intertwined with the culture. Compared with teachers in Australia, for example, Japanese teachers hold on less to the notion that the classroom is a private professional space (White & Lim, 2008). Thus, although the lesson study approach to professional development may be versatile, there are probably cultural challenges that need to be addressed in its adoption and implementation. Indeed, Hiebert and Stigler (2000) propose that in order to change instruction, teachers need to "learn in context," that is, in their actual practice, and locate "substantive decisions for improving teaching within the schools and classrooms where teaching occurs" rather than having those decisions made "up the bureaucratic ladder" (p. 9). They remind educators that systemic and cultural change is slow but that it happens, and propose using the process of lesson study as an ongoing professional development program that is carried out within schools, within the context in which teaching happens and attending to students' learning.

Lewis (2011) reported the use of toolkits and Japanese textbooks in lesson study groups in the United States. The toolkits were used to provide support for elementary teachers to teach various mathematical concepts. The initial findings of a randomized controlled trial on the topic of fractions showed that there were significant improvements in both students' and teachers' knowledge as compared with teachers in other professional development programs (Perry, Lewis, Friedkin, & Baker, 2011).

The spread, implementation, and success of lesson study as a professional development strategy can thus be traced back to the first TIMSS Video Study. The link is indirect, but it was the TIMSS Video Study that first drew worldwide attention to lesson study as a viable teacher professional development approach that had proved to be very successful in Japan. As with other cross-national adaptations, there are always limitations and challenges because of different cultural settings. Through the hard work, efforts, and creativity of teachers and researchers, lesson study has proven to be a sustainable teacher development approach, as is shown in the case of the United States (Perry et al., 2011).

Texts for teachers

One way in which countries have attempted to reach teachers is by making more information available for their use, typically translations of frameworks and results reports. We found in addition, however, two cases—Mexico and the United States—in which there were textbooks developed with the intention of training teachers in specific aspects of practice.

[14] http://www.criced.tsukuba.ac.jp/math/apec/

Mexico has participated in PISA since the first study, and its results have motivated many reactions within the country (Rizo, 2006). One of them has to do with informing and training teachers for PISA. The webpage of the INEE (National Institute for Educational Assessment) [15] contains many documents of different types: national reports, results analysis, research protocols, and materials for teachers. *PISA in the Classroom: Mathematics* (Aguilar & Loejo, 2008) explains the PISA project and the notion of mathematical competency. The core of the book contains several curriculum designs for the classroom on quantity, change and relations, and probability. These designs contain tasks that follow the PISA framework.

The sensational and consistently high performance of Singapore in the international tests has led many to inquire about their curricular organization. Their mathematics textbooks, being in English, have certainly facilitated their incorporation into classroom practice. In the United States, there is a professional development package that includes booklets, videos, and lessons that teachers can use in their classrooms and that capitalize on the "Singapore method" for solving problems. With titles such as: *Place Value, Computation & Number Sense* (Chen, 2010a, 2010b); *8-Step Model Drawing: Singapore's Best Problem-Solving MATH Strategies* (Hogan & Forsten, 2007); *Problem-Solving Secrets from the World's Math Leader* (Hogan, 2005); and *Step-by-Step Model Drawing: Solving Word Problems the Singapore Way* (Forsten, 2010a, 2010b; Walker, n.d.), the books seek to illustrate how the Singapore textbooks organize the content across strands, how models are used, and in what ways can such a presentation and way of thinking reach all students and increase students' understanding of mathematical notions. There are also publications, for teachers, to include parents in the process (Chen, 2008). All these efforts have been made at the elementary level, and there are no reports about spread of use or impact in the classroom (except for testimonials on the back covers of the books).

Textbooks (Influence 7). There are at least two ways in which we were able to track down influences of the studies on textbooks: Some countries are using the frameworks to design textbooks, whereas others are directly importing them into the classroom. In Spain, for example, textbooks are now based on the PISA conceptual framework (Lupiáñez, 2009, pp. 231–235), and as in the Mexican case, a number of books for teachers have been published that explain the PISA framework and relate the idea of competency to curriculum (see, e.g., Rico & Lupiáñez, 2008). Recent studies have shown that the majority of tasks proposed by the textbooks are at a low level of competency, the reproduction cluster, with very few tasks from the connection cluster, and almost none from the reflection cluster (González, Monterrubio, Delgado, & Codes, 2011).

The consistently high performance of Singapore and Hong Kong students has led to the development of programs that encourage U. S. districts and teachers to adopt and use their textbooks. Singapore textbooks have received more attention. A search using terms such as "Singapore," "Hong Kong," "adoption," "mathematics," "textbooks" produced about 20 results, with only three referring to Hong Kong. One study documents results of a pilot study

[15] http://www.inee.edu.mx/

investigating the impact of using Singapore's textbooks on student achievement. The study was conducted in different sites in the United States (Ginsburg, Leinwand, Anstrom, & Pollock, 2005). The main conclusion is that "under favorable conditions, Singapore mathematics textbooks can produce significant boosts in achievement, but introducing textbooks alone is insufficient to achieve improvement" (p. 127). The results were conditioned by two factors: the mobility of the student population and the amount of professional development received. Schools with relatively small and stable populations showed large improvements over the 2-year period; similarly, there was a correlation between "improvements in the Singapore pilot schools and the intensity of the schools' participation in professional training, suggesting that teacher acceptance and commitment to the new Singapore mathematics program may be key to its success" (p. 127). A second important finding related to how confident the teachers using the Singapore textbooks were, which was corroborated with classroom observation. The researchers found that elementary teachers felt underprepared to teach with the materials, admitted that they had not understood many of the concepts they were supposed to teach, and nearly 40% struggled with implementing the curriculum. About one fifth of the teachers who received intense training successfully implemented the curriculum, whereas only 7% of teachers who received less training were successful.

In general, the teachers noticed that the Singapore textbooks offer a deeper treatment of mathematical topics, returning to a topic only to teach it with more depth. They liked the books' visual explanations that explained abstract concepts concretely and the numerous multistep problems. The teachers also identified challenges in bringing these new textbooks and methods into U.S. classrooms. For example, it would be necessary to have tailored professional development to understand the Singapore method and to learn ideas for dealing with students who have weak backgrounds or who have not been exposed to the Singapore curriculum before. Other problems included the lack of real-world examples, the use of unfamiliar terms, and the unclear alignment between the Singapore content coverage and the state-mandated standards, all of which might require using supplemental material. In addition, because the Singapore curriculum assumes a spiral progression, a successful implementation would require an adoption as early as kindergarten. Finally, the Singapore textbooks do not revisit topics later on, something that is very typical in U.S. textbooks, and thus there are no provisions for teaching or reteaching notions that have not been mastered. The study concludes with suggestions for further studies on five areas: (a) comparison of content coverage and sequencing at the state level with those of Singapore; (b) extensive textbook analysis to identify features that could be used in U.S. textbooks (the emphasis on pictorial representations appears to be a feature that would benefit special education or limited English students); (c) comparison of performance on assessment by Singaporean students with national tests, such as NAEP, and by U.S. students with Singaporean examinations; (d) changes in how teachers are tested with the PRAXIS test, suggested by an examination of the alignment of this test with the entrance tests used in Singapore to select students for Singapore's National Institute of Education; and (e) extending the piloting of the Singapore textbooks in other schools, using the versions for the American market.

Summary. It appears to us that the story of how the studies have influenced teaching and learning within the participating countries is varied. Clearly there has been more reaction in countries that did not show outstanding student performance in either the TIMSS or PISA,

although there were countries (e.g., Denmark) for which an average performance was also a trigger for reflection. The effects of the international studies seem to be exemplified in countries such as the United States, Spain, and Germany, but certainly effects have been felt in other countries as well. Moreover, it is noted that these three are developed economies with the availability of more organized and advanced facilities that have perhaps made it possible for these initiatives to be planned and implemented. In the United States, it is a story of many efforts, borne out of specific interests, rather than from a common concerted collaboration to address any given particular issue. There are many initiatives, but they seem to be carried out without any agreed-upon plan or goal. They appear to be left to individual states, researchers, schools, and practitioners. In countries with a more centralized organization, the efforts appear more coherent, with possibly Germany being the country in which efforts were most focused.

There are two salient themes at the policy level that have important, and perhaps immediate, implications for the classroom: the increased interest in centrally defined standardized testing and the increased autonomy given to schools for designing curriculum that fits their local conditions. Central agencies are moving away from prescribing what should be learned, when, and how; limiting their prescriptions to a few core content topics and competencies; and giving schools and teachers the responsibility to complete the design of their curriculum. The appearance of standardized tests is the mechanism by which bureaus can control and verify that core content and competencies are indeed being implemented and achieved. One peril of this approach might be the convergence towards a narrower list of content and competencies, which might over time determine what schools and teachers will "design" locally for their classrooms. These initiatives seem to flourish in countries where there are no centralized curricula, such as the United Sates. These initiatives have not been documented in countries that fully prescribe a national curriculum, although Singapore is starting to experiment with the idea of giving more autonomy to schools while maintaining the central control of examinations.

It is unclear what the impact of the studies in the classrooms can be for countries in which student performance is high. In Singapore, for example, the Centre for International Comparative Studies (CICS) was set up in 2009 within the National Institute of Education. The aim of the CICS was to encourage further comparative analyses, using the results from international studies to provide stakeholders with findings that could predict the factors that affect student achievement. Being situated within the National Institute of Education, which conducts teacher development, the CICS could see that the findings were more immediately used and thus influence the content of teacher education courses. But although the TIMSS may play a significant role in mathematics education in Singapore, local educators have also been quick to point out that "local stakeholders are more likely to use as main indicators the performance of students in the public examinations" (Wong, Lee, Kaur, Foong, & Ng, 2009, p. 7), noting that local public examinations are high stakes, especially in the Asian region. Student performance is used as a criterion for university entrance and for awarding scholarships, which is very highly valued in Asian societies. It would therefore be safe to say that teachers in these countries are more likely to refer to items from the public examinations than from TIMSS in planning their lessons.

Proposing a Research Agenda to Investigate these Effects

The reader will have realized by now that we were unable to document the ways in which the international studies of student performance have exerted an influence on learning and attitudes toward mathematics of students in the participating countries and elsewhere, and that we were only tangentially and indirectly able to find some traces of their influence on the teaching of school mathematics in the participating countries and elsewhere; we did not find evidence of influence at the student level. So we have only partially fulfilled our task for this chapter.

Ferrini-Mundy and Schmidt (2005), referring to TIMSS, invited the mathematics education community to capitalize on the many elements of the studies to further our understanding by using their results, research methods, and theoretical constructs that were generated. As has been the case with other endeavors, the research community has been the main direct beneficiary of these studies. Because of their complexity, studies of this magnitude require substantial know-how at all levels: from design to application, and from data management to analysis and reporting. The collaboration between participating countries becomes the vehicle by which new researchers can learn new techniques and generate research agendas that use the data obtained. It looks to us that such immense benefit for the research community has yet to reach the classroom level. For example, researchers have not explored sufficiently how mathematics education stakeholders interpret the visions behind the international studies. The lack the coherence among those visions could pose constraints in the manner they get expressed in individual teachers' practice in the classroom (McNab, 2000).

We, as a research community, have not looked at the effects of international studies from a global perspective. Yore, Anderson, and Chiu (2010) claim, in relation to the PISA study, that

> many in the mathematics and science education research communities lament the lack of influence that research results have on the education profession, schools, and teaching. Academic research done in isolation of end-users—with the faint hope that teachers, politicians, and bureaucrats will access and utilise these results to inform curriculum, assessment, and instruction and to influence public policy—has not worked. (p. 593)

We believe that as researchers of student achievement in international contexts, we need to take responsibility for improving the knowledge transfer of the results from these studies into the places in which it matters: the schools, the classrooms, and the students, and to do so in a concerted and planned way. Isolated work runs the risk, as we see here, of not having an impact on the ground level, where all the policies, mandates, and guidelines are enacted. In what follows, we propose several areas that merit attention from the research community and that can have direct impact on classroom practice, with the intention of providing some coherence to the work that needs to be done to understand the impact of the international studies at the classroom level.

Some Research Questions

In the following, we propose work that is geared towards understanding how main curricular ideas (e.g., competencies) and processes (e.g., localization of curriculum design) get interpreted

and used differently at each of the intended levels, and how those interpretations mediate what teachers do for planning and enacting lessons. We, thus, assume that the influences we have described here are real (represented with gray arrows in Figure 6), and propose to study how they might influence teachers' work, and hence, students' learning and attitudes (represented by a black continuing arrow in Figure 6). The arrows make a connection to the classroom, both in the intended and the implemented levels of curriculum, because it is at these levels where teaching is planned and implemented and learning takes place. We first propose questions related to the idea of competencies and about the use of frameworks, that is, we consider how the influences represented by Arrows 1 and 2 affect the level of intention for the classroom and the implementation processes. The next set of questions refer to influences represented by Arrows 3, 4, and 5, regarding three related aspects of curriculum control, design, and management; namely, the localization of curriculum design, the emergence of standardized testing, and the use of data-driven approaches in local reform. Finally, we consider influences from Arrows 6 and 7, regarding teacher preparation and development and textbook use. In Figure 6, we highlight that we are attending to two levels of our curriculum framework: the intended classroom level, and the implemented. We mark this attention by the dark line surrounding these two levels. This is why all the arrows, except 4, point to this area in the drawing. In what follows, we identify these questions by the corresponding numbers in the figure.

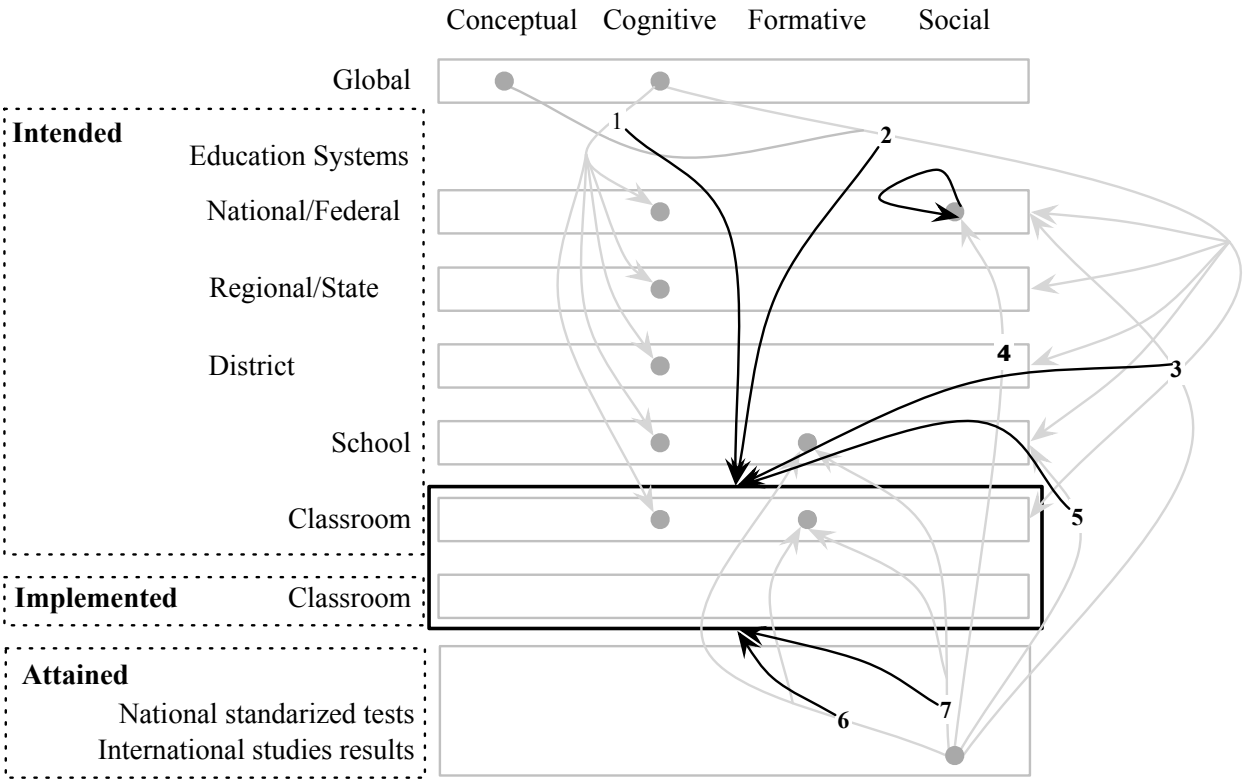

Figure 6. Proposed areas for further research on the impact of international studies in classrooms.

Competencies and frameworks. Our first set of questions refers to the impact of the notion of competencies in the curriculum and the use of the studies' frameworks. Concretely, we ask:

1. How is the idea of competency that guides the PISA framework being interpreted and transformed at the national, regional and district levels of curriculum? And how do schools and teachers interpret and use this idea when they plan and develop the curriculum in the classroom?
2. How are the frameworks of the international studies interpreted and implemented at national, regional and district levels of curriculum? And how are these interpretations and implementations used by schools and teachers at the classroom levels?

There is a growing interest in the notion of competencies—particularly in Europe—as a way of establishing longterm learning expectations by students that empower them for participating in society. It will be important to see how this notion plays out in countries in which content has been traditionally more valued than competencies and, in particular, to understand how competencies are playing a role in the push towards localization of curriculum design. Because some initial information suggests that teachers' interpretations might be at odds with original intent of competencies, understanding the way in which teachers interpret and use the notion of competencies in their daily work (for planning and for delivering instruction) is very important. We envision analyses of the transformation of the idea of competencies at different levels of the curriculum; how it is interpreted at the national, regional, or local levels; and how these interpretations influence teachers' work. This knowledge would help researchers in organizing professional development strategies that can be more effective in making the idea of competencies more transparent, and thus generating potential benefits for planning and implementing instruction in the classroom.

The different nature of the TIMSS and PISA frameworks also calls for contrasts of their influence on classroom processes. The frameworks are fundamentally the studies' backbones; yet, although we know that they have had some impact, there is little documentation on the nature of such influence. We do not know how these frameworks are interpreted at the different levels of the curriculum, and in particular, whether and how teachers use them for planning and teaching, nor whether the different nature of these frameworks is enacted in practice. If such differences were documented, then education systems could decide whether an approach that is more content oriented such as TIMSS would be better suited for its needs than an approach like PISA's that values the development of life-long skills.

Curriculum control, design, and management. Our second set of questions refers to studying the process of localization of curriculum design, the role of standardized tests, and the ways in which data-driven approaches support change. We ask:

3. What is the influence of the localization of curriculum design on the work of teachers in schools and what is its impact on the teaching and learning of mathematics? Why does it seem that this policy works in some countries and not in others?
4. How are the national standardized tests being aligned with the international studies' tests? And how is this influence reflected on what is taught and learned in classrooms?
5. In which ways do data-driven approaches influence the teaching and learning of mathematics?

We start by proposing the study of the impact of the localization of curriculum design, by which we mean the movement towards having schools and teachers produce the curriculum they will teach. Because greater responsibility is placed on schools and teachers for generating curriculum and because of concerns in some countries that teachers are not ready to assume this work, it is very important to document how such process happens. Therefore, the studies that we propose are descriptive: What does it mean for school and teachers to produce the curriculum? What resources do they use? What are the characteristics of such curriculum? Are there variations across communities that take into account the local contexts in which the curriculum is developed? What is the impact on students' learning and attitudes? Such studies would provide important information about how teachers are interpreting the task of curriculum design, the role that textbook companies play in the process, and the quality of the curriculum proposed. As mentioned before, there is some indication that in some countries teacher education programs are not preparing teachers to do this work (P. Gómez & Restrepo, 2011) and that schools have little resources to engage in this process. A second key aspect is to better understand why the localization of curriculum design is so successful in some countries and why it is harder to carry out in others. The current status of this process presents itself as ripe for such studies, as some countries have a long tradition of localized curriculum design (e.g., Finland) and some are entering this process (e.g., Colombia, Germany, and Spain).

Next, we ask about the alignment of national standardized tests and the tests used in the international studies. Because of the increased push to localize the process of curriculum design in some counties, and because of growing concerns that educators are moving slowly into a "global curriculum" defined by the content and competencies assessed in these tests, it would be important to determine the rationale behind and the extent to which newly created (and already existing) standardized tests align with the international studies tests. For some education systems, this alignment is not an issue, because the international studies tests do not have the same prominence as their national tests. In countries in which new tests are being created, and in countries in which the tests are being revised, the issue is fundamental. A close alignment would suggest a strong influence from the international studies on the education system. These studies require the collaboration between researchers in charge of the international studies in each country and their counterparts in the bureaus of education or assessment agencies. In the cases where this alignment is strong, it would be important to study the extent to which these tests are becoming an important influence on what teachers do in planning and implementing lessons: Are teachers concerned with preparing their students for the tests? In which ways? What do teachers do to prepare their students for these tests?

Finally, we ask about the extent to which data-driven approaches that use the results from the TIMSS and PISA studies can indeed influence classroom practice. We described the PROM/SE project, an effort to bring the strategy to a large number of schools to generate processes of administrative, curricular, and pedagogical change. This project is not directly tied to an international study, but it is tied to results of local standardized tests (which might become more aligned to international studies tests!). It would be important to see the extent to which such high investment of resources in collecting information about student learning, studying ways to modify practice, implement and study the changes, and revise plans, is feasible for organizations like schools, whose primary mission is to teach children.

These three areas are closely connected: The localization of curriculum requires accountability through national or common standardized tests, and the tailoring of the curriculum requires a sustained study approach in order to obtain high results on the standardized tests. A peril of this close relationship is the convergence towards an impoverished view of content and competencies both in the definition of the curriculum and in the definition of the tests (Moss, 2004). In other words, the quality of the results will depend substantially on the quality of the international studies tests and the national standardized tests because in many countries that have national standardized tests what is taught and partially learned is essentially what the agents (schools, teachers, and students) know will be assessed. National and international standardized tests more and more are not just assessing the students; the tests also assess teachers, schools, districts, and regions, and thus, students act as informants for assessing educational systems at all levels.

Teacher preparation and development and textbook use. Perhaps the most important area for research relates to the preparation of the teaching force, its continued development, and the resources that are being made available to teachers. In particular, we ask:

6. In which ways have international studies influenced the design and implementation of teacher education and professional development programs? And how do these new programs get reflected in what is taught and learned in the classroom?
7. How have textbooks from other countries been adapted and used? And what impact have these textbooks had on classroom processes?

A first step to answer questions about teacher preparation and development is documenting how widespread the use of international studies frameworks and research findings is in this area; this documentation can provide information about the ways in which the design and development of programs for teacher preparation and development have been influenced by the studies, and in which places each study has been more influential.

Beyond that use, it would be important to find out how teacher educators and teachers use resources such as lesson study, videos, or texts geared to teachers in planning and delivering their programs. It is unclear from our conversations with various teacher educators whether these resources are being used exclusively; rather, it appears that these resources are one among many that are used. This area is perhaps the one that is most important and that requires the most resources and time. It is the most important because it is the closest to studying the teaching and learning processes as they are envisioned and as teachers enact them. Such studies, however, require continuous involvement with the teachers; from the moment they participate in a teacher education program or a professional development course, to the process of planning a lesson, enacting, and assessing it. Collecting information about all the aspects of this learning process, including measures of student learning, would be necessary to understand the impact of these resources in the classroom.

Finally, there are the textbooks that have been "imported" into many classrooms and used in many different ways. If the Ginsburg et al. (2005) study is illustrative, we can anticipate many barriers to seamless adoption. Yet more interesting is to understand how teachers consider these

textbooks as additional resources that they can draw from as they plan their lessons. Because the artifacts that teachers use get transformed over several iterations (Gueudet & Trouche, 2009), longitudinal study is fundamental in order to understand such transformations and how they may affect the way lessons are designed and enacted.

Concluding Remark

The intricate connection that exists between these three areas—competencies and frameworks; curriculum control, design, and management; and teacher preparation and development and use of textbooks—is evident in our proposed research agenda. The international studies are playing an important, although subtle, role in shaping the landscape of curriculum design, implementation, and attainment. The lack of concerted efforts to trace their influence at the classroom level should be a cause for concern. To paraphrase an old saying, "All roads lead to the classroom." In the end, therefore, the substantial investments of money and effort that the international studies of student achievement require ought to have one single purpose: to improve teaching in order to improve learning. Collecting evidence in a systematic way may prove beneficial to achieving this ultimate goal.

Acknowledgements

The authors thank Ms. Heejoo Suh for her invaluable bibliographic research skills and the three reviewers who raised important points that helped us clarify our argument.

References

Aguilar, M., & Loejo, A. (2008). *PISA en el aula: matemáticas* [PISA in the classroom: Mathematics]. México, DF, México: Instituto Nacional para la Evaluación de la Educación (INEE).

Alcaide, S., & Álvarez, P. (2011, March 4). Madrid ofrece a los colegios decidir un tercio del programa escolar [Madrid offers to let schools decide a third of the school program]. *El País,* p. 38.

Arias, R. M. (2006). La metodología de los estudios PISA [Methodology in the PISA studies]. *Revista de Educación*(Extraordinario), 111–129.

Barber, M., & Mourshed, M. (2007). *How the world's best-performing school systems come out on top*. New York. NY: McKinsey.

Bassells, F. (2011, March 4). Cataluña se prepara para que cada centro fije el contenido y el orden de las materiasl [Catalonia gets ready for each center to set subject matter contents and sequencing]. *El País,* p. 38.

Beyer, L., & Liston, D. (1996). *Curriculum in conflict: Social visions, educational agendas, and progressive school reform*. New York, NY: Teachers College Press.

Blanco, L., & Rico, L. (2011, January 8). ¿Qué hacer con los datos de PISA? [What to do with PISA data?]. *El País*. Retrieved from http://www.elpais.com/articulo/sociedad/hacer/datos/PISA/elpepusoc/20110108elpepusoc_9/Tes

Blum, W. (2004). *Opportunities and problems for "quality mathematics teaching": The SINUS and DISUM projects*. Paper presented as a regular lecture at the International Congress on Mathematical Education (ICME) 10. Retrieved from http://www.icme10.dk/proceedings/pages/regular_pdf/RL_Werner_Blum.pdf

Bonnet, G. (2008). Do teachers' knowledge and behaviour reflect their qualifications and training? Evidence from PASEC and SACMEQ country studies. *Prospects, 38*(3), 325–244.

Bottani, N. (2006). La más bella del reino: El mundo de la educación en alerta con la llegada de un príncipe encantador [The fairest of them all: The world of education on alert with the arrival of a prince charming]. *Revista de Educación*(Extraordinario), 75–90.

Bracey, G. W. (1998). TIMSS, rhymes with "dims," as in "witted." *Phi Delta Kappan, 79*(9), 686–687.

Cairn, W. D. (1935). Advanced preparatory mathematics in England, France and Italy. *American Mathematical Monthly, 42*(1), 17–34.

Chen, S. (2008). *The parent connection for Singapore math: Tools to help them "get it" & get behind it*. Peterborough, NH: Crystal Springs Books.

Chen, S. (2010a). *Singapore math: Place value, computation, & number sense*. Peterborough, NH: Crystal Springs Books.

Chen, S. (2010b). *Singapore math: Place value, computation, & number sense [CD-ROM]*. Peterborough, NH: Crystal Springs Books.

Clarke, D. (2003). International comparative studies in mathematics education. In A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Second international handbook of mathematics education* (pp. 143–184). Dordrecht, the Netherlands: Kluwer.

Curriculum Development Council. (2000). *Mathematics curriculum guide (P1–P6)*. Hong Kong, China: Author. Retrieved from http://www.edb.gov.hk/index.aspx?nodeID=4907&langno=1.

Departemen Pendidikan Nasional. (2003). *Standar kompetensi: Mata pelajaran matematika* [Competency standards: Subject of mathematics]. Jakarta, Indonesia: Author.

Dunne, D. W. (2000, July 21). Why are Chicago-area students tops in the world in math and science? *Education World*. Retrieved from http://www.educationworld.com/a_curr/curr251.shtml

Education Council. (2006). Recommendation of the European Parliament and the Council of 18 December 2006 on key competencies for lifelong learning. *Official Journal of the European Union, L*(394), 10–18.

Ertl, H. (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.

Ferreras, V. (2006). Participación de Navarra en la evaluación PISA [Participation of Navarra in the PISA assessment]. *Revista de Educación*(Extraordinario), 531–542.

Ferrini-Mundy, J., & Schmidt, W. (2005). International comparative studies in mathematics education: Opportunities for collaboration and challenges for researchers. *Journal for Research in Mathematics Education, 36*(3), 164–175.

Figueroa, M. (2008). *Examination of a new method to teach elementary students mathematics* (Unpublished doctoral dissertation). Flagstaff: Northern Arizona University.

Forsten, C. (2010a). *Solving word problems the Singapore math model-drawing strategy [CD-ROM]*. Peterborough, NH: Crystal Springs Books.

Forsten, C. (2010b). *Step-by-step model drawing: Solving word problems the Singapore way*. Peterborough, NH: Crystal Springs Books.

Freudenthal, H. (1975). Pupils' achievement internationally compared. *Educational Studies in Mathematics, 6*, 127–186.

Ginsburg, A., Leinwand, S., Anstrom, T., & Pollock, E. (2005). *What the United States can learn from Singapore's world-class mathematics system and what Singapore can learn from the United States: An exploratory study*. Washington, DC: American Institutes for Research.

Givvin, K. B., Hiebert, J., Jacobs, J. K., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review, 43*(3), 311–343.

Gómez, J. A. (2006). La Rioja hacia la evaluación censal de PISA [Rioja towards a census assessment with PISA]. *Revista de Educación*(Extraordinario), 521–529.

Gómez, P., & Restrepo, A. (2011). *Procesos de planificación en matemáticas y autonomía escolar* [Planning processes in mathematics and school autonomy]. Paper presented at the XV Simposio de la Sociedad Española de Investigación en Educación Matemática, Ciudad Real, Spain.

González, M. T., Monterrubio, M. C., Delgado, M. L., & Codes, M. (2011). *Tipos de situaciones de las actividades de análisis matemático en los libros de texto de Educación Secundaria en España* [Types of situations in mathematical analysis activities in secondary education textbooks in Spain]. Paper presented at the XIIIth Inter American Conference on Mathematics Education.

Graña, J. (2006). El estudio PISA en Galicia: Pasado, presente y futuro [The PISA study in Galicia: Past, present and future]. *Revista de Educación*(Extraordinario), 515–520.

Gruber, K. H. (2006). The German "PISA-shock": Some aspects of the extraordinary impact of the OECD's PISA study on the German education system. In H. Ertl (Ed.), *Cross-national attraction in education. Accounts from Germany* (pp. 195–208). Didcot, UK: Symposium Books.

Gueudet, G., & Trouche, L. (2009). Towards new documentation systems for mathematics teachers? *Educational Studies in Mathematics, 71*(3), 199–218.

Hattori, K. (2007). Lesson study in South Africa. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 226–229). Singapore: World Scientific.

Hiebert, J., & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS video study. *Elementary School Journal, 101*(1), 3–20.

Hogan, B. (2005). *Singapore math: Problem-solving secrets from the world's math leader*. Peterborough, PA: Crystal Springs Books.

Hogan, B., & Forsten, C. (2007). *8-step model drawing: Singapore's best problem-solving MATH strategies*. Peterborough, PA: Crystal Springs Books.

Husén, T. (1983). Are standards in U.S. schools really lagging behind those in other countries? *Phi Delta Kappan, 64*(7), 455–461.

Husén, T. (Ed.). (1967). *International study of achievement in mathematics* (Vol. 1&2). New York, NY: Wiley.

Inprasitha, M. (2007). Lesson study in Thailand. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 188–193). Singapore: World Scientific.

Junta de Andalucía. (2007). Resolución de 10 de diciembre de 2010, de la Dirección General de Ordenación y Evaluación Educativa, por la que se establecen determinados aspectos de la Orden de 27 de octubre de 2009, por la que se regulan las pruebas de la evaluación de diagnóstico, para su aplicación en el curso 2010-2011 [Resolution of December 10, 2010, of the Directorate General of Educational Planning and Evaluation on establishing certain aspects of the Order of October 27, 2009, which regulates the testing for diagnostic evaluation applied during 2010–2011]. *BOJA, 12*, 22-23.

Kaur, B. (2002). TIMSS-R: Mathematics achievement of eighth graders from Southeast Asian countries. *Journal of Science and Mathematics Education in Southeast Asia, 25*(2), 66–92.

Kaur, B. (2003, 31 March - 1st April). *Evolution of Singapore's secondary school mathematics curricula.* Paper presented at the Talking it through: A cross-national conversation about secondary mathematics curricula, at the National Academy of Sciences. Retrieved from http://www7.nationalacademies.org/mseb/workshop_background_materials_kaur_spores.pdf.

Kaur, B. (2005). Schools in Singapore with high performance in mathematics at the eighth grade level. *Mathematics Educator, 9*(1), 29–38.

Kaur, B. (2009). Performance of Singapore students in Trends in International Mathematics and Science Studies (TIMSS). In K. Y. Wong, P. Y. Lee, B. Kaur, P. Y. Foong & S. F. Ng (Eds.), *Mathematics education: The Singapore journey* (pp. 439–463). Singapore: World Scientific.

Kaur, B., Koay, P. L., & Yap, S. F. (2009). International project on mathematical attainment. In K. Y. Wong, P. Y. Lee, B. Kaur, P. Y. Foong & S. F. Ng (Eds.), *Mathematics education: The Singapore journey* (pp. 494–511). Singapore: World Scientific.

Kaur, B., & Pereira-Mendoza, L. (2000a). Singapore primary school TIMSS data: Data representation, analysis and probability and patterns, relations and functions. *Mathematics Educator, 5*(1/2), 180–193.

Kaur, B., & Pereira-Mendoza, L. (2000b). TIMSS: Performance of Singapore secondary students (Part B): Proportionality, measurement, fractions and number sense. *Journal of Science and Mathematics Education in Southeast Asia, 23*(1), 54–70.

Kaur, B., & Yap, S. F. (2009). Kassel project on the teaching and learning of mathematics: Singapore participation. In K. Y. Wong, P. Y. Lee, B. Kaur, P. Y. Foong & S. F. Ng (Eds.), *Mathematics education: The Singapore journey* (pp. 479–480). Singapore: World Scientific.

Keitel, C., & Kilpatrick, J. (1998). The rationality and irrationality of international comparative studies. In G. Kaiser, E. Luna & I. Huntley (Eds.), *International comparisons in mathematics education* (pp. 242–257). London, UK: Falmer.

Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education, 61*(1–2), 172–181.

Kilpatrick, J. (1971). Some implications of the International Study of Achievement in Mathematics for mathematics educators. *Journal for Research in Mathematics Education, 2*(2), 164–171.

Kimmelman, P., Kroeze, D., Schmidt, W., van der Ploeg, A., McNeely, M., & Tan, A. (1999). *A first look at what we can learn from high performing school districts: An analysis of TIMSS data from the First in the World Consortium*. Jessup, MD: U.S. Department of Education (ERIC Document Reproduction Service No. ED 433 243).

Kimura, E. (2007). Lesson study in Honduras. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 230–235). Singapore: World Scientific.

Koseki, K. (2007). Lesson study in Indonesia. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 214–215). Singapore: World Scientific.

Kotthoff, H.-G., & Pereyra, M. A. (2009). La experiencia del PISA en Alemania: recepción, reformas recientes y reflexiones sobre un sistema educativo en cambio [The experience of PISA in Germany: Response, recent reforms and reflections on a changing educational system]. *Profesorado. Revista de Currículum y Formación de Profesorado, 13*(2), 1–24.

Landry, S. D. (2010). *Degrees of alignment among K-12 mathematics content standards of instruction: An analysis of high-performing and low-performing data sets* (Unpublished doctoral dissertation). Fort Worth, TX: Texas Christian University.

Lee, O., Deaktor, R., Enders, C., & Lambert, J. (2008). Impact of a multiyear professional development intervention on science achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 45*(6), 726–744.

Lee, O., Deaktor, R., Hart, J. E., Cuevas, P., & Enders, C. (2005). An instructional intervention's impact on the science and literacy achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 42*(8), 857–887.

Lee, V., & Zuze, T. L. (2010). *School resources and performance in Sub-Saharan Africa: Possible implications for "education for all"* (manuscript under review). Ann Arbor, MI.

Lesson Study Research Group. (n.d.). Timeline of U.S. lesson study. Retrieved from Teachers College, Columbia University: http://www.teacherscollege.edu/lessonstudy/timeline.html

Lewis, C. (2011, February 19–20). *Lesson study with Japanese curriculum materials: A randomized controlled trial.* Paper presented at the APEC Tsukuba International Conference on Lesson Study, Tsukuba, Japan.

Lewis, C., Perry, R., & Hurd, J. (2004). A deeper look at lesson study. *Educational Leadership, 61*(5), 6–11.

Lewis, C., & Tsuchida, I. (1998). A lesson is like a swiftly flowing river: How research lessons improve Japanese education. *American Educator, 22*(4), 12–17, 50–52.

Liberal Party of Australia. (2010, December 7). PISA findings back school autonomy. *Liberal*. Retrieved from http://www.liberal.org.au/Latest-News/2010/12/07/PISA-findings-back-school-autonomy.aspx

Lindner, M. (2008). New programmes for teachers' professional development in Germany: The programme SINUS as a model for teachers' professional development. *Interacções, 4*(9), 149–155.

Liptak, L. (2002). It's a matter of time. *Research for Better School Currents, 5*(2), 6–7.

López, L. S., & Toro-Álvarez, C. (2008). Formación de docentes en la enseñanza de las matemáticas a través de la resolución de problemas en la Red de Comprensión Lectora y Matemáticas–CCyM, segunda etapa [Teacher preparation in mathematical thinking through problem solving in the Reading and Mathematics Comprehension, CCyM, network, second stage]. *Universitas Psychologica, 7*(3), 753–765.

Lupiáñez, J. L. (2009). *Expectativas de aprendizaje y planificación curricular en un programa de formación inicial de profesores de matemáticas de secundaria* [Learning expectations and curricular planning in a secondary mathematics teacher preparation program] (Unpublished doctoral dissertation). Granada, Spain: Universidad de Granada.

Maestro, C. (2006). La evaluación del sistema educativo [The evaluation of the education system]. *Revista de Educación*(Extraordinario), 315–336.

McNab, D. (2000). Raising standards in mathematics education: Values, vision, and TIMSS. *Educational Studies in Mathematics, 42*(1), 61–80.

Ministerio de Educación Nacional [MEN]. (1994). *Ley 115 de Febrero 8 de 1994. Por la cual se expide la ley general de educación* [Law 115 of February 8, 1994. In virtue of which the general law of education is issued]. Bogotá, Colombia: Author.

Ministerio de Educación Nacional [MEN]. (2006). *Estándares básicos de competencias en lenguaje, matemáticas, ciencias y ciudadanas* [Basic standards of competencies in language, mathematics, science and citizenship]. Bogotá: Author.

Ministry of Education Culture Sports and Science. (2010). *Elementary school teaching guide for the Japanese course of study: Mathematics (Grade 1–6)* (M. Isoda, Trans.). Tsukuba, Japan: Center for Research on International Cooperation in Educational Development (CRICED), University of Tsukuba.

Moss, P. (2004). The risks of coherence. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability* (One Hundred Third Yearbook of the National Society for the Study of Education, Part 2, pp. 217–238).). Chicago, IL: University of Chicago Press.

Mullis, I. (1999). *Attaining excellence: TIMSS as a starting point to examine mathematics assessment. An in-depth look at geometry and algebra*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

National Council of Teachers of Mathematics [NCTM]. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics [NCTM]. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

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

National Research Council. (1999). *Global perspectives for local action: Using TIMSS to improve U.S. mathematics and science education--Professional development guide*. Washington, D.C.: National Academy Press.

Neumann, K., Fischer, H., & Kauertz, A. (2010). From PISA to educational standards: The impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education, 8*(3), 545–563.

Odani, K. (2007). Lesson study in Cambodia. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 206–209). Singapore: World Scientific.

Okubo, K. (2007). Lesson study in Egypt. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 216–217). Singapore: World Scientific.

Organisation for Economic Cooperation and Development [OECD]. (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Paris, France: Author.

Organisation for Economic Cooperation and Development [OECD]. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris, France: Author.

Organisation for Economic Cooperation and Development [OECD]. (2005). *School factors related to quality and equity: Results from PISA 2000*. Paris, Farnce: Author.

Organisation for Economic Cooperation and Development [OECD]. (2010). *Education at glance 2010*. Paris, France: Author.

Ostermeier, C., Prenzel, M., & Duit, R. (2010). Improving science and mathematics instruction: The SINUS Project as an example for reform as teacher professional development. *International Journal of Science Education, 32*(3), 303–327.

Palamidessi, M. (2006). *Desarrollos curriculares para la educación básica en el Cono Sur: prioridades de política y desafío de la práctica* [Curriculum development for basic education in the Southern Cone: Policy priorities and challenges to practice]. Geneva, Switzerland: United Nations Educational, Scientific and Cultural Organization (UNESCO).

Perry, R., Lewis, C., Friedkin, S., & Baker, E. K. (2011). *Improving the mathematical content base of lesson study: Interim summary of results.* Retrieved from http://www.lessonresearch.net/IES%20Abstract_01.03.11.pdf.

Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE]. (2006a). *Knowing mathematics: What can we learn from teachers* (Research Report, Vol. 2). East Lansing, MI: Michigan State University.

Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE]. (2006b). *Making the grade: Fractions in your schools* (Research Report, Vol. 1). East Lansing, MI: Michigan State University.

Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE]. (2008). *Dividing opportunities: Tracking in high school science* (Research Report, Vol. 4). East Lansing, MI: Michigan State University.

Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE)]. (2009a). *Content coverage and the role of instructional leadership* (Research Report, Vol. 7). East Lansing, MI: Michigan State University.

Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE]. (2009b). *Opportunities to learn in PROM/SE classrooms: Teachers' reported coverage of mathematics content* (Research Report, Vol. 6). East Lansing, MI: Michigan State University.

Promoting Rigorous Outcomes in Mathematics and Science Education [PROM/SE]. (2009c). *Variation across districts in intended topic coverage: Mathematics* (Research Report, Vol. 5). East Lansing, MI: Michigan State University.

Rico, L. (2011). El estudio PISA y la evaluación de la competencia matemática [The PISA study and the assessment of mathematical competency]. *Matematicalia, 7*(1), 1–11.

Rico, L. (Ed.). (1997). *Bases teóricas del currículo de matemáticas en educación secundaria* [Theoretical basis for mathematics curriculum in secondary education]. Madrid, Spain: Síntesis.

Rico, L., & Lupiáñez, J. L. (2008). *Competencias matemáticas desde una perspectiva curricular [Mathematical competencies from a curricular perspective]*. Madrid, Spain: Alianza Editorial.

Rizo, F. M. (2006). PISA en América Latina: lecciones a partir de la experiencia de México de 2000 a 2006 [PISA in Latin America: Lessons from the Mexican experience from 2000 to 2006]. *Revista de Educación*(Extraordinario), 153–167.

Robitaille, D. F., & Travers, K. J. (1992). International studies of achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 687–709). Reston, VA: National Council of Teachers of Mathematics.

Roth, K., & Givvin, K. B. (2008). Implications for math and science instruction from the TIMSS video study. *Principal Leadership, 8*(9), 22–27.

Saito, M. (2010). Have gender differences in reading and mathematics achievement improved? *Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) Policy Issues Series, 4*, 1–3.

Saito, N. (2007). Lesson study in Laos. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 210-213). Singapore: World Scientific.

Saracho, A. (2006). *Casos de estudio de reacciones a PISA 2000 y 2003: Alemania, Polonia, Brasil y otros* [Case studies of reactions to PISA 2000 and 2003: Germany, Poland, Brazil and others]. México, DF: Fundación IDEA.

Schleicher, A., & Shewbridge, C. (2008). *What makes school systems perform? Seeing school systems through the prism of PISA*. Paris, France: Organisation for Economic Cooperation and Development [OECD].

Schmidt, W. H., & McKnight, C. C. (1995). Surveying educational opportunity in mathematics and science: An international perspective. *Educational Evaluation and Policy Analysis, 17*(3), 337–353.

Schmidt, W. H., McKnight, C. C., Valverde, G., Houang, R. T., & Wiley, D. E. (1996). *Many visions, many aims, Vol. 1: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht, the Netherlands: Kluwer.

Shimizu, S. (2007). Lesson study in the Philippines. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 202–205). Singapore: World Scientific.

Shymansky, J. A., Yore, L. D., & Anderson, J. O. (2004). Impact of a school district's science reform effort on the achievement and attitudes of third- and fourth-grade students. *Journal of Research in Science Teaching, 41*(8), 771–790.

Sloane, P. F. E., & Dilger, B. (2005). The competence clash—Dilemmata bei der Übertragung des 'Konzepts der nationalen Bildungsstandards' auf die berufliche Bildung, Berufs- und Wirtschaftspädagogik [The competence clash—Dilemmas in conveying the "concept of national education standards" to vocational, professional, and business education]. *Berufs- und Wirtschaftspädagogik, 8*, 1–32. Retrieved from http://www.bwpat.de/ausgabe8/sloane_dilger_bwpat8.pdf

Stark, J., & Lattuca, L. R. (1997). *Shaping the college curriculum: Academic plans in action*. Boston, MA: Allyn & Bacon.

Stigler, J. W., Gonzalez, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States (NCES 1999-074)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction: An overview of the TIMSS video study. *Phi Delta Kappan, 78*(1), 14–21.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap*. New York, NY: Free Press.

Travers, K. J., & Westbury, I. (Eds.). (1989). *The IEA Study of mathematics I: Analysis of mathematics curricula*. Oxford, UK: Pergamon.

Valdés, H., Treviño, E., G., A. C., Castro, M., Carrillo, S., Costilla, R., et al. (2008). *Student achievement in Latin America and the Caribbean: Results of the Second Regional Comparative and Explanatory Study (SERCE)*. Santiago, Chile: Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, United Nations Educational, Scientific and Cultural Organization (UNESCO).

Walker, L. (n.d.). *Model drawing for challenging word problems: Finding solutions the Singapore way*. Peterborough, NH: Crystal Springs Books.

Watanabe, T. (2003). Lesson study: A new model of collaboration. *Academic Exchange Quarterly, 7*(4), 180–184.

White, A., & Lim, C. S. (2008). Lesson study in Asia Pacific classrooms: Local responses to a global movement. *ZDM: International Journal for Mathematics Education, 40*(6), 915–939.

White, A., & Southwell, B. (2003). *Lesson study project: Evaluation report*. Ryde: Australia: Department of Education and Training, Professional Support and Curriculum Directorate.

Wong, K. Y., Lee, P. Y., Kaur, B., Foong, P. Y., & Ng, S. F. (2009). Introducing the landscape of Singapore mathematics. In K. Y. Wong, P. Y. Lee, B. Kaur, P. Y. Foong & S. F. Ng. (Eds.), *Mathematics education: The Singapore journey* (pp. 1–9). Singapore: World Scientific.

Yore, L., Anderson, J., & Chiu, M.-H. (2010). Moving PISA results into the policy arena: Perspectives on knowledge transfer for future considerations and preparations. *International Journal of Science and Mathematics Education, 8*(3), 593–609.

Yoshida, M. (2007). Lesson study in Ghana. In M. Isoda, M. Stephens, Y. Ohara & T. Miyakawa (Eds.), *Japanese lesson study in mathematics: Its impact, diversity and potential for educational improvement* (pp. 222–225). Singapore: World Scientific.

Young, J. W. A. (1900). *The teaching of mathematics in the higher schools of Prussia*. New York, NY: Longmans, Green.