Chapter 2
Review of the Literature: Factors Contributing to Achievement GAP

Abstract This chapter provides a review of literature on the topic of comparative studies of mathematics achievement and summaries the key factors suggested by other studies contributing to achievement gap between the East and the West in generally, especially between China and Australia.

Keywords Comparative education · Mathematics achievement · Achievement gap · Factors

Chapter 1 has established the research problem to be addressed in this study: to account for key factors contributing to the achievement gap in mathematics between Chinese and Australian students. The purpose of this chapter is to provide a research background for comparing mathematics curriculum and assessment practices in Chinese and Australian primary schools. A rationale underpinning the broad field of comparative education is provided initially. This is followed by an overview of comparative studies on mathematics achievement and the main findings of those studies focusing on the achievement gap between East Asian and Western countries are highlighted. In Sect. 2.3, factors contributing to the achievement gap are examined, followed by a discussion of strengths and limitations of current comparative studies. Sections 2.5 and 2.6 discuss the relationship between teaching, learning, and assessment, and present a tentative model as a theoretical framework for the study.

An overview of the research literature pertaining to comparative studies of mathematics achievement is presented in the following six sections:

2.1. A rationale for comparative education
2.2. Comparative studies on mathematics achievement
2.3. Factors accounting for the achievement gap
2.4. Strengths and limitations of current comparative studies
2.5. The relationship between teaching, learning, and assessment
2.6. A tentative model for the study

D. Zhao, *Chinese Students' Higher Achievement in Mathematics*,
Mathematics Education – An Asian Perspective,
DOI 10.1007/978-981-10-0285-4_2

2.1 A Rationale for Comparative Education

Mathematics education, as an essential component of schooling, has been an enduring concern of comparative studies internationally (Clements et al. 2013; Husen 1967a, b; Mullis et al. 2000; Organization for Economic Co-operation and Development (OECD) 2014; Robitaille and Garden 1989). In this section, a rationale underpinning comparative education in general, and comparative studies of mathematics education in particular, is discussed in terms of its nature, importance, and purpose.

2.1.1 The Nature of Comparative Education

As early as 1817, Marc-Antoine Jullien, the 'father of comparative education,' defined comparative education as the act of contrasting the features and methods of education in different countries (Gauthrin 1993; Hans 2012). Eckstein and Noah (1992) explain that the concept of comparative education, in its most general sense, means inspecting two or more educational operations, which refers to any act associated with learning and teaching, in order to discover the differences and similarities between them.

It is widely accepted (Bishop 1996; Graf and Leung 2000; Hans 2012; Paris and Wellman 1998) that comparative education connotes the analysis of educational similarities and differences in two or more national environments in terms of social, political, cultural, and other contexts. Similarly, the nature of comparative studies in mathematics education is defined as a comparison of mathematics education in different countries or contexts to identify differences and similarities, and to interpret and explain the similarities and differences identified (Clements et al. 2013; Graf and Leung 2000).

2.1.2 The Importance of Comparative Studies

Comparative studies are important because they can be used as an instrument for thinking, research, and reform in the field of education in general, and mathematics education in particular (Lokan et al. 1997; Reynolds et al. 2000; Stevenson and Stigler 1992).

According to Bishop (1996), comparison is an instrument for thinking; without it neither learning nor education can happen. In terms of mathematics education, he suggests that educators seek to find out about interesting work being done in other countries and learn from the experience of colleagues elsewhere who use different practices or follow a different philosophy. He argues that other countries' and other colleagues' experiences always offer potentially interesting contrasts to one's own, which can broaden one's knowledge base and suggest alternatives to their present practices.

Eckstein and Noah (1992) assert that a fundamental premise of comparative study is that people can truly comprehend themselves only in the context of secure knowledge of other societies. Similarly, Stevenson and Stigler (1992) argue that meaning often emerges through comparison. People do not understand what children can achieve until they have seen what other children of the same age do in other countries. They suggest that international comparisons can help people discover characteristics of ones' own culture that they fail to notice by reason of familiarity. They strongly recommend using comparison to clarify and sharpen our perceptions.

Husen (1967a), as a pioneer in the field of comparative studies on mathematics achievement, argues that the importance of comparative studies is that they make use of naturally existing differences. The study of cross-national or cross-cultural variability in the educational domain can advance hypotheses which can be tested in one way or another (p. 19). He further indicates that considerable differences exist between countries in school organization, curricula, teaching, expectations of students, and many other factors potentially related to effective teaching and learning. He advocates that it is possible to use appropriately designed comparative studies to examine the conditions for students to achieve and what conditions are most likely to facilitate students' learning.

McAdams (1993) suggests that comparative studies have been used as an instrument for educational reform. He indicates that comparison of the achievement of one's own students with their peers in other nations has received special attention in the USA. For example, the school reform movement in the USA relies heavily upon comparative data as the rationale for advocating revolutionary change in the operation of American schools. According to Robitaille and Nicol (1994), comparative studies can provide opportunities for the discussion and debate of important issues such as how to provide effective and efficient mathematics education in an international context. Researchers, educators, and policymakers can be provided with data, to reflect on their own educational practice, to gain further insight into possible directions in curricula and instruction, and to challenge previous standards concerning the expectation of students' learning and what instructional activities enhance that learning.

2.1.3 The Purpose of Comparative Education

According to Arnove (2001), historically, the field of comparative education has been shaped by three principle dimensions: theoretical, practical, and global. For each of these dimensions, the purpose of comparative education differs. First, in terms of theoretical dimension, one major goal of comparative studies has been to contribute to theory building and to the formulation of general propositions about the workings of school systems and their interactions with their contexts including economic, political, cultural, and social factors. Second, in terms of a practical dimension, the reason for studying other societies' education systems is to discover what can be learned that will contribute to improved policy and practice at home. Third, in terms of the global dimension, the purpose of comparative studies has been a concern with contributing to international understanding and peace.

Halls (1990) suggests that the ultimate aim of comparative education is to improve education, so as to stimulate a higher quality life style for the society and the individual. Lokan et al. (1997) explain that since the 1960s, there has been a growing concern around the world that investments in education need to be related to the outcomes of education, which, in turn, are seen to be making a powerful contribution to the country's economic prosperity and general well-being. Therefore, the fundamental purpose of comparative studies is to improve education.

In terms of the purpose of comparative studies in mathematics achievement, Robitaille and Garden (1996) explain that most participants of comparative studies are interested in comparing their educational programs, practices, and achievements with those of other countries, especially those that are important to them for economic, cultural, and political reasons. There is also a widely held belief (Australia Association of Mathematics Teachers (AAMT) 1997; Chinese Ministry of Education 2000; Organization for Economic Co-operation and Development (OECD) 2014) that a nation's continued economic well-being and its ability to compete in the international market is heavily linked to how well that country's elementary and secondary school students achieve, particularly in mathematics and in science.

Mathematics education has been a central focus of comparative international studies (Clements et al. 2013; Graf and Leung 2000; Husen 1967a; Lokan et al. 1997; Mullis et al. 2000; Robitaille and Nicol 1994). Mathematics has always held a key position in the school system: It is one of the few subjects that is taught in most school systems internationally. Although countries differ in social, economic, and cultural contexts, there is still a surprising degree of similarity across mathematics curricula worldwide (Howson and Wilson 1986; Lokan et al. 1997; Owens and Perry 2001).

This universal status and importance of mathematics, the similarity of mathematics curricula worldwide, and strong beliefs about the link between the study of mathematics and the development of a nation's economic strength (Graf and Leung 2000; Lokan et al. 1997) make comparative studies of mathematics education of primary significance to researchers, educators, and policymakers. According to Graf and Leung (2000), mathematics educators from all around the world are continuing to make efforts to improve the quality of mathematics education, and one way to achieve this is through cross-cultural research. One purpose of comparative studies in mathematics education, therefore, is to improve mathematics teaching and learning, so as to enhance the economic viability of the country.

2.2 Comparative Studies on Mathematics Achievement

The international comparative study of mathematics education constitutes a developing field of research (Clements et al. 2013). According to Robitaille and Nicol (1994), early studies conducted at the beginning of 1900s focused on describing the organization of mathematics instruction, while current studies are generally concerned with measuring students' mathematical achievement. In this section, a review of IEA studies of mathematical achievement is presented firstly in Sect. 2.2.1. An overview of other comparative studies of mathematical achievement is discussed in Sect. 2.2.2. The key findings of current comparative studies of mathematics achievement are highlighted in Sect. 2.2.3.

2.2.1 IEA Studies on Mathematics Achievement

The International Association for the Evaluation of Educational Achievement (IEA), established in 1959 is a cooperative network of research centers that conduct international comparative studies on schools and aid the research community in developing international tests and statistical analyses (Genova 2015; Husen 1967a). Today, membership of the IEA consists of institutions from more than fifty countries with the common goal of investigating the potential influence of alternative curricula, teaching strategies, and administration strategies on student achievement (Hayes 1991). According to Jacobsen (1996) and Genova (2015), for the last five decades, the major influence in the field of comparative studies has been the larger-scale comparative studies on mathematics achievement conducted by the IEA. Since conducting the *First International Mathematics Study* (*FIMS*) in 1964, the IEA has undertaken four large-scale studies of mathematics achievement by the end of 2000, which are reviewed in the following sections.

The First International Mathematics Study (FIMS)

According to Husen (1967a), FIMS was the first large-scale comparative study measuring achievement across various mathematics topics in twelve countries.[1] Countries that participated in FIMS included Australia, Belgium, England, the Federal Republic of Germany, France, Finland, Israel, Japan, the Netherlands, Scotland, Sweden, and the USA. The student populations sampled for this study consisted of 13-year-old students and students in their last year of secondary school.

Husen (1967a) indicated that the aim of this study was to test a number of fundamental hypotheses related to the outcomes of different patterns of educational organization set in a variety of social and cultural contexts. By using a two-dimensional item-specification grid consisting of a content-by-cognitive behavior matrix, topics involving arithmetic, algebra, geometry, and calculus

[1]China did not take part in this study.

were selected for the achievement tests of FIMS at each population level. Test items were presented in multiple-choice format. Students, teachers, school principals, and national experts were also asked to complete questionnaires dealing with such variables as type of school, size of mathematics class, number of those teachers with specialist training, and so on. Students responded to an attitude survey providing information about their personal views of mathematics, their schoolwork, and career aspirations.

In terms of test results, at the lower level (13-years-olds), where 100 % of the age group was in full-time schooling, differences between countries in average score were quite marked. Japanese students had the highest mean (31.2) and Sweden the lowest (15.7). Australia ranked fifth and had a mean just above the average.

In terms of findings, many of the hypotheses posed at the outset of the study could neither be confirmed nor be rejected. The author indicated that neither the productivity of an educational system of a country, nor the effect of the instruction given, could be assessed reliably from national mean scores. The only consistent finding was large variability in student achievement (Husen 1967b, p. 288). Husen (1967b) correlated achievement with other variables, but without more detailed explanation of these variables in the contexts of the countries involved, these correlations meant very little. For example, the findings were that per-student expenditure and teacher training displayed negative correlations with total mathematics scores at the 13-year-old level. The report attributed this finding to the fact that Sweden and the USA had the lowest score and at the same time the highest per-student expenditure, while in Japan the opposite was the case (Husen 1967b, p. 290).

In terms of limitations, according to Freudenthal (1975), FIMS was heavily criticized for its lack of attention to the significance of differing curricula in the participating countries, the poor contextualization of the outcomes, and the heavy reliance on multiple-choice question formats. Freudenthal argued that the tests used were not valid, that the meanings of many of the variables overlapped, and that some important variables such as examinations were not included. Husen (1967b, p. 309) also acknowledged the limitations of FIMS and made a number of suggestions for further IEA research. One of these suggestions was that case studies of particular countries needed to be conducted:

> In the mathematics stage of the IEA study, no detailed case studies of particular countries have been carried out. Such studies would have to imply thorough analyses of curricula, textbooks, and other learning materials. In some instances, special attention would have to be directed to incentives used in instruction, the relation between learning at school and at home, to what extent the parents are supporting the school, etc. Considering the high 'yield' in mathematics in Japan, for example, it would seem to be especially fruitful to apply case study method to this country.

The Second International Mathematics Study (SIMS)

After the FIMS in the 1970s, an extensive curriculum reform movement took place throughout the world. The SIMS was conducted in the early 1980s to compare and contrast, in an international context, the varieties of curricula, instructional practices, and student outcomes across schools of twenty countries and educational systems (Travers and Westbury 1989, p. 1).

According to Robitaille and Nicol (1994), the design of SIMS was quite different from the FIMS. The SIMS was curriculum-based and focused on the study of mathematics through three different levels:

- the *intended* curriculum: the mathematics intended for learning by national and system-level authorities
- the *implemented* curriculum: the curriculum as interpreted by teachers and presented to students
- the *attained* curriculum: the curriculum learned by students and determined by their achievement and attitudes (Travers and Westbury 1989, pp. 6–7)

This framework provided direction for the development of instruments and the interpretation of results. It also included an analysis of curriculum content of participating countries. Taylor and Jenkins (2000) indicated that SIMS took much greater account of the influences of curricula, teaching approaches, school organization, teacher training, and student motivation than did the earlier attempts by FIMS.

Twenty education systems, including Belgium (Flemish), Belgium (France), Canada (British Columbia), Canada (Ontario), England and Wales, Finland, Hong Kong, Hungary, Israel, Japan, Luxembourg, Netherlands, New Zealand, Nigeria, Scotland, Swaziland, Sweden, Thailand, and United States of America, participated in one or more parts of SIMS, in which two population levels were chosen (populations sampled for this study were the same as FIMS). Teachers were asked to respond to questionnaires concerning the instructional strategies they used in teaching particular topics. Test items were presented in the form of multiple-choice, and test content was selected from areas of arithmetic, algebra, geometry, measurement, and descriptive statistics (Robitaille and Garden 1989; Travers and Westbury 1989).

An analysis of the results from this study is presented in a couple of volumes (Robitaille and Garden 1989; Travers and Westbury 1989). The major findings deal with changes since the FIMS in the *intended* curriculum. Many countries experienced curricula reform and change, mainly due to the influence of the new mathematics movement. In general, it was found that for most countries, curricula emphasis given to the study of geometry had decreased, while importance given to the study of algebra had increased. The results of the SIMS provided countries with important data that was used to make decisions regarding further change in curriculum and instruction.

However, in terms of limitations, the case study approach, suggested by Husen (1967b), was not adopted in the SIMS. Like FIMS, SIMS relied on teachers' reports on the way the curriculum was implemented in the classroom, and there was no attempt to validate the reported practice through classroom observation. Robitaille and Travers (1992, p. 696) recognized that 'self-report data must always be interpreted with caution.' Taylor and Jenkins (2000) pointed out that some countries assessed students at the beginning and end of the same school year, revealing some instability in the scoring which in itself raised questions related to the reliability of these assessment instruments.

In terms of achievement, the results showed that Japanese and Hong Kong students' performance among the highest in the SIMS, but the study did not provide an explanation for the achievement gap between East Asian and Western Countries. Robitaille and Travers (1992) suggest that to explain the achievement, inclusion of an elementary school population could provide very important information. They indicate that it would be useful to find out whether the kinds of achievement differences which are apparent by the time students reach the age of 13 are present as early as age 9 or 10 (p. 701).

The Third International Mathematics and Science Study (TIMSS)
The TIMSS was the largest and the most comprehensive international study of student achievement ever conducted by IEA at the time. In 1994–1995, it was conducted at five grade levels in more than 40 countries[2] (the third, fourth, seventh, and eighth grades, and the final year of secondary school). Students were tested in mathematics and science and extensive information about the teaching and learning of mathematics and science was collected from students, teachers, and school principals by questionnaires (Beaton et al. 1996). TIMSS also investigated the mathematics and science curricula of the participating countries through an analysis of curriculum guides, textbooks, and other curricular materials.

The TIMSS results were released from a series of reports. For example, the report *Mathematics Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study* (Mullis et al. 1997) describes student achievement in mathematics and science, respectively, for third and fourth graders in twenty-six countries. In terms of results of mathematics, Singapore, Korea, Japan, and Hong Kong were the top-performing countries at both the fourth and the third grades. In upper grades' achievement in mathematics, Australia ranked eleventh; in the lower grades' achievement in mathematics, Australia ranked ninth (Lokan et al. 1997).

Key findings from the study are summarized as follows (Mullis et al. 1997):

- First, in terms of gender differences, for most countries, the differences in mathematics achievement were small or essentially nonexistent.
- Second, in terms of attitudes to mathematics, the overwhelming majority of the fourth graders in nearly every country indicated that they liked mathematics.
- Third, in terms of educational resources, having educational resources in the home (i.e., computer, dictionary, own study desk, and 100 or more books in the home) was strongly related to mathematics achievement in every country.
- Fourth, in terms of teaching approaches, small-group work was used less frequently than other instructional approaches. Across countries, teachers reported that working together as a class with the teacher teaching the whole class, and having students work individually with assistance from the teacher were the most frequently used instructional approaches.

[2]Mainland China did not take part in this study.

- Last, in terms of teaching resources, the textbook was the major written source mathematics teachers used in deciding how to present a topic to their classes.

In terms of limitations, Wang (1998) suggests that in contrast to the former IEA studies, the limitations of TIMSS were less obvious, partly because of its thoughtful response to most research concerns in the past. However, two problems related to assessment instruments were highlighted. According to Wang (1998), the first problem was the free-response items (the first time presented in an IEA study) which constituted an important component of the TIMSS tests. This item by its definition should have different types of responses. But, the TIMSS Technical Report only suggests one correct answer. Second, although it was claimed that the item selection was grounded on an international consensus, a group of TIMSS researchers asserted that due to the tremendous curricula variability across nations and the desire to over-sample some topic areas, the TIMSS test varied in its match to any particular curriculum (Jakwerth et al. 1997).

TIMSS-Repeat and Trends in International Mathematics and Science Study (TIMSS)

TIMSS-Repeat is the second assessment in the series of TIMSS studies to measure trends in students' mathematics and science achievement. It was conducted by the International Study Center at Boston College in 1999 and included thirty-eight countries (see Footnote 2). The assessment measured the mathematics and science achievement of eighth-grade student (ages 13 and 14 years) and collected extensive information from students, teachers, and school principals about mathematics and science curricula, instruction, home contexts, and school characteristics and policies by questionnaires. Of the thirty-eight participating countries, twenty-six also participated in the 1995 TIMSS assessment, which enabled these countries to measure trends in their children's mathematics and science achievement and in school and home learning contexts.

According to *TIMSS 1999 International Mathematics Report* (Mullis et al. 2000), the findings of TIMSS-Repeat are summarized as follows:

- First, in terms of students' mathematics achievement, Singapore, Korea, Taiwan, Hong Kong, and Japan had the highest average performance, with Singapore and Korea having significantly higher achievement than all other participating countries. Australia ranked thirteen in the study.
- Second, in terms of students' environment and attitudes toward mathematics, on average internationally, students from homes with a high level of educational resources (e.g., more than 100 books) had higher mathematics achievement than students from homes with fewer resources.
- Third, in almost every country (except East Asian countries), there was a positive association between educational expectations and mathematics achievement. In each country, a more positive self-concept in mathematics was associated with higher average achievement. However, several countries with low percentages of students reporting a strong self-concept had high average mathematics achievement, including the five East Asian countries (Singapore,

Hong Kong, Chinese Taipei, Korea, and Japan). The TIMSS-Repeat data however did not provide an explanation for this finding.

- Fourth, in terms of the mathematics curriculum, in thirty-five of the thirty-eight countries, specifications for students' curriculum goals in mathematics were developed as national curricula. The exceptions were Australia, Canada, and the USA. System-wide testing and individual assessment were widely used methods to support curriculum implementation in most countries. Across countries, the *intended* curricula for eighth grade most commonly placed major emphasis on mastering basic skills and understanding mathematical concepts. Moderate to major emphasis was placed on assessing student learning, 'real-life' applications of mathematics, and communicating mathematically. Thirty-three countries reported at least moderate emphasis on solving non-routine problems, but working on mathematics projects was given minor or no emphasis in the *intended* curriculum of most countries.
- Fifth, in terms of instructional contexts and practices, internationally, 84 % of students were taught by teachers having mathematics and/or mathematics education as a major area of study. The TIMSS 1999 results showed that higher achievement was related to higher levels of teachers' confidence in their preparation to teach mathematics. Across the participating countries, teachers reported that the two most predominant activities encountered in mathematics classes were direct teacher instruction and teacher-guided student practice, accounting for nearly half of class time. Students in classes emphasizing reasoning and problem solving had higher achievement than those in classes with a low emphasis on these activities.
- Sixth, in terms of school factors, students in schools that reported being well-resourced generally had higher average mathematics achievement than those in schools where across-the-board shortages affected instructional capacity in mathematics.

In terms of their limitations, Taylor and Jenkins (2000) suggest that trend data from TIMSS to TIMSS-Repeat are very difficult to measure and interpret. They indicate that the same item may take on different significance from administration to administration, because of curriculum changes or wider societal influences, so the effects may be hard to disentangle. Regarding the methodology of TIMSS and TIMSS-Repeat, like the former IEA studies, many data were gathered through questionnaires. According to Stevenson and Nerison-Low (2002), the limitations of the questionnaire are obvious:

> As is the case with all methods used for the collection of behavioral and attitudinal data, questionnaires have both strengths and weaknesses…but interpretation of data from questionnaires is often more difficult than when there is an opportunity to interact with the respondents and to probe for details or elaboration of answers, as is possible with case studies. (p. 13)

Similarly, in terms of students' mathematics achievement, the Trends in International Mathematics and Science Study (TIMSS) 2007 (Mullis et al. 2008) provide the most recent and compelling achievement data. At the fourth grade, Hong Kong students had the highest average mathematics achievement relative to students from all other 36 countries. At the eighth grade, Chinese Taipei had the highest average mathematics achievement relative to students from 49 countries.

In summary, the IEA studies especially TIMSS and TIMSS-Repeat have found that Chinese students from Hong Kong and Taipei consistently outperform their Western counterparts in mathematics, and tested a considerable number of variables related to mathematics achievement, thereby providing a basis for this study. Case studies, then, might be used to explore the factors accounting for achievement differences between East Asian and Western countries (Husen 1967b; Stevenson and Nerison-Low 2002).

2.2.2 Other Comparatives Studies on Mathematics Achievement

Beside the IEA studies, there have been a number of other large-scale and small-scale comparative studies. In this section, three studies including the Michigan Studies, the International Assessment of Educational Progress (IAEP), and the Program for International Students Assessment (PISA) are reviewed.

The Michigan studies

For more than ten years the University of Michigan studies, under the direction of Harold Stevenson, has extended psychological studies of children's learning to include comparison of mathematics achievement (Paris and Wellman 1998). According to Stevenson and Stigler (1992), the research started in the mid-1970s and focused on the roots of academic achievement in East Asian countries and the USA. Five large metropolitan areas in three countries were chosen for the studies including Sendai, Japan; Taipei, Taiwan; Beijing, Mainland China; Minneapolis and Chicago, USA. Data were collected through interviews with teachers, pupils, parents and school heads, classroom observation, and tests. For mathematics, the samples included Year 1 and Year 5 students. To devise a test of children's mathematics achievement in different cultures, they extensively analyzed the mathematics textbooks and curricula used in Japan, China, and USA. Only concepts taught at the same grades in all three countries were included on the test. Beyond testing mathematics skills, their research program encompassed tests of reading, information acquisition, and basic cognition as well as observation and analyses of curricula, in-class teaching behavior, and at-home informal teaching (Paris and Wellman 1998, pp. xii–xiii).

The results of the achievement tests showed that students in the three Asian cities outperformed their American counterparts. There was a considerable learning 'gap' between American and East Asian countries. Stevenson and Stigler (1992) identified *cultural context* as the key factor causing the 'learning gap.' They indicated that the goal of primary school education in Japan and China was to teach children academic skills and knowledge. Educationalists and students attributed academic success to effort and showed relative disregard for innate abilities. In contrast, the US society placed a high priority on non-academic aspects of students' lives, and educationalists and students attributed academic success to innate abilities. They also found both Chinese and Japanese classrooms provided more opportunities than US classrooms for students to understand the goals and the sequence of the activities in which they were engaged (pp. 152–154).

In contrast to IEA studies, the strength of the Michigan studies was their stress on how mathematics was taught and learned in classrooms of different cultures. Independent classroom observation and interviews formed the basis of data collection, which was certainly more valid than relying on teacher reporting of their own classroom practice. Based on contextual data, it also provided some insights into understanding the 'learning gap.' The Michigan studies have been widely disseminated by respected journalists and educationalists. They certainly influenced Americans' view of mathematics and science achievement in American schools (Bracey 1993, 1996, 1998).

In terms of limitations, there are two aspects that need consideration. First, the studies involved countries where three different languages were used, yet most of the researchers were not trilingual. So the studies had to rely on different researchers collecting data in different countries, and on employing bilingual coders to summarize the classroom observation data (Stigler and Perry 1988, p. 213). This limitation creates a problem of validity of translation and observer–observer reliability in data collection. Secondly, the studies limited themselves to students in the first and fifth grades in primary school. In order to gain a more holistic picture of primary schooling, observation of other grades would be necessary.

The International Assessment of Educational Progress (IAEP)

According to Lapointe et al. (1992), in 1990–1991 twenty education systems including Brazil, Canada, China (Mainland), England, France, Hungary, Ireland, Israel, Italy, Jordan, Korea, Mozambique, Portugal, Scotland, Slovenia, Soviet Union, Spain, Switzerland, Taiwan, and the USA were surveyed by the Center for the Assessment of Educational Progress in Princeton. The survey focused on the mathematics and science performance of 13-year-old students (14 countries also assessed 9-year-olds in the same subjects) as part of the Second IAEP. The curriculum model adopted by the study was similar to SIMS. While recognizing the fundamental differences in intended curriculum from country to country, the participants assembled tests that focused on the common elements of their curricula. In order to form the context for interpreting these student achievement data, they added sets of questions about students' home background and classroom experiences and characteristics of the schools they attended.

The major findings of the IAEP were that at age 13, based on average percent correct scores, Mainland China (80 %), Korea (73 %), and Taiwan (73 %) outperformed all other countries[3] in the study. In almost all 13-year-old samples, at least 10 % of the students performed very well (20 points or more above the IAEP

[3] Australia did not take part in this study.

average). In Mainland China, however, *even students in the 10th lowest percentile performed close to the IAEP average*.

Lapointe et al. (1992) suggested that this result is counter-intuitive because countries such as Mainland China with less well-trained teachers, large classes, and poor-quality instructional materials produce students achieving the highest scores, while countries such as the USA, with better trained teachers, who work in schools that are more generously supported, perform under average (p. 5). To understand this paradox, it was suggested (Lapointe et al. 1992) that factors impacting on academic performance interact in complex ways and operate differently in various cultural and educational systems and there is no single formula for success. It was also proposed that factors such as teaching practices, types of instructional materials, teacher background, and classroom organization vary from country to country for children at age 13. However, these factors did not distinguish between high-performing and low-performing populations. The IAEP study itself did not provide adequate explanation to justify its findings. Clearly, like IEA studies, questionnaires and pencil-and-paper tests were used as the major method to collect data, which limited the possibility of validating the data and explaining the findings in depth.

In terms of generalizability of the findings, Lapointe et al. (1992) recognized the limitations of the study. They indicated that in some countries participants assessed virtually all age-eligible children in their countries and in other countries they confined samples to certain geographic regions, language groups, or grade levels. In some countries, significant proportions of age-eligible children were not represented because they did not attend school. It was also found that in some countries, low rates of school or student participation meant results could be biased (p. 3). They also admitted 'the resulting tests do not match all countries' curricula equally well' (p. 9). As Bell and Kang (1995) point out:

> Although reputable projects such as The Third International Mathematics and Science Study and The International Assessment of Educational Progress establish 'frameworks' which aim to control for the effects of intended, implemented and attained curricular differences, variables such as rate of learning, educational goals and expectations, and even the meaning of achievement itself (as it affects test performance) can never be effectively calibrated. (pp. 43–44)

In short, from the results of the Michigan and the IAEP studies, it can be inferred that in terms of mathematics achievement, Mainland China is on a par with other high achieving East Asian countries.

The Program for International Student Assessment (PISA)

The Program for International Student Assessment (PISA) comprises a series of surveys about the knowledge and skills of 15-year-olds in principal industrialized countries coordinated by governments of participating countries through the Organization for Economic Co-operation and Development (OECD). PISA is an on-going program of assessment that gathers data on the reading, mathematics and science skills of 15 year-olds in a cycle every three years. According to the report (Organization for Economic Co-operation and Development (OECD) 2000), the survey was conducted initially in 2000 with 265,000 students from 32 countries (of which 28 are OECD members). Australia, Austria, Belgium, Brazil, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Latvia, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Russian Federation, Spain, Sweden, Switzerland, the United Kingdom, and the USA participated in pencil-and-paper student assessment. Students and their school principals also completed questionnaires about their learning experiences and their schools.

In contrast to the IEA and other comparative studies, PISA claims that it assesses young students' capacity to use their knowledge and skills in order to meet real-life challenges, rather than merely looking at how well students have mastered a specific school curriculum. Assessment items are a mixture of multiple-choice test items and questions requiring the student to construct their own responses. The items are organized in groups based on real-life situations. According to Schleicher (2000), the underlying principle of PISA is a dynamic model of lifelong learning in which new knowledge and skills necessary for successful adaptation to changing circumstances are continuously acquired over the life cycle (p. 66).

The major findings of PISA 2000 (Organization for Economic Co-operation and Development (OECD) 2000), in terms of mathematical literacy, were that students in Japan and Korea displayed the highest mean scores (see Footnote 1). Although the tasks for the PISA assessment of mathematical literacy were designed so that students not using calculators would not be disadvantaged, students were allowed to use their own calculators or those provided by test administrators. But there was no indication that the use of calculators provided an advantage to students in terms of their performance in PISA.

The comparison between spending per-student and mean student performance across countries cannot be interpreted in a causal way. Nevertheless, these data reveal a clear positive association between the two. At the same time, spending on educational institutions is considered a necessary prerequisite for the provision of the high-quality education. The comparison also suggested that spending alone is not sufficient to achieve high-level outcomes and that other factors, including the effectiveness with which resources are invested, play a crucial role. But like other large-scale studies, the report itself did not adequately explain the rationale underlying its findings.

In terms of mathematics achievement gap, the OECD PISA 2012 provided the most recent example. About 510,000 15-year-old students from 65 countries and economies took part in the assessment of mathematics literacy; Shanghai-China achieved the highest score, followed by Singapore, Hong Kong and Taipei scored higher than all other countries including Australia, who ranks 17th (Organization for Economic Co-operation and Development (OECD) 2014; Thomson et al. 2013).

However, being a large-scale comparative study, PISA also has both strengths and limitations, as Schleicher (2000) points out:

> OECD's objectives are ambitious. For the first time an international assessment of school students aims to determine not just whether they have acquired the knowledge specified in

> the school curriculum, but whether the knowledge and skills that they have acquired in childhood have prepared them well for adult life. Such outcome measures are needed by countries which want to monitor the adequacy of their education systems in a global context. The ideal will not be instantly achieved, and some of OECD/PISA's goals will initially be constrained by what is practicable in an assessment instrument that needs also to be reliable and comparable across many different cultures. (p. 77)

It is recognized that one of the contributions of PISA to comparative studies is its dynamic model of lifelong learning which extends comparisons beyond the school curriculum (Lokan et al. 2001). But, like other large-scale studies, the limitations in methodology and instruments used may have influenced the validation of its findings.

2.2.3 Key Findings of Comparative Studies

From the preceding overview of comparative studies on mathematics achievement, it is clear that students of East Asian countries, such as China, Korea, Japan, and Singapore, consistently outperform their counterparts in mathematics. China includes Mainland China, Taiwan. Hong Kong and Macao, but in the comparative studies they were regarded as separate education systems. Taking TIMSS as an example, in 1995, 9-year-old students from Hong Kong outperformed all other Western countries (Mullis et al. 1997). This pattern of achievement was confirmed by TIMSS-Repeat in 1999, in which the mathematics performance of eighth-grade students in Hong Kong was superior to that of any Western countries (Mullis et al. 2000). It was further confirmed by TIMSS 2007 (Mullis et al. 2008), in which at the fourth grade Hong Kong students had the highest achievement in mathematics; at the eighth grade, Chinese Taipei had the highest achievement in mathematics. It was recently confirmed by PISA 2012 (Organization for Economic Co-operation and Development (OECD) 2014), in which, China's Shanghai achieved the highest score. It was also confirmed by the former IAEP (Lapointe et al. 1992) in which Mainland China, Taiwan, and Korea were the top three in terms of average percent correct scores, and by former Michigan Studies (Stevenson and Stigler 1992) in which Chinese and Japanese primary school students outperformed American primary school students in mathematics. These comparative studies corroborate evidence that the 'achievement gap' in mathematics between China and Western countries is continuing and may possibly be widening. This promotes further research questions as to the underlying reasons for this apparent 'gap.'

2.3 Factors Contributing to the 'Achievement Gap'

According to Graf and Leung (2000), comparative studies such as SIMS and TIMSS produced survey data indicating that there may be some system-based reasons for differences in achievement between East Asian countries (such as

China, Japan, Korea, and Singapore) and Western countries (such as North America, Europe, and Australia). To explain the 'achievement gap,' researchers have described various factors that contribute toward it. In this section, key factors contributing to the achievement 'gap' are discussed as follows:

2.3.1 Cultural context
2.3.2 Teachers' mathematics knowledge
2.3.3 Classroom teaching
2.3.4 Out-of-school curricula
2.3.5 Standards and content of curriculum
2.3.6 Form and content of assessment

2.3.1 Cultural Context

It has been widely acknowledged that mathematics education in any country is influenced by its cultural context (Bishop 1988; Brown et al. 1989; Graf and Leung 2000; Lave 1988; Leung 2001; Stevenson and Stigler 1992; Stigler and Baranes 1988). According to Graf and Leung (2000), East Asian countries by and large can be regarded as countries sharing a Confucian culture. In contrast, Western countries can be broadly identified as countries sharing a Greek/Latin/Christian culture.

In the Michigan Studies, Stevenson and Stigler (1992) suggest that cultural difference is the main reason accounting for the 'learning gap' between East Asian and Western countries. One of the key issues considered in their argument supporting cultural factors is that of students' attitudes toward ability and effort in achievement. They assert that students in East Asian countries attributed mathematics achievement to effort. In contrast, American students attributed mathematics achievement to innate abilities.

Leung (2001) explains that there are distinctive features about mathematics education in East Asia, and that those features are expressions of distinctive underlying cultural values. These features and values of East Asian mathematics education, which may be contrasted with those of the West, are described as six dichotomies (Leung 2001):

- *Product (content) versus process*
  In East Asia, mathematics is viewed as a body of knowledge dealing with mathematics principles and relationships of numbers and geometry. Mathematics education focuses on acquiring this knowledge through teacher-directed guidance. In contrast, the contemporary Western view emphasizes a constructivist view of knowledge where learning is perceived of as a process of active construction by the learner.

- *Rote versus meaningful learning*
  In East Asia, memorization and repeated practice are the accepted ways of learning. In contrast, Western mathematics educators advocate learning for meaning, emphasizing that students should first understand before they memorize (although rote learning is still practiced).
- *Studying hard versus pleasurable learning*
  The traditional view in East Asian countries holds that studying is a serious endeavor and students are expected to work hard to achieve success in learning. In contrast, pleasurable learning has been advocated in a number of Western countries. Western educators advocate that learning should be a positive experience for the child; learning should be enjoyable.
- *Extrinsic versus intrinsic motivation*
  Educators in East Asia emphasize extrinsic motivation and completing examinations have traditionally been an acceptable source of motivation for student learning. According to Leung, the East Asian education systems are characterized by highly competitive examinations, and this is deeply rooted in Chinese culture. China was the first country in the world where a national examination system was established. Paper-and-pen examinations are considered a fair method for selection of the able from the less able, and are highly valued. Hence, in China, examinations are considered a legitimate source of motivation for students to learn and to demonstrate that they have achieved mathematics knowledge. In contrast, educators in the West promote intrinsic motivation in students' learning mathematics. They consider the best way of motivating students is by getting students interested in mathematics. Extrinsic motivation, such as that derived from examination pressure, is considered somewhat 'harmful' to students' learning and student self-esteem, particularly in the primary school.
- *Whole-class teaching versus individualized learning*
  Because of the strong social orientation of East Asians, learning together in a collaborative social setting is highly treasured. Teaching usually takes place in a large group setting with the teacher directing the learning to the whole class. In contrast, Western culture stresses independence and individualism, and individualized teaching and learning in as small a group as possible is considered the ideal. However, whole-class teaching is still common practice.
- *Teacher as an expert versus facilitator*
  The image of the mathematics teacher in the East Asian tradition is as an expert in mathematics (teacher-scholar). This is deeply rooted in the Confucian culture where the teacher should primarily be a scholar before they are able to fulfill the role of a facilitator of learning. *In primary schools teachers usually subject specialize (i.e., mathematics).* In contrast, the primary role of the teacher in Western culture is often perceived as a facilitator of learning, helping, and guiding students to learn. In Australian primary schools, teachers usually teach all subject areas.

In short, these factors underlying some differences between mathematics education in East Asian and Western countries, suggested by Leung, are all based essentially on different cultural values.

This debate on cultural versus non-cultural factors is difficult to resolve through published comparative studies, because nearly all of them are focused on comparing mathematics achievement in different countries where cultural influence is confounded with other factors. To determine the influences of cultural factors on students' achievement, there is a need for new studies investigating the same cultural group in different countries, such as a comparison of the achievement of Chinese students in China and in other countries such as Australia and America. This can provide a different form of comparative data that elucidates cultural influence in authentic settings.

2.3.2 Teachers' Mathematics Knowledge

In contrast to cultural factors, the role of teacher knowledge has received much recent attention in the cultural versus non-cultural debate. Ma (1999) conducted a comparative study of US and Chinese elementary school teachers' mathematical knowledge as it related to classroom teaching practice. She suggested that Chinese teachers began their teaching careers with a better understanding of elementary mathematics than that of most US elementary teachers (p. xvii). She argued that 'unlike factors outside of classroom teaching, teachers' knowledge might directly affect mathematics teaching and learning. Moreover, it might be easier to change than cultural factors' (pp. xix–xx).

Ma (1999) compared the responses of twenty-three American elementary school teachers and seventy-two Chinese teachers to the same four questions. The interview questions were designed to present mathematics through a hypothetical classroom situation. She found that there was a striking contrast in the knowledge of the two groups of teachers studied. US teachers tended to be 'procedurally focused.' Most showed sound algorithmic competence in the two elementary topics, but had difficulty with the two more advanced topics. Most of the Chinese teachers demonstrated algorithmic competence as well as conceptual understanding of all four topics.

Based on these data, *Ma argued that Chinese teachers have a much better grasp of the mathematics they teach than do American teachers*. She suggested:

> Elementary mathematics can be viewed as fundamental mathematics. Elementary teachers need 'Profound Understanding of Fundamental Mathematics' (PUFM), which means an understanding of the terrain of fundamental mathematics that is more than sound conceptual elementary mathematics. It is deep, broad, and thorough. A teacher with PUFM is not only aware of the conceptual structures and basic attitudes towards elementary mathematics, but also is able to teach them to students effectively. (p. 124)

Following this conceptual framework, Ma indicated that elementary mathematics is constructed very differently in China from the USA. She found that ten percent of Chinese teachers displayed such understanding of this difference, and eighty percent displayed at least some of that understanding. American teachers, however, appeared to be lacking in this area.

After she interviewed three teachers, whom she identified as having PUFM, Ma found several factors contributed to PUFM attainment: learning from colleagues, learning mathematics from students, learning mathematics by doing problems, teaching class by class (in China, it means teaching in a cycle from Year 1 till Year 6, then back again), and studying materials intensively. According to Ma, Chinese teachers develop PUFM during their teaching careers, stimulated by a concern for what to teach and how to teach it, inspired and supported by their colleagues and teaching materials.

According to Sherman (2000), Ma's work has been well received on both sides of the so-called math wars, an engaging vitriolic debate on the proper direction of mathematics reform in the USA. Her study was praised for her focus on real understanding as opposed to mere computational competence. However, Sherman (2000) points out:

> But most NCTM reformers believe that understanding comes from 'constructivist', or discovery learning. They give low priority to computational skills (unnecessary, they say, if students have calculators). Their goal is not to teach traditional mathematical subjects as logically coherent systems but to develop 'higher-order' thinking skills that are supposed to be transferable to all subjects. Ma, in contrast, stresses the importance of mastering facts and procedures. She does not advocate child-centered classrooms ('teach the child, not the subject') in which the curriculum is secondary to fostering juvenile self-esteem. Unlike NCTM reformers, Ma does not see computational competence and real understanding as antithetical ('drill and kill') but as mutually reinforcing. (p. 86)

It must be acknowledged that there was no detailed information about Ma's selected sample. She did not indicate how the 23 teachers from the USA were chosen for the American sample. Surely the US 'procedural teachers' recognized by Ma's study were not following the NCTM guidelines. Although, Ma's sample of 72 Chinese teachers showed they had better mathematics knowledge, student learning was not necessarily better because of teachers' knowledge. This inconsistency raises the question about whether the Chinese sample is representative. For international comparison and generalizations to be made, the samples of students and teachers must be representative of the nation being compared, and they must be comparable to each other (Bracey 1998).

Recently, Leung (2002) replicated Ma's study in Hong Kong and Korea and found that unlike Ma's Chinese samples, Hong Kong and Korean teachers' reported *teaching strategies that were very procedural*. When probed, it was found that these teachers fully understood the concepts behind the procedures, but deliberately taught in a procedural manner for pedagogical reasons. This evidence indicates that *teacher knowledge alone does not necessarily determine teacher's classroom practices and students' mathematics achievement*. To find out what is really happening in Chinese classrooms, we cannot rely on what teachers say they do; there is need for classroom observation to validate what teachers report.

2.3.3 Classroom Teaching

Researchers in comparative education are increasingly becoming aware of the need to look at instructional practices in the classroom (not what teachers *say* they do in the classroom) as the basis for interpreting data on mathematics achievement (Burstein 1992; Cobb et al. 1992; Leung 1995; Stigler and Hiebert 1999). In this section, three studies are reviewed.

The Michigan study

In their Michigan Studies, Stevenson and Stigler (1992) observed teaching practices in Chinese, Japanese, and American primary schools. They found that the techniques used by Chinese and Japanese teachers are ones often recommended by American educators, but not broadly applied in the USA. They indicate:

> [Japanese and Chinese classroom] consist of coherent lessons that are presented in a thoughtful, relaxed, and no authoritarian manner. Teachers frequently involve students as sources of information. Lessons are oriented toward problem solving rather than rote mastery of facts and procedures, and make use of many different types of representational materials. The role assumed by the teacher is that of knowledgeable guide, rather than that of primary dispenser of information and arbiter of what is correct. There is frequent verbal interaction in the classroom as the teacher attempts to stimulate students to produce, explain, and evaluate solutions to problems. (Stevenson and Stigler 1992, pp. 176–177)

Stevenson and Stigler (1992) indicate that these techniques have not been broadly applied in the USA, because American teachers are lacking the training and time necessary to prepare lessons and the opportunities to share experiences with each other; American teachers find it difficult to organize well-crafted lessons. Based on their classroom observations, they suggest that different teaching techniques may lead different learning outcomes between East Asian countries and the USA. According to Robitaille and Travers (1992), the Michigan Studies have made an important contribution to our knowledge of the teaching and learning of mathematics across cultures. Through classroom observation, they provided data on the contexts in which teaching and learning take place in different cultures (p. 706).

Classroom teaching in Beijing, Hong Kong, and London

Building on his own study of *The Mathematics Classroom in Beijing, Hong Kong and London*, Leung (1995) attributes different beliefs pertaining to mathematics, and beliefs about mathematics teaching and learning, to differences between Chinese and Western cultures. He argues that differences in classroom practices relate primarily to differences in beliefs. He also found that the schools in these three cities differed considerably in their physical setting, facilities, and structure, and these differences must be taken into account.

There was much more concern for individual differences in London, as manifested by schools adopting individualized learning programs, by teachers spending less time on whole-class instruction, more time on seatwork and allowing students to proceed at different paces.

In contrast, in Beijing and Hong Kong, students were more exposed to direct instruction from the teacher. In Beijing, mathematics lessons were more structured and there was greater stress on mathematics concepts, while in Hong Kong more emphasis was placed on practicing mathematical skills. Leung indicates that Beijing teachers emphasized the content of mathematics; Hong Kong teachers emphasized mathematical skills, while London teachers emphasized active learning and enjoying mathematics (p. 312). He suggested that the classroom practices of Beijing teachers reflected a more rigid view of mathematics than the practices in Hong Kong and London. There was also a much greater stress on memorization in the teaching at Beijing. He also found that classroom teaching in all three places was influenced to some extent by examinations, but this was not explained in detail.

It is apparent that classroom teaching practices in Beijing described by Leung (focusing on content and memorization) are quite different from those observed by Stevenson and Stigler (focusing on process and understanding). These inconsistent descriptions of Chinese mathematics classrooms raise questions about the range of classrooms selected for observation. The observations of Stevenson and Stigler may have reflected to a large extent the ideal classroom practices the teachers wanted to portray. According to Leung (1995), teachers in Beijing followed the textbook very closely and the teacher's guide in which the ideal lessons were presented in detail. This may not have reflected the apparent deep teacher's knowledge of mathematical pedagogical practice.

TIMSS video study

As part of TIMSS, Stigler and Hiebert (1999) analyzed videotapes from dozens of randomly selected eighth-grade mathematics classrooms in the USA, Japan, and Germany. They found that American teachers presented about twice as many definitions as Japanese and German teachers. They also discovered that American lessons were devoid of mathematical proofs. About ten percent of German lessons and more than half of the Japanese lessons contained such proofs. They also indicated that Japanese and American teachers organized their lessons quite differently. In a typical American lesson, a teacher reviewed homework, demonstrated how to solve the problem of the day, gave students classroom practice, corrected students' work, and assigned homework.

In contrast, Japanese teachers reviewed the previous lesson, presented the problem of the day, and set the students to working on its solution either individually or in groups. The class then discussed problem solutions (some problems had more than one solution), often led from the blackboard by students who thought they had successfully solved the problem. American students almost never led such a discussion. They suggested that most efforts to improve mathematics education in America have failed because they simply do not have sufficient impact on the quality of teaching inside classrooms. They argue that American teachers are not incompetent, but the methods they use are severely limited, and American teaching practice has no system in place for improvement.

However, Bracey (2000) disagreed with Stigler and Hiebert's view on classroom teaching as the major factor accounting for the achievement gap. Bracey argued that American and Japanese teachers organize their lessons differently largely because they believe different things about what mathematics is and how to teach it. He asserts that these differences are in no way related to differences in achievement. German teachers' methods are similar to Japanese teachers, yet Japanese students attained much better scores than American students, and German students scored the same as Americans. Bracey (2000) clarifies this:

> It could well be that the Japanese students score higher because so many of them attend juku-cram schools-after school and on weekends. Juku specialize in teaching students how to take tests. 'The Teaching Gap' does not discuss the role of Juku. A Japanese educator who, at my request, watched some mathematics and science study tapes, concluded that, indeed, the Japanese classes were more conceptually oriented. He felt, though, that Japanese teachers are free to teach conceptually because they can count on family support to ensure that less glamorous mathematics activities will be completed at home. American teachers cannot count on such support. (p. 54)

Bracey's view was supported by Schumer's study. In the paper, *Mathematics Education in Japan* (Schumer 1999), the author argued that we cannot limit observation to mathematics in Japanese public schools when we want to gain insights into the achievement of Japanese students. Homework, voluntary studies at home, and private supplementary lessons are more important in Japan than in Germany or in the USA.

However, like the debate on cultural versus non-cultural factors, this influence of classroom and non-classroom factors is also difficult to be measured by current comparative studies, because these studies focus on comparative classroom factors. Little attention is paid to out-of-classroom factors. To understand out-of-classroom influences on students' achievement, data on students' after-school learning and home study are needed in further comparative studies.

2.3.4 Out-of-School Curricula

In the November 2001 issue of *Mathematics Education Dialogues* (NCTM), educators discussed what the research community should learn from comparing the performance of Asian and North American students and about the factors that might contribute to the superior performance of Asian students. A considerable number of contributors (Ng 2001; Shimizu 2001; Uy 2001; Whang 2001) suggested that out-of-school curricula might be one of the most important factors accounting for Asian students' high achievement in mathematics. For example, Lin (2001) argues that the high performance of 13-year-old Taiwanese students on TIMSS is a result of out-of-school studying. He indicates that Taiwanese children have no choice but to attend out-of-school 'cram' lessons or to receive private tutoring to achieve their expectations of a university education. Students work very hard to pass examinations, which puts considerable stress not only on the child but also on the child's family. It becomes a natural responsibility for parents to look for 'star' teachers at good 'cram' schools for their children. According to Lin, thirteen-year-old Taiwanese students spend as many as 15 h per week for studying mathematics.

Similarly, in Japan, Shimizu (2001) argues that an explanation for the achievement gap is complex and we must consider students' perspectives. In 2000, Shimizu surveyed eighty students (who took part in TIMSS in 1995). He found that Juku School (where students spend extra hours after school and where mathematics was regarded as one of the key subjects) is an important factor identified by the students as contributing to their high achievement in mathematics.

However, according to Leung (2002), TIMSS data suggest that students in East Asian countries such as Japan do not spend more time out of school doing mathematics homework than that of their counterparts in Western countries. In contrast, Ueno (2002) pointed out that TIMSS data do not reflect the reality of Japanese out-of-school mathematics learning. According to him, Japanese students regard 'Juku' as part of school learning and using mathematics textbooks at home is not considered as doing homework. Thus, the TIMSS data may not have accurate data for much of the out-of-school work. This inconsistency raises an interesting question about the validity of data from student questionnaires. In order to understand what mathematics education really happens in school, or out-of-school, more in-depth qualitative study is needed.

2.3.5 Standards and Content of Curriculum

As published comparative studies of mathematics achievement include curriculum analysis in the early stages of designing instruments for testing students' achievement, a comparison of standards and curriculum content between countries is a natural process in comparative studies of mathematics achievement. Lindquist (2001) indicates that curriculum standards and content might account for differences in mathematics performance between East Asian countries and the USA. McLeod et al. (1996) suggest that when compared to the American curriculum, the Japanese curriculum is more intense. American textbooks are about three times as thick as those of other nations, but covering topics only briefly and shallowly. In an in-depth analysis of mathematics curricula, Schmidt, McKnight, and Raizen (1997) describe the vision shaping US mathematics education as 'splintered.' This was seen through inconsistencies in curriculum planning, in textbooks, and what teachers actually teach. In contrast, both an analysis of mathematics curricula and the video study of the TIMSS showed Japanese mathematics teaching as coherent throughout the curricula and also within classroom lessons. There was a close match between *intended, implemented,* and *achieved* curriculum.

MacNab (2000) argues that standards of attainment in school mathematics are closely connected to belief systems regarding values and purpose; that American education systems do not always collectively offer a credible and coherent vision for mathematics education which can be effectively implemented in school classrooms; and that this coherence of vision is to a large extent what characterizes the higher performing TIMSS countries.

However, Campbell and Kyriakides (2000) challenge the view that setting high national curriculum standards contributes to high achievement. They made a comparison of the standards in primary school mathematics in England and Cyprus and found that despite higher-expected standards, Cypriot pupils achieved lower standards in TIMSS than their English counterparts. They suggest four possible explanations for this apparent paradox, namely problems with the TIMSS methodology, cultural influences, differential time frames, and poor understanding of the different conceptions of 'standards.'

This debate on the relationship between curriculum standards and mathematics achievement cannot be resolved by current comparative studies, because the test items in these studies assumed common standards of mathematics across countries.

2.3.6 Form and Content of Assessment

Recently, Cai (2000) found that the superior performance of East Asian students might be related to the form and content of tests. He indicated that Chinese students' superiority was limited to certain aspects of mathematics achievement, including basic skills of computation and solving routine problems. Western students performed better than their Asian counterparts in aspects such as using visual and graphical representations and solving open-ended problems. Cai (2000) indicated that students in the USA rely on concrete strategies and experiences to solve mathematics problems more frequently than their Chinese counterparts. But he suggested that the use of manipulatives or concrete experience does not guarantee students' conceptual understanding and high performance in paper-pen examinations, in which more efficient and generalized solution strategies are needed.

There are considerable differences between Cai's studies and other comparative studies such as the Michigan Studies. According to Stevenson and Stigler (1992), in the Michigan Studies, they used both typical mathematics problems (routine questions) and novel problems (open-ended questions) to test students. The content of their tests included word problems, number concepts and equations, estimation, operations, geometry, graphing, visualization and mental folding, mental calculation, and oral problems. They found that for nearly every item the mean score for the American students was the lowest. Stevenson and Stigler (1992) claimed that the East Asian students' superiority was not restricted to a narrow range of well-rehearsed, automatic computational skills, but was manifest across all the tasks including open-ended tasks (pp. 39–41). These findings raise an interesting question related to the instruments used in comparative studies for testing students' achievement. However, this cannot be resolved by current comparative studies, because nearly all of them used a similar format in which most items were presented in multiple-choice format. More analysis on the assessment process may well explain inherent differences in comparative studies (Geary et al. 1998).

2.4 Strengths and Limitations of Current Comparative Studies

It is widely acknowledged that large-scale international surveys in mathematics education have both strengths and limitations (Martin and Mullis 2000; Taylor and Jenkins 2000; Theisen and Adams 1990; Wang 1998). In this section, the advantages and disadvantages of some current studies are discussed in terms of their research frameworks and methodologies.

2.4.1 Research Frameworks

According to Howson (1998), large-scale comparative studies such as the IEA studies have adopted an inductive framework, analyzing situations and contexts, collecting data and analyzing the results (most in terms of rankings, percentiles, and correlations), and generating some hypotheses along the way which the data can help to illuminate. In terms of the strengths of this approach, Theisen and Adams (1990) indicated that the IEA activities have collected a large amount of data and have made significant contributions to education. They have drawn the attention of key educational and political actors to the substantial differences that exist among the educational systems of the world in terms of the academic achievement they produce. Comparative studies have also shed light on why some of these differences occur and have highlighted a range of inefficiencies in the schooling process (p. 290). Taylor and Jenkins (2000) indicate that the IEA studies represent a collaboration of over forty-five countries and provide data that have offered a distinctive contribution to our understanding of the effects of curriculum and school organization upon learning and the relationship between achievement and pupil attitudes (pp. 13–14).

On the other hand, in terms of limitations, Howson (1998) points out that large-scale comparative studies seldom had a hypothetical deductive model, in which specific hypotheses are tested, which from the beginning determine the whole design, analysis, and reporting. Taylor and Jenkins (2000) suggest that small-scale, in-depth explorations are needed complementing the insights generated by other kinds of approaches. Other kinds of questions are needed, which can yield insights often missed (or not sought) in large-scale enterprises, an end product that is justification in itself as well as providing potential hypotheses for wider investigation.

The IEA studies have tried to address curriculum as a broad explanatory factor underlying student achievement (Martin and Mullis 2000; Robitaille and Garden 1989; Schmidt et al. 1997). However, countries' rankings, published by IEA studies, are often the first focus point for politicians and policymakers (Cohen 1998; Taylor and Jenkins 2000). Taylor and Jenkins (2000) suggest:

> We need also not just to relate educational achievements to these wider influences but also to the institutions and values of the society, and to a model of how educational systems as a whole work and reflect curriculum-based learning. In an international context, the model also needs to be sufficiently general to measure the global effects of schooling on student learning. We are still far from understanding the link between the curriculum as 'intended', the curriculum as 'implemented' and the curriculum as 'received' by pupils, particularly in such a wide range of socio-cultural and pedagogic contexts. (p. 21)

2.4.2 Limitations in Methods

The IEA activities have made a significant contribution to methodologies and use of analytic tools available (Lokan 2000; Martin and Mullis 2000; Theisen and Adams 1990). According to Martin and Mullis, the design and development of IEA studies were complexes involving a variety of representatives from the participating countries, technical advisors, mathematics and science educators and subject matter specialists, and the major funding agencies. Taking TIMSS as an example, Wang (1998) indicates that it is the largest, most comprehensive, and most rigorous international study in comparative education. The TIMSS researchers employed multiple research methodologies, including assessments, questionnaires, curriculum analyses, and videotapes to enrich the public understanding of the findings.

However, there are considerable challenges associated with designing large-scale comparative studies. According to Eckstein and Noah (1992), in both their form and content, the IEA tests were required to be internationally uniform and acceptable, to make possible comparisons of levels of cognitive achievement across nations. In meeting this aim, IEA tests may have omitted important elements of curriculum and assessment in particular nations. Other researchers have noted that in the effort to make the studies comparable across nations, methodological and content sacrifices have been made at the expense of a more complete understanding of variations within individual countries. That is, if more carefully researched, these data might better inform national policies designed to improve the educational status of children (Theisen and Adams 1990). Martin and Mullis (2000) confirm that 'among the major components of TIMSS that posed methodological challenges were curriculum analysis, achievement testing and estimation of proficiency' (p. 34).

Theisen and Adams (1990) argue that since nearly all of the international comparisons were affected through the variation in administration of test items and questionnaires, their limitations are obvious. Eckstein and Noah (1992) confirmed:

> Naturally much of the information required about the class and school context is difficult to measure precisely using questionnaires, especially since much of the information is retrospective. Moreover, IEA work has not been considered in light of the implications of assessment in different countries. (p. 166)

As suggested by Theisen and Adams (1990, p. 708), to find out what really goes on in mathematics classrooms, we need to observe what teachers actually do there, and not rely solely on what they say. In addition, most former international comparative researchers only focused on the 'attained curriculum' by using their uniform examinations to test the student achievement but neglected assessment practices already completed at the local school and classroom level.

Researchers in comparative curriculum studies are increasingly aware of the need to look at instruction and assessment practices in the classroom as characteristic of the curriculum and as the basis for interpreting data on student outcomes (Leung 1995). There are some comparative studies (e.g., Leung 1995; Stevenson and Stigler 1992; Stigler and Hiebert 1999) on the mathematics curriculum that include a component of classroom observation to account for differences in *implemented* curriculum, but little research has been done on comparative study of assessment practices at school and class level. As Eckstein and Noah (1992) noted, the study of assessment practices seems to have been a largely neglected aspect of comparative education.

Clearly, there is a lack of in-depth research that investigates curriculum and assessment factors contributing to apparent differences in student achievement. There is especially a lack of comparative study incorporating descriptive knowledge of assessment practices and examining similarities and differences in 'authentic' settings. This requires qualitative studies of cases of apparent differences in curriculum and assessment practices, internationally.

In terms of case studies, Husen (1967b) suggested that for further research, the IEA might consider case studies of particular countries. He recommended that such studies would have to employ thorough analyses of curricula, textbooks, other learning materials, and the relation between learning at school and at home, and so forth (p. 309).

> We never seriously considered an alternative strategy, for example, limiting ourselves to selecting of a few schools and classrooms that could be subjected to intensive, qualitative observations. We certainly expected too much from the broadly collected information that was obtained by questionnaires from the students about their home background and from the teachers about how they taught. (Husen and Postlethwaite 1994, p. 646)

In terms of methodology, Arnove (2001) suggests that:

> Case studies are likely to continue to be the most commonly used approach to studying education-society relations. Given the limited resources of most researchers working in the academy, the tendency of most individuals is to study that with which they are most familiar. More than just convenience, Charles Gagin argues that "the comparative method is essentially a case-oriented strategy of comparative research" because of the need to take into account the contingencies of particular sociocultural milieus and historical information. The value of case studies resides in their contribution to the refinement and modification of

> extant theory and ultimately to the creation of new theory when existing explanatory framework are not applicable. (p. 490)

In summary, current comparative studies have attempted to address most of the concerns related to students' mathematics achievement raised in the research community by using a variety of methods. Although a mindful position was taken by researchers to downplay the 'horse race' feature of comparative studies on mathematics achievement, due to the limitations of large-scale studies and its focus on achievement, little attention has been paid to problems of assessing empirical factors in educational contexts. There is a real need to use case studies to compare the apparent differences in curriculum and assessment practices cross-culturally, in order to understand the 'achievement gap' comprehensively.

2.5 The Relationship Between Teaching, Learning, and Assessment

It is widely acknowledged that student's achievement is affected not only by officially prescribed mathematics curricula, but also by the classroom teaching and assessment (Black and Wiliam 1998; Bright and Joyner 1998; Carraher et al. 1990; Clarke 2000; Clarke and Stephens 1996; Leder and Forgasz 1992; National Council of Teachers of Mathematics (NCTM) 1995, 2000). In this section, some key studies on mathematics assessment are briefly reviewed including the nature of mathematics assessment, and the relationship between teaching, learning, and assessment.

According to Clarke (1996), contemporary mathematics assessment reflects an international consensus regarding the nature of mathematics learning. He argues that assessment should be recognized, not as a neutral element in the mathematics curriculum, but as a powerful mechanism for the social construction of mathematics competence.

According to Leung (2001), the East Asian education systems are characterized by highly competitive examinations. Teachers and parents attach great importance to education, and there is high expectation for their students and children to succeed. Academic achievement has been considered a means of bringing honors to one's family. This high expectation on students to succeed provides another important source of motivation for students to learn. In addition, the difference in social and economic status between the academically high qualified and low qualified is much greater in East Asia than in the West. This sends a message to students about the importance of academic achievement and constitutes another source of extrinsic motivation for student learning. Influenced by the traditional 'examination culture,' assessment has been regarded as the 'baton' directing teaching and learning in China (Zhao 1999).

Assessment practices have assumed an increasingly important role in mathematics education across nations. There is a need for educators to realize and exploit the significant role that assessment plays on mathematics teaching and learning

(Clarke 1996; Clarke and Stephens 1996; Dweck 1986; Zhao 2000). Leder and Forgasz (1992) indicate that the link between student learning of mathematics, the teaching methods adopted in the mathematics classroom, and the assessment procedures used to determine and measure student knowledge is addressed by the research community and closely related to the nature of assessment. There is a clear acceptance of the close interdependence of the three areas. They argue that by limiting the domain of mathematics sampled, assessment has often been a restrictive influence. Broadening the areas and skills evaluated, and reinterpreting assessment as a way of communicating students' current understandings to teachers, can lead to positive rather than destructive reinterpretations of mathematics learning.

A number of researchers (Black and Wiliam 1998; Bright and Joyner 1998; Leder and Forgasz 1992; Lokan and Doig 1997) have reported that developing pedagogical knowledge through assessment of students' mathematical learning has a substantial impact on teachers' instruction and hence on students' learning. According to Clarke (1996) and Little and Wolf (1996), the emphasis on assessment is now evident worldwide, both in developing and in developed countries. It has been commonly accepted that reform in assessment is a necessary condition for any reform of the mathematics curriculum.

Crooks (1988) indicates that regular classroom testing does increase these aspects of attainment in mathematics (p. 20). According to Peterson, Fennema, and Carpenter (1989), the more teachers knew about their pupils' mathematical knowledge, the better the pupils were at word problem solving. Knowledgeable teachers questioned their pupils about problem-solving processes and listened to their responses, while less knowledgeable teachers tended to explain problem-solving processes to pupils or just observed their pupils' solutions. Bright and Joyner (1998) suggest that if the educational system is to help more students achieve at higher levels in mathematics, greater attention must be focused on day-to-day linkages among teaching, learning, and assessing. They explained that student performance is influenced by instruction, and one way that instruction can be improved is by helping teachers develop better information about students' thinking. Daily classroom assessment provides a vehicle for teachers to use to acquire knowledge about student's performance and thinking. With high-quality assessments, teachers can help students develop mathematical power than will be revealed in greater student achievement (pp. 3–4).

Black and Wiliam (1998) suggest that the standards of learning can be elevated by improving the formative assessment process in the classroom. They indicate that firm evidence shows formative assessment as an essential component of classroom work and that its development can raise standards of achievement.

There is little doubt (Barnes et al. 2000; Brown et al. 1989; Clarke 1996; Clarke and Stephens 1996; Hiebert and Carpenter 1992; Webb 1992) that assessment exerts a strong influence on teachers and pupils. Assessment sends clear messages about what is valued and can motivate pupils, although this can have both positive and negative consequences. Bright and Joyner (1998) advocate that high-quality classroom assessment is likely to have a great impact on the ultimate learning of mathematics. If students are given clear learning targets (feedforward) and helpful, timely feedback about the quality of their work, and if as a result they develop solid understanding of the content, they are more likely to be able to use what they have learned in new and challenging situations. They also recommend that along with providing information for students, high-quality classroom assessment should also inform the teacher and students' families. They indicate that:

> It seems reasonable to expect that there will be stronger, ongoing encouragement from home for students to make the needed effort to master mathematical ideas and to remain engaged in mathematics course-taking if students know what the learning targets are and have an understanding of what quality performance looks like, if teachers better understand the strengths of their students, and if teachers can share this information with students' family. (p. 4)

In summary, the current discussion suggests that mathematics assessment has powerful influence on mathematics teaching and learning throughout the world, and especially in East Asian countries. To understand the 'achievement gap' between East Asian and Western countries, the influence of assessment on achievement must be the focus of investigation.

2.6 A Tentative Model for the Study

According to Bishop (1988), mathematics education must be recognized as being a social process and there are five significant levels involved:

- Cultural
- Societal
- Institutional
- Pedagogical
- Individual

Although Bishop's work did not focus on comparing mathematics achievement between East Asian and Western countries, he suggests that student's mathematics achievement is influenced by cultural and social factors at five levels. First, at the cultural level, ideological factors such as beliefs about mathematics have been a driving force in the development of mathematics and it has shaped the character of mathematics education. He asserts that mathematics is a cultural phenomenon and there is a need to analyze mathematics education from a cultural perspective. Second, at the societal level, mathematics is mediated by the various institutions in society and is subject to the political and ideological forces in that society which use various formal and informal educational institutions to shape mathematics education to meet societal goals and aspirations. Third, at the institutional level, the intended curriculum which includes the mechanisms for grouping pupils, for examining them, for resourcing the subjects, and for teaching them have profound effects on students' mathematical education. Fourth, at the pedagogical level, the students' mathematical education is influenced by the teacher knowledge and approach and the classroom group. Fifth, at the individual level, students bring personal dimensions to mathematics education derived from the influence of their family, from their traditions and history and the culture of their home and community.

Influenced by Bishop's five levels of social aspects of mathematics education, Leung (2002) suggests that in order to explain differences in mathematics education between the East Asia and the West, variables at five levels need to be considered:

- Student's level (attitudes and self-concept)
- Teacher's level (teaching style, teacher's mathematics competence)
- Education system level (curriculum policy, practice and assessment, instructional time and pacing, of instruction and time out of school doing mathematics and tutoring)
- Societal level (GNP per capita, education expenditure, and school and home resources)
- Cultural level (beliefs, traditions, attitudes, and aspirations).

Like Bishop, Leung argues that there are different cultural values pertinent to education that may explain the difference in mathematics education between East Asian and Western countries.

This study adopts a tentative model (Fig. 2.1) identifying main factors proposed as influencing students' achievement into three levels:

- A base level—a cultural context including cultural values about education and mathematics achievement.
- The second level—standards and content of curriculum, implemented assessment, and teachers' knowledge.
- The third level—classroom teaching methods and out-of-school curricula

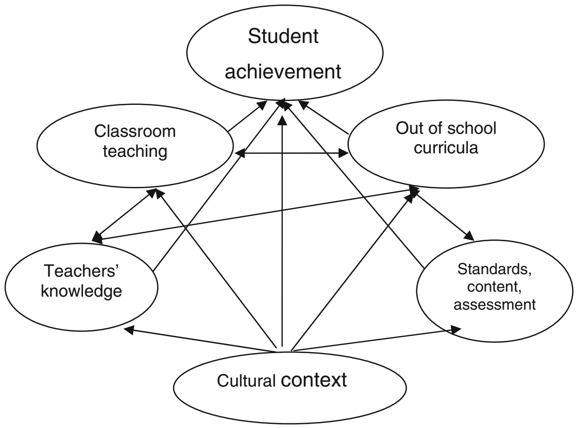

Fig. 2.1 Tentative model of key factors contributing to 'achievement gap'

These factors are reviewed in terms of the three levels and the core influence of cultural context through analyses of case studies. This model is served as a framework for the study and a discussion of factors contributing to mathematics achievement is presented in Chap. 10.

In summary, the literature reviewed in this chapter covers a wide range of studies concerned with comparative studies on mathematics achievement. It is suggested that due to limitations of large-scale assessment, there is a need to use qualitative study to explore curriculum and assessment practices in authentic settings cross-culturally. This review has provided a background and framework for designing a comparative study on mathematics curriculum and assessment policy and practice between Australian and Chinese primary schools, which will be described in the following chapters.

References

Arnove, R. F. (2001). Comparative and international education society facing the twenty-first century: Challenges and contributions. (Presidential Address). *Comparative Education Review, 45*(4), 477–503.

Australian Association of Mathematics Teachers (AAMT). (1997). *Numeracy = everyone's business: The report of the numeracy education strategy development conference May 1997*. Adelaide, Australia: AAMT.

Barnes, M., Clarke, D., & Stephens, M. (2000). Assessment: the engine of systemic curriculum reform? *Journal of Curriculum Studies, 32*(5), 623–650.

Beaton, A. E., Mullis, V. S., Martin, M. O., Gonzales, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's TIMSS*. Massachusetts: Boston College.

Bell, G., & Kang, O. (1995). Intended curriculum and the mathematical attainments of year 6 students in Australia and Korea. *Curriculum Perspectives, 15*(3), 43–52.

Bishop, A. J. (1988). *Mathematical enculturation: A cultural perspective on mathematics education*. Dordrecht, Netherland: Kluwer Academic Publishers.

Bishop, A. J. (1996). Introduction. In A. J. Bishop, et al. (Eds.), *International handbook of mathematics education* (pp. 1–4). Dordrecht, Notherland: Kluwer Academic Publishers.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Bracey, G. W. (1993). American students hold their own. *Educational Leadership, 50*(5), 66–69.

Bracey, G. W. (1996). International comparisons and the condition of American education. *Educational Research, 25*(1), 5–11.

Bracey, G. W. (1998). Are US students behind? *The American Prospect, 9*(37), 64–70.

Bracey, G. W. (2000). 'Diverging' American and Japanese science scores. *Phi Delta Kappan, 81*(10), 791–792.

Bright, G., & Joyner, J. (Eds.). (1998). *Classroom assessment in mathematics: Views from a national science foundation working conference*. Lanham, MD: University Press of America.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Burstein, L. (Ed.). (1992). *The IEA study of mathematics III: Pupil growth and classroom processes*. Oxford, New York: Pergamon.

Cai, J. F. (2000). Mathematical thinking involved in US and Chinese students' solving process-constrained and process-open problems. *Mathematics Thinking and Learning, 2*(4), 309–340.

Campbell, R. J., & Kyriakides, L. (2000). The national curriculum and standards in primary schools: A comparative perspective. *Comparative Education, 36*(4), 383–395.

Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1990). Mathematics in the streets and in schools. In V. Lee (Ed.), *Children's learning in school* (pp. 91–102). London, England: Hodder and Stoughton in association with the Open University.

Chinese Ministry of Education. (2000). *National mathematics curriculum standards for nine-year compulsory education*. Beijing, China: Beijing Normal University Press.

Clarke, D. (1996). Assessment. In A. J. Bishop, et al. (Eds.), *International handbook of mathematics education* (pp. 327–370). Dordrecht: Kluwer Academic Publishers.

Clarke, D. (2000). Assessment for teaching and learning. *Learning Matters, 5*(3), 2–5.

Clarke, D., & Stephens, W. M. (1996). The ripple effect; the instructional impact of systemic introduction of performance assessment in mathematics. In M. Birenbaum & F. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 63–69). Dordrecht: Kluwer Academic Publishers.

Clements, M. A., Bishop, A., Keitel, C., Kiparick, J., & Leung, F. K. S. (2013). *Third international handbook of mathematics education*. New York, N.Y.: Springer.

Cobb, P., Yackel, E., & Wood, T. (1992). Interaction and learning in mathematics classroom situations. *Educational Studies in Mathematics, 23*(1), 99–122.

Cohen, D. (1998). World league tables: what's the score? *Principal Matters, 10*(1), 3–7.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438–481.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist (Special Issue: Psychological Science and Education), 41*(10), 1041–1048.

Eckstein, M. A., & Noah, H. J. (Eds.). (1992). *Examinations: Comparative and international studies*. Oxford, New York: Pergamon Press.

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

Gautherin, J. (1993). Marc-Antoine Jullien (1775–1848). *Prospects: The Quarterly Review of Comparative Education*, *23*(3/4), 757–773.

Geary, D. C., Hamson, G. O., Chen, G. P., Liu, F., & Hoard, M. K. (1998). In S. G. Paris & H. M. Wellman (Eds.), *Global prospects for education: development, culture and schooling* (pp. 31–64). Washington, DC: American Psychological Association.

Genova, T. (2015). Torsten Husen—a co-founder and chairman of IEA from 1962 to 1978. In N. Popov (Ed.), *Quality, social justice and accountability in education worldwide* (pp. 34–40). Sofia, Bulgaria: Bulgarian Comparative Education Society.

Graf, K., & Leung, F. K. S. (2000). The thirteenth ICMI Study on mathematics education in different cultural traditions. *Educational Studies in Mathematics, 43*(1), 95–116.

Halls, W. D. (Ed.). (1990). *Comparative education: Contemporary issues and trends*. London, England: Jessila Kingsley Publishers.

Hans, N. (2012). *Comparative education: A study of educational factors and traditions*. New York, NY: Routledge.

Hayes, W. (1991). *IEA guidebook 1991: Activities, institutions, and people*. The Hague, Netherlands: IEA.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65–97). New York: Macmillan.

Howson, A. G. (1998). The value of comparative studies. In G. Kaiser, E. Luna & I. Huntley (Eds.), *International comparisons in mathematics education,* (pp. 165–188). London: Falmer Press.

Howson, G., & Wilson, B. (Eds.). (1986). *School mathematics in the 1990s*. Cambridge, UK: Cambridge university Press.

Husen, T. (1967a). *International study of educational achievement in mathematics: A comparison of twelve countries* (Vol. I). Stockholm, Hamburg: Almquist & Wisksell.

Husen, T. (1967b). *International study of educational achievement in mathematics: A comparison of twelve countries* (Vol. II). Stockholm, Hamburg, German: Almquist & Wisksell.

Husen, T. & Postlethwaite, T. N. (Eds.), (1994). *The international encyclopedia of education* (2nd, ed, v2). London: Pergamon Press.

Jacobsen, E. (1996). International co-operation in mathematics education. In A. J. Bishop, et al. (Eds.), *International handbook of mathematics education* (pp. 235–1256). Dordrecht, Netherland: Kluwer Academic Publishers.

Jakwerth, P., Bianchi, L., Houang, R., Schmidt, W., Valverde, G., Wolfe, R., et al. (1997). *Validity in cross-national assessments: Pitfalls and possibilities*. Paper presented at the 1997 American Educational Research Association Annual Meeting, Chicago, Illinois.

Lapointe, A. E., Mead, N. A., & Askew, J. M. (1992). *Learning mathematics*. Princeton, NJ: Educational Testing Service.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Leder, G. C., & Forgasz, H. J. (Eds.). (1992). *Assessment and learning of mathematics*. Melbourne, Australia: Australia Council for Educational Research.

Leung, F. K. S. (1995). The mathematics classroom in Beijing, Hong Kong and London. *Educational Studies in Mathematics, 29*(4), 297–325.

Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics, 47*(1), 35–51.

Leung, F. K. S. (2002). Mathematics education in East Asia and West: Does culture matter? *A Presentation Prepared for the 13th ICMI Comparative Study Conference*. Hong Kong: The University of Hong Kong.

Lin, F. L. (2001). The pressure for high achievement in Taiwanese education. *Dialogues, 2001* (11). http://www.nctm.org/dialogues/2001/20022213.htm. Accessed on December 16, 2001.

Lindquist, M. (2001). NAEP, TIMSS, and PSSM: Entangled influences. *School Science and Mathematics, 101*(6), 289–291.

Little, A., & Wolf, A. (Eds.). (1996). *Assessment in transition: Learning, monitoring and selection: An international perspectives*. Oxford, New York: Pergamon.

Lokan, J. (2000). Messages for mathematics education from TIMSS in Australia. In D. S. Taylor & E. W. Jenkins (Eds.), *Learning from others* (pp. 259–277). Dordrecht: Kluwer Academic Publishers.

Lokan, J., & Doig, B. (Eds.). (1997). *Learning from children: Mathematics from a classroom perspective*. Melbourne, Australia: Australian Council for Educational Research.

Lokan, J., Ford, P., & Greenwood, L. (1997). *Maths and science on the line: Australian middle primary students' performance in the third international mathematics and science study*. Melbourne, Australia: Australia Council for Educational Research.

Lokan, J., Greenwood, L., & Cresswell, J. (2001). *15-up and counting, reading, writing and reasoning: How literate are Australia's students?*. Melbourne, Australia: Australia Council for Educational Research Press.

Ma, L. P. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, N.J.: Lawrence Erlbaum Associates.

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

Martin, M., & Mullis, I. (2000). International comparisons of student achievement: Perspectives from the TIMSS international study centre. In D. S. Taylor & E. D. Jenkins (Eds.), *Learning from others* (pp. 29–47). Dordrecht, Netherland: Kluwer Academic Publishers.

McAdams, R. P. (Ed.). (1993). *Lessons from abroad: How other countries educate their children*. Lancater, PA: Techonomic Publishing Company Inc.

Mcleod, D. B., Stake, R. E., Schappelle, B. P., Mellissinos, M., & Gierl, M. J. (1996). Setting the standards: NCTM's role in the reform of mathematics education. In S. Raizen & T. Britton

(Eds.), *Bold ventures: Case studies of US innovations in mathematics education* (pp. 13–133). Dordrecht, The Netherlands: Kluwer.
Mullis, I., Martin, M., Beaton, A., Gonzalez, E., Kelly, D. & Smith, T. (1997). *Mathematics achievement in the primary school years: IEA's third international mathematics and science study*. Massachusetts: Centre for the Study of Testing, Evaluation and Educational Policy, Boston College.
Mullis, I., Martin, M., & Foy, P. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's international mathematics and science study at the fourth and eighth grades.* Massachusetts: Centre for the study of Testing, Evaluation and Education Policy, Boston College.
Mullis, I., Martin, M., Gonzalez, E., Garden, R., O'Connor, K., Chrostowski, S. et al. (2000). *TIMSS 1999: International mathematics reports, finding from IEA's repeat of third international mathematics and science study at the eighth grade.* Massachusetts: Centre for the Study of Testing, Evaluation and Educational Policy, Boston College.
National Council of Teachers of Mathematics. (1995). *The assessment standards for school mathematics*. Reston, Australai: NCTM.
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, Australia: NCTM.
Ng, S. F. (2001). The Singapore story: A nation's effort to develop human resources. http://www.nctm.org/dialogues/2001/20022202.htm. Accessed on December 16, 2001.
Organization for Economic Co-operation and Development (OECD). (2000). *Knowledge and skills for life: First results for PISA 2000.* Paris: OECD Publishing.
Organization for Economic Co-operation and Development (OECD). (2014). *PISA 2012 results: What students know and can do-students' performance in mathematics, reading and science* (Vol. 1, Revised). Paris: OECD Publishing.
Owens, K., & Perry, B. (2001). *Mathematics K-10 literature review*. Sydney, Australia: NSW Board of Studies.
Paris, S. G., & Wellman, H. M. (Eds.). (1998). *Global prospects for education: Development, culture, and schooling*. Washington DC: American Psychological Association.
Peterson, P. L., Fenema, E., & Carpenter, T. (1989). Using knowledge about how students think about mathematics. *Educational Leadership, 46*(4), 42–46.
Reynolds, D., Creemers, B., Stringfield, S., & Teddlie, C. (2000). World class schools: Some preliminary methodological finding from the international school effectiveness research project. In D. S. Taylor & E. W. Jenkins (Eds.), *Learning from others: International comparison in education* (pp. 115–136). Dordrecht, Notherland: Kluwer Academic Publishers.
Robitaille, D. F., & Garden, R. A. (1989). *The IEA study of mathematics II: Contexts and outcomes of school mathematics*. Oxford, New York: Pergamon.
Robitaille, D. F., & Garden, R. A. (Eds.). (1996). *TIMSS monograph no. 2: Research questions and study design*. Vancouver, Canada: Pacific Educational Press.
Robitaille, D. F., & Nicol, C. (1994). The development of comparative studies in mathematics education. In R. Biehler, R. W. Scholz, R. Strabaer, & B. Winkelmann (Eds.), *Didactics of mathematics as scientific discipline* (pp. 403–414). Dordrecht, Netherland: Kluwer Academic Publisher.
Robitaille, D. F., & Travers, K. (1992). International studies in mathematics achievement. In D.A. Grouws (Ed.), *Handbook for research on the teaching and learning of mathematics* (pp. 687–709). New York: MacMilian.
Schleicher, A. (2000). Monitoring student knowledge and skills: The OECD programme for international student assessment. In D. S. Taylor & E. D. Jenkins (Eds.), *Learning from others* (pp. 63–77). Dordrecht, Netherland: Kluwer Academic Publishers.
Schmidt, H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of US. Science and mathematics education*. Dordrecht, Netherland: Kluwer Academic.
Schumer, G. (1999). Mathematics education in Japan. *Journal of Curriculum Studies, 31*(4), 399–427.

Sherman, M. J. (2000). Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States. (Book reviews). *American Scientist, 88*(1), 86–87.

Shimizu, Y. (2001). Why the mathematics performance of Japanese students is higher than that of students in Western countries: Listening to the voices from the inside. http://www.nctm.org/dialogues/2001/20022204.htm. Accessed on December 16, 2001.

Stevenson, H. W., & Nerison-Low, R. (2002). *To sum it up: Case studies of education in Germany, Japan and the United States*. Washington, DC: US Department of Education.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Summit Books.

Stigler, J. W., & Baranes, R. (1988). Culture and mathematics learning. In E. Rothkopf (Ed.), *Review of research in education* (pp. 253–306). Washington, DC: American Research Association.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.

Stigler, J. W., & Perry, M. (1988). Mathematics learning in Japanese, Chinese, and American classrooms. In G. B. Saxe & M. Gearhart (Eds.), *Children's mathematics* (pp. 27–54). San Francisco: Jossey-Bass.

Taylor, D. S., & Jenkins, E. W. (Eds.). (2000). *Learning from others*. Dordrecht: Kluwer Academic Publishers.

Theisen, G., & Adams, D. (1990). Comparative education research. In R. M. Thomas (Ed.), *International comparative education: practices, issues and prospects* (pp. 277–298). New York: Pergammon.

Thomson, S., De Bortoli, L., & Buckley, S. (2013). *PISA 2012: How Australia measures up*. Victoria, Australia: Australian Council for Educational Research.

Travers, K. J., & Westbury, I. (Eds.). (1989). *The IEA study of mathematics: Analysis of mathematics curricula*. Oxford, New York: Pergamon Press.

Ueno, K. (2002). From Wasan to Yozan (Western Mathematics). *A Presentation Prepared for the 13th ICMI Comparative Study Conference*. Hong Kong: University of Hong Kong.

Uy, F. L. (2001). The Asian advantage: Fact, fiction, fear or fantasy? *Mathematics Education Dialogues, November*. http://www.nctm.org/dialogues/2001-11/20001112.htm. Accessed on December 16, 2001.

Wang, J. (1998). A content examination of the TIMSS items. *Phi Delta Kappan, 80*(1), 36–38.

Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 661–683). New York: Macmillan.

Whang, W. H. (2001). Speculating on the high achievement of Korean students. http://www.nctm.org/dialogues/2001/20022214.htm. Accessed on December 16, 2001.

Zhao, D. (1999). For the students' future: A review of the reform on high school assessment in New South Wales. *Studies in Foreign Education, 26*(4), 18–22.

Zhao, D. (2000). Establishing an evaluation system suitable for quality education: A review of the evaluation system of primary school pupils in Australia. *Studies in Foreign Education, 27*(1), 39–43.

http://www.springer.com/978-981-10-0283-0

Chinese Students' Higher Achievement in Mathematics
Comparison of Mathematics Education of Australian
and Chinese Primary Schools
Zhao, D.
2016, XVII, 307 p. 15 illus., Hardcover
ISBN: 978-981-10-0283-0