

A Conceptualization of Indicators for Mathematics Teacher Education Quality for International Studies

Feng-Jui Hsieh, Chiu-Keung Law, Haw-Yaw Shy, Ting-Ying Wang,
Chia-Jui Hsieh, and Shu-Jyh Tang

Abstract The Teacher Education and Development Study in Mathematics, sponsored by the International Association for the Evaluation of Educational Achievement, is the first data-based study about mathematics teacher education with large-scale samples; this article is based on its data but develops a stand-alone conceptual framework to investigate the quality of teacher education among various countries. The framework includes five indicators: future teacher achievement, instructor effectiveness, coherence between universities and schools, courses/content arrangement, and overall effectiveness of teacher education programs. One of the findings provides indications that it is necessary to combine theoretical knowledge with practical teaching into teacher education; another finding is that for all countries involved, future teachers are less approving of the courses/content arrangement of teacher education programs than are program educators, thus perhaps lowering educators' motivation to improve the arrangement. The data also indicate that there is a high degree of synchronization and organization in teacher education programs in the United States; however, these programs still require further development and promotion of their future teachers' knowledge achievements.

Keywords Teacher education · Teacher development · Mathematics education · International education · International study · School effectiveness · Teacher effectiveness

This paper is an adjusted version of Hsieh et al. (2011).

F.-J. Hsieh (✉) · T.-Y. Wang · C.-J. Hsieh · S.-J. Tang
National Taiwan Normal University, Taipei, Taiwan
e-mail: hsieh fj@math.ntnu.edu.tw

C.-K. Law
National Hsinchu University of Education, Hsinchu, Taiwan

H.-Y. Shy
National Changhua University of Education, Changhua, Taiwan

1 Introduction

The world is flat! Countries around the world affect one another, no matter their economic structures, thoughts, beliefs, or values. Therefore, we know that no country or person can be independent. Acknowledging the situations and trends of other countries is thus essential. According to this view, the international Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) enable different nations to understand one another's educational situations.

Results from the TIMSS and PISA studies reveal that there are significant differences among countries in mathematics competency (National Center for Education Statistics [NCES] 2009; Organisation for Economic Co-operation and Development [OECD] 2007). These include mathematics knowledge and its application, analysis and problem solving, and model utilization and augmentation, among other proficiencies (OECD 2007). In addition, these studies indicate that different countries have different curricula, a fact that might constitute the reason for various students' significant levels of difference. However, these indications are the total of our current interpretations; we still do not know how and in what way this difference of curricula affects our students' vast differences in knowledge.

To understand the factors that can affect students' levels of achievements, the TIMSS 1999 video study (NCES 2003) focused on three aspects of mathematics teaching: the way lessons are organized, the nature of content implemented in lessons, and instructional practices. The study found that there were detectable differences in the relative emphasis or arrangement by mathematics teachers in different countries. It further suggested that teaching methods should align with what teachers want their students to learn and that one cannot say which teaching method may be best to implement in a given country.

In these three aspects, a common latent factor is noticed—the quality of teachers. Many studies have shown that teacher quality is the most important school-related factor influencing student achievement (Goe 2007; Kaplan and Owings 2001; Rice 2003). Some also have found that the methods and content used by teachers have a definite influence on their students' learning (Abell Foundation 2001; Fetler 1999; Goldhaber and Brewer 2000).

Various nations have therefore established teacher certification to control the quality of teachers (e.g., NCES 1999; see also Goldhaber and Anthony 2004). Different certifications thus exist in order to assess candidates' different types of knowledge for mathematics teaching. Among these different types of knowledge are subject matter knowledge, subject specific knowledge for teaching, and pedagogical knowledge, all of which enjoy considerable favor in modern certifications, such as the Praxis Series (Hill et al. 2007; NCES 1999). These types of knowledge have also gained attention in academic circles (Hill et al. 2007). Many researchers further claim that a given teacher's knowledge of mathematics and knowledge of how to translate mathematics into a form that can be understood by students play the most important role in effective teaching (Ferrini-Mundy et al. 2005). These two types of

knowledge are indeed consistent with, if not identical to, the two of Shulman's categories for teachers' knowledge that are applicable to mathematics, namely, mathematics content knowledge and mathematics pedagogical content knowledge (Shulman 1987).

How a country can guarantee that its teacher quality is high has been a seriously considered issue. One straightforward inference is to guarantee the quality of the basic learning environment in which we train and equip our future teachers, that is, teacher education programs (TEPs). The different features and practices involved in TEPs are therefore worth investigating. A recent study conducted by the National Research Council (2010) in the United States addressed various issues about teacher preparation, including faculty and staff qualifications, the requirements for subject matter knowledge, general pedagogy and professional knowledge, and field experience. However, this study provided results only from the United States. Therefore, we realized that it is vitally important to globalize the study of the knowledge of future teachers and the features as well as the practices of TEPs and to be able to compare these results among various countries.

To reflect on the demands of globalization, the international Teacher Education and Development Study in Mathematics (TEDS-M) was launched to study and compare the policies, practices, and outcomes of teacher preparation programs among different countries. TEDS-M, sponsored by the International Association for the Evaluation of Educational Achievement (IEA), is the first cross-national study about mathematics teacher education with large-scale samples. The TEDS-M study team developed a thorough analytical framework and completed a process of data collection (Tatto et al. 2008). An important issue for TEDS-M was to describe and compare teacher education quality among diverse countries. However, the TEDS-M data analysis is still in its initial stage; therefore, the available resources are limited in scope. This article consequently is based on a stand-alone study that we conducted and uses the data collected by TEDS-M, while also referring to the earlier results of Taiwan's TEDS-M national report (Hsieh et al. 2010). The main purpose of this study is to depict the phenomena, patterns, and comparisons of the participating countries' TEPs in terms of effectiveness.

2 Framework

Based on the purpose of this study, we face the following problem: What features of TEPs can be treated as indicators of effectiveness?

Darling-Hammond (2000) claims that policies regarding teacher education, licensing, hiring, and professional development might make an important difference in the qualifications and capacities that teachers bring to their work. Wang et al. (2010) agree with this point and propose that teacher education should prepare and retain sufficient numbers of high-quality teachers who can work effectively with students in order to establish a credible public image of what they do. Many researchers have attempted to figure out how best to evaluate teacher effectiveness

and which criteria should be included in such an evaluation. Among all types of teacher knowledge, there are essentially two types of mathematics teacher knowledge: mathematics content knowledge (MCK) and mathematics pedagogical content knowledge (MPCK). Some researchers have noted that MCK is necessary for mathematics teachers to be effective (Allen 2003; Goldhaber and Brewer 1997), whereas others have posited that MPCK is an important element to effectiveness (Hill et al. 2007; Ingvarson et al. 2007; Shulman 1987). Accordingly, as both MCK and MPCK are regarded as essential ingredients in *future teacher achievement*, both types of knowledge make up the first indicator of the effectiveness of TEPs in our conceptual framework. This indicator concerns the issue of outcome, an essential part of the effectiveness of TEPs, particularly relating to the quality of people being cultivated.

Many studies have shown that there is a strong correlation between students' achievement and the quality of their instructors' teaching (Ferguson 1998; Goe 2007; Kaplan and Owings 2001; Rice 2003). Others have paid particular attention to levels of quality regarding how well instructors can teach (Clark 1992; Ducharme and Ducharme 1999; Howey 1995). From these points, there is no doubt that the quality of instructors' teaching is an important factor in determining the quality of the TEP. Regarding mathematics TEPs, instructors in mathematics-related courses (MR-instructors), who play a crucial role in helping future teachers learn to teach mathematics (Tatto et al. 2008), and school-based supervising teachers (SB-supervisors), who have the important responsibility of mentoring future teachers' learning during field-based experiences (Putnam and Borko 2000), cannot be ignored when evaluating the effectiveness of TEPs. Therefore, the second indicator of the effectiveness of TEPs in our conceptual framework is the *effectiveness of instructors*, composed of the effectiveness of both MR-instructors and SB-supervisors.

On the one hand, MR-instructors provide future teachers with theoretical concepts about teaching ideas, principles, and standards as well as demonstrating models, evaluations, and reflections in college. On the other hand, SB-supervisors offer practical knowledge in field experiences and teaching methods and an understanding of pupils at school sites. Thus, future teachers should apply these teaching theories in real classroom teaching and advance their field experiences in developing knowledge (Bates et al. 2009; Wang et al. 2010; Zeichner 2010). For a TEP to be effective, it is mandatory to evaluate whether or not those ideas and principles taught in colleges, or the standards provided by them, are coherent with the experiences needed in schools. The *teaching coherence* between teacher education universities and schools is therefore integrated into the framework of this study as the third indicator of the effectiveness of TEPs.

Although the coherence of teaching between a university and a school is important and contributes to successful teaching for future teachers, course arrangement is also an important characteristic of TEPs (Tatto et al. 2008) and one that is usually expected to meet the main needs as an effective teacher (Florida State Department of Education 1983). In light of this, the effectiveness of *courses/content arrangement* in TEPs is considered and incorporated into the framework of this study as the fourth indicator of the effectiveness of TEPs. In this article, future teachers are

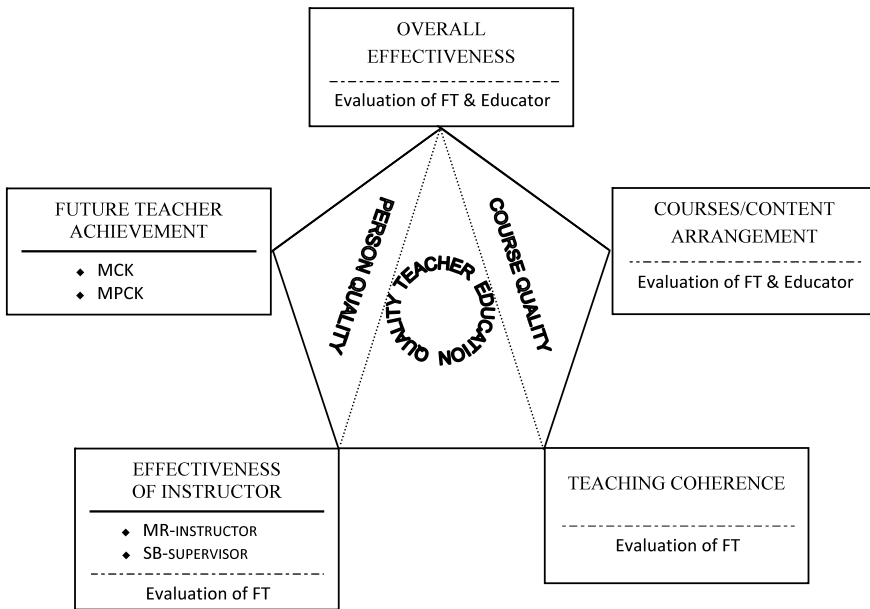


Fig. 1 Conceptual framework utilized in this study. *FT* = future teacher; *MCK* = mathematics content knowledge; *MPCK* = mathematics pedagogical content knowledge; *MR-instructor* = instructor in mathematics-related courses; *SB-supervisor* = school-based supervising teacher

assumed to provide the pragmatic view, and the educators, as the planners and executors, are assumed to represent the advanced view. Their evaluations together may depict the effectiveness of the courses/content arrangement.

Each of these four indicators, while referring to some part of the quality of teacher education, nevertheless paints an incomplete picture. This study fills this gap with a final indicator—the *overall effectiveness* of TEPs, the inclusive nature of which lends itself to evaluation by both future teachers and educators. Together, these five indicators depict the effectiveness of TEPs from different perspectives: the one being educated, the educator, and the circumstances under which this education takes place. This article is based on the proposed framework shown in Fig. 1 and seeks to investigate teacher education quality among various countries from an effectiveness point of view. This framework consists of five indicators that fall into two major categories: person quality and course quality.

Based on this conceptual framework, we therefore addressed three research questions to guide the analysis and discussion of this study:

1. What are the phenomena or patterns regarding effectiveness for each of the five indicators among the participating countries?
2. What are the levels of effectiveness for each of the five indicators for each country?
3. What are the correlations among these five indicators and the possible concomitant interpretations?

3 Research Method

The target populations in this study included future primary and lower secondary teachers in their last year of training to teach mathematics and teacher educators who instructed these future teachers in the fields of (a) mathematics and mathematics pedagogy and (b) general pedagogy. All persons with regular and repeated responsibilities in teaching future primary and lower secondary mathematics teachers were classified as teacher educators in this study.

The sampling plan contained in this study followed a stratified multistage probability sampling design (Tatto et al. 2009), and samples of teacher preparation institutions were randomly selected with probability proportional to size within explicit strata according to the specific context of each country. For each selected teacher preparation institution, individuals including educators and future teachers were randomly selected or a census was used. The sampling designs and processes for all the countries were developed in consultation with IEA sampling referees and the regulations of the IEA-developed sampling guide.

The future teacher samples of this study either met the IEA's threshold (at least 75 %) or met the criterion to use with an annotation (60 %–75 %; for a more detailed description see Chap. "Framing the Enterprise: Benefits and Challenges of International Studies on Teacher Knowledge and Teacher Beliefs—Modeling Missing Links" of this book). For the educator samples, this study used a threshold rate of 50 %. This rate takes into account previous research that shows that it is difficult to get satisfactory response rates when surveying adults, and it was chosen to ensure the inclusion of more information. Some studies have accepted participation rates much lower than 50 % for adult samples (e.g., Archambault and Crippen 2009; Enochsson 2010). IEA advises that a sample with a participation rate of 30 %–60 % is to be reported separately. However, our choice of a rate above or equal to 50 % is not far from 60 % and also equates to more than half of the sample. All participation rates were calculated and are reported in Table 1.

The international TEDS-M data set did not distinguish teacher educators by levels. The entire group is thus designated as *all educators* in this article. To match the levels with future teachers, these educators were further recategorized by this study into primary level or lower secondary level based on the levels of the teacher preparation units in which these educators served. If an educator served at both levels, he or she would be counted in each level.¹ For the purpose of this article, a distinction is made between *teacher* and *educator*, where *educator* refers to an educator of future teachers. The groups of educators that serve in preparing lower secondary and primary level future teachers are named *lower secondary educators* and *primary educators*, respectively, in this article.²

¹Levels are indistinguishable from Germany's data. Thus, when comparing data of educators with the levels distinguished, the data of all educators are used for Germany.

²This study uses the newest release of data sets TEDS_MS_NRC-USE_IDB_20091209_v30 for national research coordinators from the international Teacher Education and Development Study in Mathematics.

Table 1 Participation numbers and participation rates of each level

Country	Educator		Lower secondary future teacher		Primary future teacher	
	<i>n</i>	PR	<i>n</i>	PR	<i>n</i>	PR
Botswana	43	x	53	x	86	x
Chile	392	50–60	746	60–75	657	60–75
Georgia	62	x	78	60–75	506	x
Germany	482	50–60	771	x	1032	x
Malaysia	255	50–60	389	60–75	576	x
Norway ^a			550		551	
Oman	84	x	268	x	–	–
Philippines	589	x	733	x	592	x
Poland	734	60–75	298	60–75	2112	60–75
Russia	1212	x	2141	x	2266	x
Singapore	77	x	393	x	380	x
Spain	533	x	–	–	1093	x
Switzerland	220	50–60	141	x	936	x
Taiwan	195	x	365	x	923	x
Thailand	312	x	652	x	660	x
US-Public ^b			607	60–75	1501	x
Total	5190		8185		13871	

Note. *n* = number of unweighted participants; PR = range of the participation rate. An “x” indicates the participation rates that meet International Association for the Evaluation of Educational Achievement’s threshold. Blanks indicate the data were collected but not yet processed. Dashes indicate the country did not participate in the denoted level

^aPR cannot be confirmed yet for Norway

^bThe data from the United States include only public institutions

For future teachers and educators, TEDS-M developed three instruments: a future primary teacher questionnaire, a future secondary teacher questionnaire, and an educator questionnaire. The future teacher questionnaires included both tests and Likert-type scale items, whereas the educator questionnaire included only the latter. Using a self-report method to study and measure the effectiveness of TEPs has its limitations, as respondents’ self-impressions may be different from reality. Although other methods may overcome some of these limitations, self-report questionnaires are economical and simple to administer to large numbers of respondents, especially for a cross-national study involving different cultures and languages. Moreover, direct evaluation of effectiveness by future teachers constitutes a pragmatic benefit that is similar to customer evaluation. Therefore, we adopt the data obtained by TEDS-M using both testing and a self-report method of data collection in this study.

According to our proposed conceptual framework and research questions, several variables from those instruments were adopted in this study. The variables MCK and MPCK were used as the indicator of future teacher achievement. The rest of the variables all came from Likert-type scale items: Effectiveness of instructors consisted of MR-instructors' and SB-supervisors' effectiveness in teaching, teaching coherence was concerned with the connection between the teaching of universities and the teaching of schools, courses/content arrangement dealt with the consistency of courses and/or content within a university itself, and overall effectiveness treated the TEP as a whole. Both future teachers and educators were involved in the last two indicators. For the indicators measured based on Likert-type scale items, factor analyses were done to put the items together. Except for overall effectiveness, all variables for the indicators in this study were estimated by using the partial-credit Rasch model with a center at the value 10 as an essentially neutral position. In other words, a logit score of 10 represents a neutral rating toward the rated index. According to the attributes of logit scores, a higher score therefore means a higher index. The data collection period for participating countries varied from late 2007 to early 2009.

In addition to the statistics used by TEDS-M, this study further utilized a variety of statistical analyses and statistical procedures, which will be delineated as they become applicable.

4 Research Findings

International comparisons are widely used to indicate the degree of success of a nation's education system and also the levels of performance to which a given country should aspire. To some degree, cross-national comparisons of education can serve as indicators of a country's educational qualities and have thus constituted a powerful impetus for educational reforms. Thus, from an international perspective, we propose to focus our concentration and begin our discussions on the phenomena, patterns, and comparisons of those indicators within the conceptual framework that are relevant to the effectiveness of TEPs.

4.1 Future Teacher Knowledge Achievements

The results of international analyses show that the mean differences between the highest and lowest rated countries were strikingly large in terms of the standard deviation of 100. The least divergence ($SD = 2.48$) appears in MPCK at the primary level, whereas the most ($SD = 3.13$) appears in MCK at the lower secondary level. The dispersions of the means reported here were in comparison with those of the fourth and eighth graders' achievements in TIMSS, and it appeared that the variability of future teachers' knowledge among these countries was bigger than that of

school-level students. This may mean that in these countries the differences in future teachers' MCK scores have a more serious impact than the achievement scores of school students. One point worth mentioning is that the primary-level MCK items included only those at the school level. In this case, the significant differences between countries should be a cause of concern to the teacher education field, as it seems that in some countries primary-level teachers lack some of the basic mathematics knowledge that is commonplace among future primary teachers in other countries. In contrast, the lower-secondary-level mathematics tests included items from the primary level to the college level. Though this may cause a greater difference in achievement among countries, it also demonstrates that some countries emphasize mathematics up to the college level, whereas others do not.

Two more interesting phenomena emerge if we investigate the data further by school levels and knowledge types. The range for MCK was larger than that for MPCK at both school levels, and the lower-secondary-level MCK means were spread out much more widely than those of the primary. It is difficult to reach a sure interpretation concerning these phenomena because each country presents a separate contextual element. However, one possible conclusion is that between the countries there is a greater difference in the emphasis on MCK than there is in the emphasis on MPCK. The wider spread of lower-secondary-level MCK scores may further confirm that the inclusion of tests from the primary level to the college level may yield a big achievement difference between countries with a narrow range in mathematics and those with a wide range including college mathematics.

In terms of all countries' means, the ranks varied case by case, with some relatively more stable than others. By taking a look at only six countries that have achieved levels beyond the international mean of 500 on all four measures, primary and lower secondary MCK and primary and lower secondary MPCK, we found that Singapore and Taiwan ranked consistently within the top-three highest achieving countries, and Germany and the United States almost always remained in the middle, with means a little higher than the international mean (see Fig. 2).³ Figure 2 also shows that the Russian Federation exhibited a trend of means similar to but lower than that of Taiwan, whereas Singapore exhibited a trend of means similar to but higher than that of the United States, and Switzerland had a trend of means similar to but higher than that of Germany.

Because the future teachers of Singapore, Taiwan, Germany, and the United States scored higher than the international mean in the primary vital indicator of knowledge achievements, they demonstrate an evenly kept balance among different school levels as well as in the types of knowledge in their teacher education policies. They also exemplify both the high-achieving countries, Singapore and Taiwan, and the mid-achieving countries, Germany and the United States, in terms of knowledge. For this reason, we anticipated that it would be informative to explore their strengths and limitations; therefore, these four countries are used as examples,

³None of the countries persistently stayed in the middle range, but Germany and the United States stayed frequently in the middle range, each having three means in between ± 0.25 standard deviations from the international means.

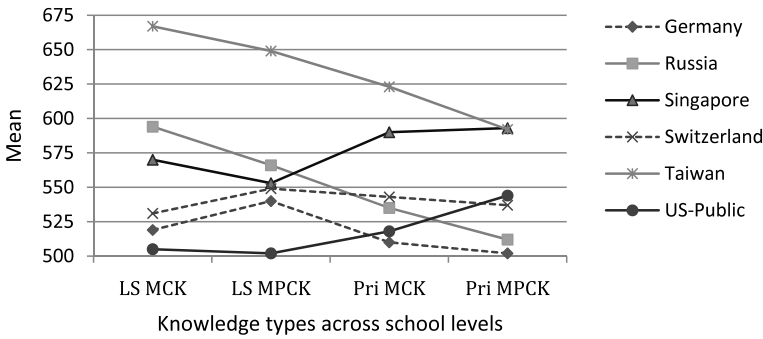


Fig. 2 The within-country trends of knowledge types across school levels for the six countries having all means above the international mean of 500. *LS* = lower secondary; *Pri* = primary; *MCK* = mathematics content knowledge; *MPCK* = mathematics pedagogical content knowledge

and the concomitant analyses use their performances in other indicators whenever possible. In some cases, when the Russian Federation and Switzerland demonstrate unique features, they are also included in discussion.

4.2 Effectiveness of Instructor

As discussed earlier in the conceptual framework, two kinds of instructors were evaluated in this study with respect to their effectiveness: MR-instructors (mathematics-related instructors) and SB-supervisors (school-based supervisors). Based on TEDS-M future teacher questionnaires, the effectiveness of MR-instructors was determined by demonstrating good models in their teaching, evaluations, and reflections; drawing on and using research that is relevant to the content of their courses; and valuing future teachers’ learning and experience. The effectiveness of SB-supervisors was measured by whether their feedback could help future teachers improve their understanding of pupils, curricula, teaching methods, and knowledge of mathematics content. The ratings for both types of these instructors were obtained through a set of Likert-type scale items.

In Table 2, we present the means of MR-instructors’ and SB-supervisors’ scores for each of the participating countries at both the lower secondary and primary levels.

From Table 2, one can notice that all means go beyond the neutral rating of 10, which means that every participating country had positive ratings regarding the effectiveness of their MR-instructors and SB-supervisors. This indicates that future teachers can benefit from both academic and school-based instructors, and it also shows the necessity and appropriateness of integrating theoretical knowledge and practical teaching in teacher education. By going into more detail concerning the ranks of countries across all categories, we see that all Eastern and Southeastern Asian participating countries, other than Malaysia, ranked in the upper half, whereas

Table 2 Country means of future teachers' logit scores of MR-instructor and SB-supervisor

Lower secondary FT				Primary FT			
MR-instructor		SB-supervisor		MR-instructor		SB-supervisor	
Country	M	Country	M	Country	M	Country	M
Philippines	14.00	Philippines	14.49	Philippines	14.17	Philippines	14.71
Thailand	13.81	Taiwan	14.12	Thailand	13.64	Thailand	14.26
Singapore	13.04	Thailand	14.02	US-Public	13.35	Taiwan	13.95
Malaysia	12.82	Singapore	13.77	Russia	13.28	Singapore	13.89
US-Public	12.81	Botswana	13.55	Singapore	13.25	Russia	13.68
Taiwan	12.62	US-Public	13.52	Malaysia	13.21	US-Public	13.46
Russia	12.48	Oman	13.39	Taiwan	12.15	Botswana	13.33
Chile	11.86	Russia	13.05	Botswana	12.10	Malaysia	12.84
Oman	11.85	Malaysia	12.87	Chile	11.99	Poland	12.02
Georgia	11.81	Chile	12.19	Georgia	11.78	Norway	11.99
Botswana	11.79	Norway	12.09	Poland	11.62	Chile	11.88
Poland	11.33	Poland	12.08	Switzerland	11.25	Spain	11.86
Germany	11.19	Switzerland	10.97	Spain	10.83	Georgia	11.42
Switzerland	10.73	Germany	10.68	Norway	10.68	Switzerland	10.80
Norway	10.47	Georgia	10.56	Germany	10.57	Germany	10.54

Note. MR-instructor = instructor in mathematics-related courses; SB-supervisor = school-based supervising teacher

the United States and the Russian Federation were the only other two countries that also ranked in the upper half. Although one cannot say that the MR-instructors and SB-supervisors of those countries ranked in the upper half provided more professional help to facilitate their students in becoming well-trained teachers, we can say that they earned a stronger endorsement from their students, namely, the future teachers.

Based on the homogeneity and heterogeneity of the duties of MR-instructors and SB-supervisors, it would be fascinating to know what these future teachers regard as important in order to be more functionally effective. Thus, we started our inspection by focusing on the mean differences within countries, and an interesting pattern emerged. More than two thirds of the countries gave evidence that the SB-supervisors are more effective than MR-instructors in helping future teachers become well trained.

This study further examined whether the effectiveness of these two groups correlated with each other, and the Spearman's correlation coefficients at the lower secondary and primary levels were $r_s = 0.76$, $p < 0.01$, and $r_s = 0.88$, $p < 0.01$, respectively.⁴ This means that the rankings of the effectiveness of MR-instructors

⁴Spearman's rank correlation analyses were chosen because of the small number of countries sampled in this study.

were highly associated with those of the SB-supervisors. This result indicates why the placements of countries are so consistent in these rankings.

With these strong correlations, we sought to investigate whether the effectiveness of these two groups of instructors influenced future teachers' MCK and MPCK achievements. We employed Pearson's correlation analyses to examine each country's situation and showed that of these countries the results were either without significance or with small significant coefficients ($-0.2 < r < 0.2$), with the exception of Georgia.

These findings revealed that the effectiveness of MR-instructors and SB-supervisors did not have a noticeable influence on future teachers' knowledge achievement in any country except Georgia. The future teachers' high ratings for effectiveness of instructors thus did not guarantee high knowledge achievement and vice versa. For example, Germany generally placed midrank in achievement scores, but the effectiveness of Germany's instructors is ranked near the bottom. Operating on a premise that students learn from their instructors, why are there no significant correlations between the future teachers' knowledge and the effectiveness of their instructors? By reviewing the items examining the effectiveness of instructors in the future teacher questionnaires, we noticed that the content of these items is highly related to real teaching instead of knowledge accomplishment. This probably explains why the correlations are not significant.

4.3 Teaching Coherence Between Universities and Schools

As an indicator, teaching coherence reveals the effectiveness of the education future teachers receive at their universities in relation to their future needs as teachers. Being experienced teachers, the SB-supervisors not only play the role of mentors in TEPs but are also in the position of inspecting whether the content or approaches of courses taken by future teachers in their universities are consistent with the needs of teaching in schools. Because both the SB-supervisors and the future teachers possess firsthand observations, experiences, and a sense of the learning consistency, summaries that include a probe of SB-supervisors' views will be more informative. In TEDS-M, the evaluation of teaching coherence was obtained by the future teachers' ratings on five Likert-type four-point scale items. These items took into account the extent to which SB-supervisors appreciated the teaching ideas, approaches, and standards employed in their teacher education universities in terms of applicability to the real classroom settings.

Table 3 shows the means of teaching coherence for each of the participating countries at both the lower secondary and primary levels.

Similar to the effectiveness of instructors, all means go beyond the neutral rating of 10, which indicates that every participating country had positive ratings regarding the coherence of the content taught in their teacher education universities compared to what future teachers should know in real classroom teaching. A noticeable phenomenon emerged: The United States, a mid-achieving country, compiled the

Table 3 Country means of future teachers' logit scores of teaching coherence

Lower secondary FT		Primary FT	
Country	M	Country	M
US-Public	13.00	US-Public	13.13
Philippines	12.91	Russia	13.04
Russia	12.61	Philippines	13.02
Botswana	12.56	Thailand	12.52
Thailand	12.52	Singapore	12.46
Singapore	12.45	Chile	12.18
Chile	12.29	Poland	11.77
Oman	12.09	Switzerland	11.63
Poland	11.81	Botswana	11.51
Malaysia	11.79	Norway	11.41
Switzerland	11.59	Spain	11.36
Norway	11.39	Malaysia	11.35
Taiwan	10.98	Georgia	11.27
Georgia	10.95	Taiwan	11.06
Germany	10.60	Germany	10.80

Note. FT = future teacher

highest average scores in teaching coherence at both school levels, but Taiwan, a high-achieving country, remarkably descended toward a bottom position.

This phenomenon told us that there was most likely no correlation between the indicators of teaching coherence and future teachers' knowledge achievement. Therefore, Pearson's correlation analyses were employed to determine whether these correlations existed. Although some countries' correlation coefficients reached the 0.05 level of statistical significance, all of these coefficients were small ($-0.2 < r < 0.2$). On the other hand, Spearman's rank correlation analyses showed no significant correlations between the countries' means of teaching coherence and MCK or MPCK achievements. These results revealed that the degree to which teaching coherence between universities and schools related to the teaching ideas, principles, and standards was not statistically associated with future teachers' performance on MCK or MPCK.

This result does not seem to be predictable. A common concept is that learning will be motivated and promoted if we can reinforce it. A teacher education system with a high rating of teaching coherence seems to be reinforced at the second learning location: the school. Why is there no statistical relationship between teaching coherence and knowledge achievement? One possible explanation for this is that teaching coherence as employed in this study evaluates the degree of the coherence between universities and schools in the dimension of real teaching, not the dimension of knowledge achievement; therefore, their correlation is not significant.

At the country level, Spearman's rank correlation analyses showed that the ranking of teaching coherence was significantly correlated with the ranking of

MR-instructors' ratings ($r_s = 0.58$ and $r_s = 0.69$, respectively, for the lower secondary and primary level) and the ranking of SB-supervisors' ratings ($r_s = 0.63$ and $r_s = 0.55$, respectively, for the lower secondary and primary level). These analyses further indicate that, for both kinds of instructors, the more effectively the instructors were ranked, the higher the TEPs' coherence was also rated. Although these indicators have significant correlations, country evaluation differences still exist, as exemplified by Singapore, whose future teachers rated teaching coherence around 12.5, about 0.5 to 1.5 logits less than their ratings for effectiveness of instructors. Nevertheless, we still found that some countries' means of teaching coherence were closer to those of effectiveness of instructors, such as the United States.

It is not easy to change a person's attitude toward something in a short period of time, for example, in a few classes. Therefore, if the ratings of teaching coherence and effectiveness of instructors are nearly the same and at a high level, it should be perceived that the teaching ideas, principles, and standards taught by university instructors; the teaching models, evaluations, and reflections demonstrated by MR-instructors; and the field experiences, teaching methods, and understanding of pupils induced by SB-supervisors are tightly integrated. We perceive the teaching of these types of programs as being synchronized. By taking the sum of the rating scores for teaching coherence and effectiveness of instructors, with each of the MR-instructors and SB-supervisors weighted a half, as a score for the degree of synchronization, we found that the United States and Singapore presented the most synchronized programs at the lower secondary level and that the United States and the Russian Federation shared first place at the primary level. For the six countries included here, the United States demonstrated the most synchronized teaching in TEPs at both the lower secondary and the primary levels.⁵

4.4 Courses/Content Arrangement

Not only is the connection of teachings between a university and a school important but also the connection of teachings within a university itself. The evaluation of the courses/content arrangement can be an indicator of the effectiveness of courses and the practicality of the materials being taught. Both future teachers and educators are involved in this indicator. On the one hand, the future teachers, who are those persons being directly exposed to the courses/content arranged by their TEP, can evaluate from a practical standpoint. On the other hand, with higher academic backgrounds and richer experiences involving researching or teaching, educators can represent a more advanced standpoint. In fact, educators usually play the most important role in developing and executing the content or even planning courses for their

⁵For the six countries in concern, at the primary level, the top-three countries, in order, were the United States, Singapore, and the Russian Federation. For the lower secondary level, the United States remained the first, but the second and third countries exchanged positions. For both levels, Taiwan, Switzerland, and Germany were sequentially ranked. The rating scores of the United States and those of the countries ranked third or lower at both levels were significantly different.

Table 4 Country means of future teachers’ and educators’ logit scores of courses/content arrangement and the differences in the means

Lower secondary level				Primary level			
Country	FT	Edor	DM	Country	FT	Edor	DM
	M	M			M	M	
Philippines	13.76	14.88	-1.11**	Philippines	13.98	15.23	-1.24**
US-Public	13.50			Russia	13.50	13.61	-0.12
Russia	12.96	13.13	-0.16	US-Public	13.35		
Thailand	12.94	14.17	-1.23**	Malaysia	13.10	14.03	-0.93**
Botswana	12.70	14.06	-1.35	Thailand	13.06	14.17	-1.11**
Malaysia	12.70	14.24	-1.54**	Singapore	12.68	13.43	-0.75**
Oman	12.28	12.92	-0.64*	Botswana	12.57	13.58	-1.01
Singapore	12.01	13.53	-1.52**	Chile	11.88	13.63	-1.75**
Taiwan	11.96	13.29	-1.34**	Georgia	11.62	14.07	-2.45**
Chile	11.79	13.50	-1.71**	Taiwan	11.47	12.24	-0.77**
Poland	11.43	12.20	-0.77**	Poland	11.26	12.43	-1.18**
Georgia	10.83	14.31	-3.49**	Spain	10.30	11.15	-0.86**
Switzerland	10.42	11.70	-1.28**	Norway	10.22		
Norway	9.96			Switzerland	10.20	11.79	-1.59**
Germany	9.22			Germany	9.07		

Note. FT = future teacher; Edor = educator; DM = difference of means, which were obtained by subtracting future teachers’ mean of logit scores by educators’ for each country. The German mean for all educators, which is not distinguishable by levels, is 11.17. It is much higher than the average of their future teachers’ means, 9.15, of both levels. * $p < 0.05$, ** $p < 0.01$

students. To determine the effectiveness of courses/content arrangement, TEDS-M focused on the organization of the sequences, the links of the courses/content, and whether the courses/content met the needs of future teachers. Sets of six Likert-type scale items were included in both the future teachers’ and educators’ questionnaires in order to obtain their ratings.

Table 4 presents the means of courses/content arrangement of both the future teachers and educators for each of the participating countries at both the lower secondary and primary levels.

Although almost all countries’ future teachers, regardless of their levels, approved of the courses/content arrangement, some were below 10, as shown in Table 4. Because teaching involves the use of the different kinds of knowledge taught in universities and any effective teaching method is subject to different kinds of learners or situations, it therefore requires teachers to incorporate a large block of ideas and skills simultaneously. However, in most TEPs, subject matter knowledge and didactical methods often are separated, letting future teachers integrate related concepts and skills by themselves. This situation made the sequences and links of the courses/content, a part of courses/content arrangement, vitally important. Con-

sequently, if a TEP has a high rating of effectiveness from future teachers for its courses/content arrangement, we call this program's curriculum well organized.

At both of these levels, the Russian Federation and the United States ranked in the top three, meaning that in comparison to other countries, their courses/content arrangements were appropriate and, from the view of future teachers, met their needs. These two countries demonstrate good examples of programs with well-organized curricula. In contrast to the United States, Germany showed a lack of organization in its curriculum. Because the arrangement of courses and teaching content for future teachers should always consider the targeted levels of instruction, we conducted a comparison between the teaching grade spans and specializations among these three countries.

The Russian Federation prepared generalists at the primary level (up to the fourth grade) and specialists in mathematics at the upper primary and lower secondary levels. The United States was similar to the Russian Federation, the only exception being that there was a mix of generalists and specialists at the Grade 4–5 levels. These models of program organization are probably better in terms of the courses and teaching content arrangement. Germany, on the other hand, had complicated program types. Not only are there some mixes of the four types of future teachers—generalists with mathematics, generalists without mathematics, specialists in two subjects with mathematics as one of these two, and specialists in two subjects without mathematics—but also there are programs that prepare future teachers to teach grades with wide spans, such as 1–10 and 5–13. What kinds of courses or content can a program offer for a future teacher to be eligible to teach from Grade 1 to Grades 5–8 and further into Grades 9 and 10? It seems reasonable to conclude that the German TEP's model of specialization and teaching grade spans does not produce positive results with respect to the courses/content arrangement. Another possible reason for Germany's situation may be the fact that its TEP has struggled with different forms of revisions and reforms since the 1970s and that its state (*Länder*) ministries are formally in charge of the structure, course content, and methods of teacher education, causing considerable differences among the 16 states. Consequently, this may lead Germany's future teachers to feel at a loss as to what to do (Foraker 1999).

The correlation of this indicator with knowledge achievement was again calculated. The results show that the degree of well-organized curriculum from future teachers' views and their MCK and MPCK achievements are not statistically correlated. These results suggest that the future teachers' MCK and MPCK achievements and the organization or arrangement of courses they received in TEPs are not necessarily related. One possible reason for this result is that there might be other factors relating to courses, such as the amount and difficulty level of the course content, that influence knowledge achievement.

With regard to the educators' viewpoints, the data in Table 4 primarily indicate that no matter what types of educators exist within a given country, the means of the logit scores exceeded the neutral score of 10, which implies that educators in every country approved of their courses/content arrangements on average. Thus, an

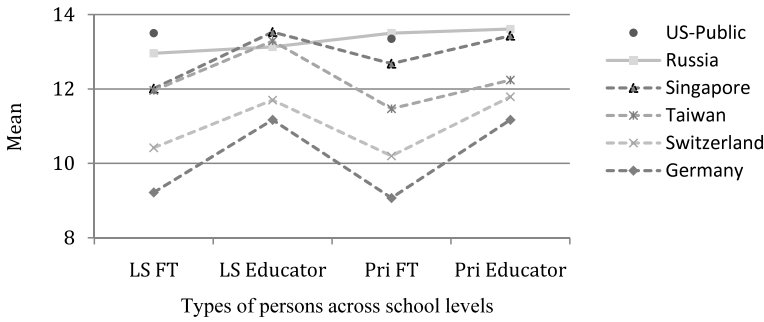


Fig. 3 The within-country trends of courses/content arrangement across school levels for the six countries that have all means of knowledge achievement above the international mean of 500. *LS* = lower secondary; *Pri* = primary; *FT* = future teachers. The lines of US-Public are not drawn since its data of educators were not processed

inconsistency in the evaluations among future teachers and educators on the effectiveness of courses/content arrangement appears, prompting us to start our inspection by focusing on the mean differences of future teachers and educators within countries. A common pattern emerged in that educators at both levels gave statistically significantly higher ratings than future teachers did for all applicable countries, with the exception of two, the Russian Federation being one. This phenomenon tells us that educators had more confidence in courses/content arrangements in TEPs than did future teachers. However, educators are often the planners and executors of their curricula; therefore, the higher ratings they provided may translate into a lack of motivation to improve. These results in issues worth considering: For example, does this idea mean that educators are unfamiliar with the lower secondary and/or primary level, or are they just more optimistic? What kinds of courses/content do future teachers desire or need?

Undeniably, in terms of courses/content arrangement, each TEP possesses a different degree of focus on either the advanced or the pragmatic standpoints. Based on our scales, when the arrangement of courses/content are rated to the same degree from both the educators’ advanced standpoints as well as the future teachers’ pragmatic standpoints, then the TEP shows an arrangement that possesses equilibrium. From the six countries that all have achievement means above the scale mean of 500, we discovered two different patterns (see Fig. 3). Germany, by using the type of all educators as an estimation of both the lower secondary and primary educators, shares the same pattern as Switzerland, where the arrangements are much more valued from the advanced view at both school levels. Taiwan and Singapore also are similar to Germany and Switzerland but with a slight difference at the primary level in that the degree of equilibrium between both views is slightly better.

The Russian Federation, however, presented a totally different pattern in that its TEP is in perfect equilibrium. The United States, like the Russian Federation, having almost the highest ratings from future teachers, unfortunately did not have the educators’ data, and this fact hindered an investigation on whether they would fall

in the same pattern as the Russian Federation in terms of the degree of equilibrium between both views in courses/content arrangement.

4.5 Overall Effectiveness

The overall effectiveness of TEPs in educating future teachers on mathematics teaching is taken as the last indicator of teacher education quality with respect to the ratings of persons inside the TEPs. Again, both the pragmatic view and the advanced view are valuable in this indicator. The international TEDS-M included a question at the end of both the future teachers' and the educators' questionnaires inquiring about this topic. Four levels of ranks (*very ineffective*, *ineffective*, *effective*, and *very effective*) were provided as the levels of satisfaction in association with effectiveness, ranging from one to four points, respectively.

Table 5 presents the means of overall effectiveness for each of the participating countries in both the lower secondary and primary levels.

One observation from the lower secondary future teachers' data was that the countries whose overall effectiveness ranked in the top six in terms of the means of the effectiveness points were all among the eight countries that prepared only specialists teaching in one subject. The Russian Federation and Taiwan fell into this category. For those countries that prepared some generalists, namely, Chile, Norway, and Switzerland, the effectiveness means were low. An important inference made from these preliminary results is that effectiveness is influenced by the degree of specialization for the lower secondary level.

With regard to primary future teachers, the means of the effectiveness points of the nine countries preparing only generalists were spread across the rating scale. The only two countries that prepared only specialists—Thailand (in one subject) and Malaysia (in two subjects)—ranked high, at the third and fourth positions. From the means of the effective points, we can see that some of the countries' programs had reached the effectiveness threshold of three points from both the future teachers, who represent a pragmatic standpoint, and the educators, who represent an advanced standpoint. These countries include Taiwan at the lower secondary level and the Russian Federation at the primary level. Unfortunately, at both levels, Germany and Switzerland, as well as the United States at the primary level, did not meet the threshold of three points.

Another interesting issue relates to the differences in ratings from the future teachers' pragmatic standpoint and the educators' advanced standpoint. We started our inspection by focusing on the mean differences of future teachers and educators within countries. The differences were very small for each country (absolute values less than 0.3), indicating near equilibrium of the two standpoints in a system. These results were different from those for the courses/content arrangement indicators. For the Russian Federation, the degrees of equilibrium concerning advanced and pragmatic views were high for both the courses/content arrangement and overall effectiveness in preparing future teachers to teach mathematics at both

Table 5 Country means of effectiveness points of future teachers and educators rating their teacher education programs and the differences in the means

Lower secondary level				Primary level			
Country	FT	Edor	DM	Country	FT	Edor	DM
	M	M			M	M	
<i>Botswana</i>	3.27	3.26	0.01	Russia	3.24	3.02	0.22**
<i>Russia</i>	3.24	2.97	0.26**	Philippines	3.19	3.24	-0.05
<i>Philippines</i>	3.20	3.14	0.06	<i>Thailand</i>	3.15	3.04	0.11**
<i>Oman</i>	3.19	3.15	0.04	Malaysia	3.14	3.00	0.14**
<i>Taiwan</i>	3.14	3.28	-0.14**	Singapore	3.14	2.93	0.21*
<i>Thailand</i>	3.11	3.04	0.07*	Botswana	3.13	2.85	0.28
Malaysia	3.04	3.02	0.02	Georgia	3.08	3.08	0.00
Singapore	3.04	2.95	0.08	US-Public	3.03		
<i>US-Public</i>	2.98			Taiwan	2.83	2.88	-0.04
Georgia	2.94	3.15	-0.21	Switzerland	2.81	2.73	0.08
<i>Poland</i>	2.94	2.99	-0.05	Poland	2.77	2.91	-0.14**
Switzerland	2.84	2.80	0.04	Chile	2.66	2.82	-0.16**
Germany	2.75			Norway	2.59		
Chile	2.53	2.82	-0.29**	Spain	2.51	2.66	-0.15*
Norway	2.48			Germany	2.40		

Note. FT = future teacher; Edor = educator; DM = difference of means, which was obtained by subtracting future teachers' means of the ratings of program effectiveness by educators' means for each country. The significances of the differences of means were tested by *t* tests. The italicized countries are those that prepare only specialists teaching in one subject. The German mean for all educators, which is not distinguishable by levels, is 2.75. It is higher than the average of their future teachers' means, 2.58, for both levels. * $p < 0.05$, ** $p < 0.01$

levels. This point means that the TEP of the Russian Federation showed balance between advanced and pragmatic views. On the other hand, the educators of Germany, Singapore, Switzerland, and Taiwan gave higher ratings for courses/content arrangement than their future teachers did; however, ratings for overall effectiveness were balanced among educators and future teachers. This development may be a result of the educators' confidence in the curricula that they have formed, thus causing their ratings for courses/content arrangement to be higher. However, for the educators of these countries, with the incorporation of other considerations into the overall evaluation of the effectiveness of a country in preparing future teachers to teach mathematics, such as the courses taken and the level of knowledge of their students, their overall ratings became lower. In contrast, the future teachers did not find the courses/content arrangements effective, but their ratings for the other factors, such as the effectiveness of instructors or the satisfaction of their knowl-

edge achievement, managed to bring their overall ratings close to that of the educators.

But were future teachers' ratings of the overall effectiveness of their TEPs associated with their own MCK and MPCK achievements? Spearman's rank correlation analyses showed that in all countries, except the United States and Germany, at both school levels the correlation coefficients (significant) were comparatively low ($-0.2 < r < 0.2$). Germany and the United States, however, reached the 0.05 level of statistical significance.⁶ In addition, conducting Spearman's rank correlation analyses to examine whether or not the correlation exists between countries' means of overall effectiveness and their MCK and MPCK scores also returned a negative result.

Finally, this study tried to examine what kinds of effectiveness indicators influenced future teachers or educators in evaluating the overall effectiveness of their TEPs. Spearman's rank correlation coefficient analyses were used at the individual level for every country because all the scales of effectiveness were ordinal. Results showed that, regardless of the school level, the correlations between the overall effectiveness and other indicators of effectiveness ranked in the following order: MR-instructors, courses/content arrangement, teaching coherence, and SB-supervisors (see Table 6). At the lower secondary level, there were two major groups of five and six countries categorized in terms of which indicators influenced the countries' overall effectiveness the most. Germany, the Russian Federation, Taiwan, and the United States all fell into the first group, where their overall TEP effectiveness in preparing future teachers to teach mathematics was most influenced by the effectiveness of the MR-instructors. Singapore and Switzerland fell into the second group, with their overall effectiveness most influenced by the effectiveness of both MR-instructors and courses/content arrangement. At the primary level, the dominating pattern was the second group, which consisted of eight countries. The United States did not fall into this group because the indicator of teaching coherence between the teaching of universities and that of schools further influenced its overall effectiveness. The results indicated that there are more factors involved in making primary future teachers feel that their TEPs are capable of preparing them for mathematics teaching than in the lower secondary level. The US primary-level programs especially need to take care of at least three indicators in order for future teachers to feel satisfied with its overall TEP effectiveness.

From the data shown in Table 6, we are aware that there are some countries in which none of the four indicators has an effect on the overall effectiveness ratings. The reason for this phenomenon is still unknown and thus needs further study.

⁶At the lower secondary level, only the United States had significant correlation coefficients for mathematics content knowledge (MCK) and mathematics pedagogical content knowledge (MPCK), 0.22 and 0.23, respectively. At the primary level, Germany had significant correlation coefficients for MCK and MPCK, 0.37 and 0.33, respectively.

Table 6 Noticeable significant correlations between future teachers’ ratings of overall effectiveness and other indicators of the effectiveness of TEP

Lower secondary level					Primary level				
Country	MR-I	CA	TC	SB-S	Country	MR-I	CA	TC	SB-S
Oman	•	•	•	•	Botswana	•	•	•	•
Norway	•	•			US-Public	•	•	•	
Poland	•	•			Taiwan	•	•		
Switzerland	•	•			Spain	•	•		
Chile	•	•			Poland	•	•		
Singapore	•	•			Thailand	•	•		
Thailand	•	•			Norway	•	•		
US-Public	•				Chile	•	•		
Russia	•				Switzerland	•	•		
Taiwan	•				Singapore	•	•		
Botswana	•				Russia	•			
Germany	•				Germany	•			
George					Malaysia				
Malaysia					Philippines				
Philippines					Georgia				

Note. MR-I = instructor in mathematics-related courses; CA = *courses/content arrangement*; TC = *teaching coherence*; SB-S = school-based supervising teacher. A dot (•) indicates $r_s \geq 0.3$ at $p < 0.01$, with the exception of Botswana at $p < 0.05$. The significances for $r_s < 0.3$ are not shown in the table

By considering the comparisons across countries, some factors influential to overall effectiveness became apparent. One factor is future teachers’ knowledge achievement. Taking the United States and Taiwan at the lower secondary level as examples, with Taiwan’s ratings in overall effectiveness higher than those of the United States, we see that the United States, having all other indicators of effectiveness rated highly with only knowledge achievement placing at the middle, did not receive higher ratings in overall effectiveness compared with Taiwan, which had effectiveness indicators that usually did not reach the levels of the United States but which had extremely high achievement scores. In the same way, Singapore and Taiwan, being high-achieving countries, together with Germany, a mid-achieving country, and Switzerland, an upper-half-achieving country, did not reach the same ranks in their overall effectiveness as their knowledge achievement, given that their other indicators of effectiveness were not as positive as their achievement levels. The Russian Federation, the only country that ranked high for all indicators of effectiveness and achievements, always remained within the top two ranks in overall effectiveness. Perhaps this point is exemplary of a country’s TEP that can make its future teachers feel that they are being aptly prepared to teach mathematics, therefore proving that all indicators of effectiveness and achievements are necessary.

5 Conclusion

Through international comparison, countries around the world acquire opportunities to learn from each other and reflect on themselves for the future. This initial study looks at the picture of mathematics teacher education quality in terms of effectiveness across countries by constructing a number of indicators and using various TEDS-M collected or scaled data. These indicators not only consider TEPs in terms of the outcome of knowledge future teachers possess but also the persons involved at the other end of TEPs—academic and school-based instructors; effectiveness is evaluated both from the perspectives of these persons and from the circumstances. Certain types of these data had never been collected prior to this study, such as statistics relating to educators. Several reflections and implications can be drawn from this study.

5.1 *Effectiveness of Instructors and Teaching Coherence*

Whether practical teaching and theoretical knowledge should both be included in the TEP, and how much of each is necessary, has always been a topic of discussion. This study shows that future teachers report that they benefit from both academic and school-based instructors in every participating country, and this result supports the necessity and appropriateness of integrating theoretical knowledge and practical teaching into teacher education. Based on the fact that future teachers in more than two thirds of the countries gave evidence that the effectiveness of SB-supervisors is higher than that of MR-instructors in helping them become well-trained teachers, it seems reasonable for countries to raise the following question: Should we reorganize our TEPs in order to allow future teachers more time in practicum?

Regardless of from whom the future teachers have benefitted the most, all countries' future teachers rated both positions of instructors as effective in providing professional help to facilitate them in becoming well-trained teachers. However, the effectiveness of instructors does not produce any noticeable influences on future teachers' knowledge achievement. Given that the future teachers feel their instructors are effective in educating them, they must have learned something and been influenced by their instructors. So what aspects of the future teachers' experience did the educators influence—their future classroom teaching or their knowledge achievement? This question is worth investigating and reminds us that a paper-and-pencil measure may not provide the whole picture for the achievements of future teachers.

The MR-instructors usually provide a theoretical foundation, and the SB-supervisors can use the future teachers' qualifications and what they have learned in the universities to strengthen their real classroom teaching. This study produces a concept of synchronization by joining the three indicators—teaching coherence, the effectiveness of MR-instructors, and the effectiveness of SB-supervisors. When

the three indicators are highly rated, this reveals that university and school instructors are effective and make use of tightly integrated teachings. A program with this characteristic is regarded as synchronized.

The United States demonstrated an effective model in terms of having the most synchronized teaching in TEPs at both the lower secondary and primary levels. This means that we can expect that US teachers will be good at real classroom teaching in terms of building their instructional frameworks together with theoretical support. Other countries, for example, Taiwan, should reflect on what they could learn from the features of the US TEP. Further study regarding this issue is needed.

5.2 Courses/Content Arrangement and Overall Effectiveness

A high-quality TEP not only is synchronized in its teaching but also needs to attend to the organization of the sequences and links of the courses/content to meet the needs of future teachers. For this reason, the indicator courses/content arrangement emerges. This indicator serves as a criterion to determine whether a program has a well-organized curriculum in educating future teachers. Ideally, a program is well organized if it is perceived as being equipped with a well-organized curriculum from both the advanced and pragmatic points of view. However, given that some of the countries involved did not provide educator samples, the rating from an advanced view could not be obtained; therefore, this study employs only the pragmatic view of future teachers.

The United States and Russian Federation are good examples of well-organized programs. In contrast to the United States, Germany shows a lack of organization in its curriculum. Further analysis of these countries' cases shows that the complicated mixes of specializations and teaching grade spans may influence the organization of curricula in TEPs. It is easy to recognize the difficulty of building a curriculum that encourages competence in preparing a future teacher to be eligible to teach from the primary to senior high grades.

A complicated mixture of different specializations and teaching grade spans also shows a negative influence on overall effectiveness, the last and most comprehensive indicator of the quality of TEPs. The six countries chosen for further investigation in this study provide evidence showing the tendency that TEPs preparing only specialists at the lower secondary level and TEPs preparing generalists at the primary level are better in terms of overall effectiveness; however, the mixture of specializations and grade spans is not the only influence involved. Other factors, such as the effectiveness of MR-instructors and/or the courses/content arrangement, also show moderate influences on future teachers' ratings. Among the six countries we chose to further investigate, we found that in order for a TEP to make its future teachers feel that they are being aptly prepared to teach mathematics, all indicators of effectiveness and achievement are necessary and influential to overall effectiveness.

Last but not least, focus should be put on the equilibrium of a TEP between both the advanced and pragmatic views. For most of the countries involved in

this research the overall effectiveness is in equilibrium; however, the indicator courses/content arrangement is not balanced. This phenomenon produces an issue worthy of consideration by the mathematics education community. Another fact is that among all participating countries, all levels of educators, being the planners and executors of courses/content arrangements, rated their arrangements in providing suitable courses/content for their students much higher than their students did. The higher ratings they provided may translate into a lack of motivation to improve. To us, as teachers of teachers, this is not only a heavy blow but also a wake-up call raising issues worth considering: Does this mean that teacher educators are too unfamiliar with the situations at the lower secondary and/or primary levels, or are they just being more optimistic? What kinds of courses/content do future teachers desire or need? These are issues that the academic community should immediately pursue.

This study discovered that many effectiveness indicators do not correlate with the knowledge indicator, which is regarded as the most important indicator by some people. Somewhat based on the results from the section titled Overall Effectiveness, we may put forth a hypothesis that other factors exist that may be combined with our indicators to guarantee the knowledge of future teachers. For example, the mathematics knowledge of future teachers at the entry point of the TEP or the amount and depth of the courses taken in the TEP may be other factors that influence future teachers' knowledge at the exit point. Further research is suggested.

Research concerning teacher education has always been highly valued; yet, how many national studies expose this reality? Although some studies (e.g., Judge et al. 1994) criticize US teacher education as a "non-system" in that it is not under national control but has a great deal of autonomy for teacher educators, it is worth noting that, as observed from this international comparison study, the US TEP is synchronized and well organized from the pragmatic views of its future teachers. One thing to which the United States should pay more attention is the elevation of its future teachers' MCK and MPCK, which may be the reason why the overall effectiveness of the US TEP does not stand out in the international ranks.

From the abundance of information this research has obtained, it is reasonable to say that this international comparison study provides new information to many countries. The insufficient aspect of this study was the small number of participating countries; it therefore lacked complete international representation. Furthermore, some countries, like the United States, provided insufficient data concerning their educators; this caused certain pieces of information to remain unresolved, and therefore they could not be presented. Nevertheless, the results of this initial analysis show that teacher education matters and that international teacher education studies are valuable. This should not mark the end of teacher education studies; instead, this is the beginning.

References

- Abell Foundation (2001). *Teacher certification reconsidered: stumbling for quality*. Baltimore: Author. http://www.abell.org/pubsitems/ed_cert_1101.pdf.

- Allen, M. (2003). *Eight questions on teacher preparation: what does the research say?* Denver: Education Commission of the States. <http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED479051>.
- Archambault, L., & Crippen, K. (2009). Examining TPACK among K-12 online distance educators in the United States. *Contemporary Issues in Technology and Teacher Education*, 9(1). <http://www.citejournal.org/vol9/iss1/general/article2.cfm>.
- Bates, A., Ramirez, L., & Drits, D. (2009). Connecting university supervision and critical reflection: Mentoring and modeling. *Teacher Educator*, 44(2), 90–112.
- Clark, D. (1992). Search for a more effective future. In H. Gideonse (Ed.), *Teacher education policy: narratives, stories, and cases* (pp. 269–295). Albany: SUNY Press.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1), 1–50.
- Ducharme, E. R., & Ducharme, M. K. (1999). Teacher educators and teachers: The needs for excellence and spunk. In R. A. Roth (Ed.), *The role of the university in the preparation of teachers* (pp. 41–58). London: Falmer.
- Enochsson, A. (2010). ICT in initial teacher training: Sweden country report (OECD document). <http://www.oecd.org/dataoecd/43/15/45046846.pdf>.
- Ferguson, R. (1998). Can schools narrow the Black–White test score gap? In C. Jencks & M. Phillips (Eds.), *The Black–White test score gap* (pp. 318–374). Washington: Brookings Institution.
- Ferrini-Mundy, J., Floden, R., McCrory, R., Burrill, G., & Sandow, D. (2005). *A conceptual framework for knowledge for teaching school algebra*. East Lansing: Authors.
- Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives*, 7(9). <http://epaa.asu.edu/ojs/article/viewFile/544/667>.
- Florida State Department of Education (1983). *Improving the quality of teacher education in Florida: report and recommendations of the joint legislative and executive task force for teacher education quality improvement*. Tallahassee: Author.
- Foraker, W. C. (1999). The educational structure of the German school system. In H. Stevenson, S. Lee, & R. Nerison-Low (Eds.), *Contemporary research in the United States, Germany, and Japan on five educational issues: Structure of the education system, standards of education, the role of school in adolescents' lives, individual differences among students, and teachers' lives* (pp. 151–313). Washington: US Department of Education.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. <http://nctq.learningpt.org/publications/LinkBetweenTQandStudentOutcomes.pdf>.
- Goldhaber, D., & Anthony, E. (2004). Can teacher quality be effectively assessed? <http://www.urban.org/publications/410958.html>.
- Goldhaber, D., & Brewer, D. (1997). Evaluating the effect of teacher degree level on educational performance. In W. Fowler (Ed.), *Developments in school finance* (pp. 197–210). Washington: US Department of Education/National Center for Education Statistics.
- Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129–145.
- Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What knowledge matters and what evidence counts? In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 111–155). Charlotte: Information Age.
- Howey, K. (1995). *Tomorrow's schools of education*. East Lansing: Holmes Group.
- Hsieh, F.-J., Wang, T.-Y., Hsieh, C.-J., Tang, S.-J., Chao, G.-H., Law, C.-K., et al. (2010). A milestone of an international study in Taiwan teacher education—an international comparison of Taiwan mathematics teacher (Taiwan TEDS-M, 2008). <http://tedsm.math.ntnu.edu.tw/eng/result.htm>.
- Hsieh, F.-J., Law, C.-K., Shy, H.-Y., Wang, T.-Y., Hsieh, C.-J., & Tang, S.-J. (2011). Mathematics Teacher Education Quality in TEDS-M: Globalizing the Views of Future Teachers and Teacher Educators. *Journal of Teacher Education*, 62, 172–187.

- Ingvarson, L., Beavis, A., & Kleinhenz, E. (2007). Factors affecting the impact of teacher education courses on teacher preparedness: Implications for accreditation policy. *European Journal of Teacher Education*, 30(4), 351–381.
- Judge, H., Lemosse, M., Paine, L., & Sedlak, M. (1994). *Oxford Studies in Comparative Education: Vol. 4. The university and the teachers: France, the United States, and England*. Oxford: Symposium Books.
- Kaplan, L. S., & Owings, W. A. (2001). Teacher quality and student achievement: Recommendations for principals. *NASSP Bulletin*, 85(628), 64–73.
- National Center for Education Statistics (1999). *Teacher quality: a report on the preparation and qualifications of public school teachers (NCES 1999-080)*. Washington: US Department of Education/Office of Education Research and Improvement. <http://nces.ed.gov/pubs99/1999080.pdf>.
- National Center for Education Statistics (2003). *Teaching mathematics in seven countries results from the TIMSS 1999 video study (NCES 2003-013)*. Washington: US Department of Education. <http://nces.ed.gov/pubs2003/2003013.pdf>.
- National Center for Education Statistics (2009). *Trends in International Mathematics and Science Study (TIMSS)*. <http://nces.ed.gov/timss/video.asp>.
- National Research Council (2010). *Preparing teachers: building evidence for sound policy*. Washington: National Academies Press (Committee on the Study of Teacher Preparation Programs in the United States, Division of Behavioral and Social Sciences and Education). http://www.nap.edu/catalog.php?record_id=12882.
- Organisation for Economic Co-operation and Development (2007). *OECD Programme for International Student Assessment (PISA)*. http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1,00.html.
- Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4–15.
- Rice, J. K. (2003). *Teacher quality Understanding the effectiveness of teacher attributes*. Washington: Economic Policy Institute.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
- Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): conceptual framework*. East Lansing: Michigan State University Press.
- Tatto, M. T., Schwille, J., Senk, S., Rodriguez, M., Bankov, K., Reckase, M. D., et al. (2009). *Teacher Education Study in Mathematics (TEDS-M): technical summary*. East Lansing: Michigan State University Press.
- Wang, J., Spalding, E., Odell, S., Klecka, C., & Lin, E. (2010). Bold ideas for improving teacher education and teaching: What we see, hear, and think. *Journal of Teacher Education*, 61(1–2), 3–15.
- Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college- and university-based teacher education. *Journal of Teacher Education*, 61(1–2), 89–99.