IEA POLICY BRIEF

Researching education, improving learning

NUMBER 13 FEBRUARY 2017

Exploring cross-national changes in instructional practices: evidence from four cycles of TIMSS

BY **MOJCA ROŽMAN** & **ECKHARD KLIEME** (GERMAN INSTITUTE FOR INTERNATIONAL EDUCATIONAL RESEARCH; DIPF)

SUMMARY

The IEA's Trends in International Mathematics and Science Study (TIMSS) can reveal a great deal about national teaching cultures and policies. Changes in instructional practices over more than a decade may be studied from a cross-cultural perspective. Using teacher and student reports on the frequency of seven distinct practices, as assessed by TIMSS in 1995, 1999, 2003 and 2007, it is possible to identify a number of significant changes between adjacent time points for individual countries. Teaching practices seem to be shaped by national educational cultures or policies. While this brief does not identify the causes or effects of these changes, there is mixed evidence for global "mega-trends" in education, as discussed by researchers and policymakers. Constructivist pedagogy (working in groups, applying mathematical content to daily life) was boosted on a large scale during the mid-2000s. However, there is only limited support for a rise in assessment-based instruction and a shift from computational practice to problem-solving in mathematics. TIMSS evidence suggests a quest for systematic development of teaching practices, classroom cultures, and teacher quality at the national level.

INTRODUCTION

Classroom processes are the most immediate and probably strongest factors involved in explaining school effectiveness (Hattie, 2009). Student learning and student outcomes are driven by the quality of teaching and learning within the classroom. The IEA's Trends in International Mathematics and Science Study (TIMSS) offers broad information on instructional practices and has been used to study their effects on student achievement scores (see Schwerdt, & Wuppermann, 2009; Zuzovsky, 2013; Mullis, Martin, & Loveless, 2016; Nilsen, & Gustaffson, 2016). However, most of the research work so far has focused on patterns and relationships within a single assessment wave of TIMSS[1]. This brief aims to widen the perspective, studying changes in instructional practices over more than a decade from a cross-cultural perspective. We focus on grade 8 mathematics teaching and learning.

1 For a rare exception, see the sophisticated analysis of changes in instruction in Israel 1999–2003 by Zuzovsky (2008).

CONTENTS

Although instructional practices impact on individual student learning at the classroom level, we examined these practices from a system-level perspective. The aim was to observe change in the use of specific instructional practices in mathematics teaching over four cycles of TIMSS (1995–1999–2003–2007) across 18 different countries.

From a policy perspective, we aimed to determine whether there was any significant change in instructional patterns for individual countries, or even across countries worldwide; if so, instructional practices could be considered malleable. Policies promoting certain practices may affect changes in classroom practices. If no significant change is evident, it seems reasonable to conclude that instruction is deeply rooted in pedagogical culture and traditions, with little chance of inducing change.

"Changes occurring in the frequency and effectiveness of certain instructional modes often reflect changing pedagogical fashions worldwide" (Zuzovsky, 2008, p. 66). Sharing this assumption, we start from the following hypothesis: from 1995–2007, some mega-trends occurred in debates on mathematics teaching and learning worldwide, which, if taken up by educational policy and practice, should be visible in trend data on frequency of related practices.

MEGATRENDS IN EDUCATIONAL ASSESSMENT, PEDAGOGY, AND SUBJECT-MATTER DIDACTICS

Over the last decades, several mega-trends have been observed in the international educational debate. This brief refers to, and intends to study empirically, three different trends:

1. Educational assessment: The literature on assessment, evaluation, and accountability policies (Bayer, Klieme, & Jude, 2016), as well as reanalyses of international data reported by school principals (Teltemann, & Klieme, 2016), have shown that there has been a growing interest in regular student testing for some decades, in the form of both formative and summative assessments, to support classroom learning and school evaluation.
2. Pedagogy: Teacher educators and researchers have increasingly advocated constructivist approaches involving student-oriented teaching and self-regulated student activity, rather than teacher-centered instruction (see, for example, Seidel, & Shavelson 2007; Tobias, & Duffy, 2009).
3. Subject-matter didactics: Inspired by, for example, the "Standards for Curriculum and Evaluation" established by the National Council of Teachers of Mathematics (NCTM, 1989), with support from distinguished scholars such as Alan Schoenfeld (2006), mathematics educators in the USA and worldwide promoted mathematics education that aimed at reasoning and problem-solving activities rather than training computational and procedural skills.

DATA

We used data from the TIMSS populations of 8th grade students that were assessed in 1995 (the first TIMSS study), 1999, 2003 and 2007. Unfortunately, the frequencies of specific teaching practices were not assessed in the 2011 and 2015 TIMSS questionnaires, limiting the current analysis to the period 1995–2007. The selection of education systems was limited to those that had participated in all four cycles. We used both student and teacher data from the following 18 education systems, hereafter referred to as countries:

- Five English-speaking countries: Australia, Canada (Ontario), Canada (Quebec)[2], United States, and England;
- Five Central and Eastern European countries: Hungary, Lithuania, Romania, Russian Federation, and Slovenia;
- Four East Asian countries: Hong Kong SAR, Japan, Republic of Korea, and Singapore; and
- Four additional countries: Cyprus, Islamic Republic of Iran, Israel, and Italy.

2 Canada participated as a country in 1995 and 1999, and as separate provinces in 2003 and 2007. Thanks to the support of Statistics Canada and the National Centre, we were able to disaggregate Canada's 1995 and 1999 data (when the provinces were used as a stratification variable). Accordingly, the sample size, especially for Quebec in 1995 and 1999, was smaller than in 2003 and 2007.

INDICATORS OF INSTRUCTIONAL PRACTICE

The choice of variables was limited to items that had been administered similarly or unchanged over the four cycles. After reviewing all the items concerning instructional practices in mathematics in the student and teacher questionnaires, five items were retained from the student questionnaire (relating the mathematics content to daily life, working in groups, having a quiz or test, using a calculator, and beginning homework in class) and two items from the mathematics teacher questionnaire (practicing computational skills and working on problems without an immediately obvious solution).

Five of these seven items refer to the deeper level of mega-trends mentioned previously, while two items report on surface-level aspects of classroom instruction for which we do not endorse clear expectations or hypotheses, namely "Using a calculator" and "Beginning homework in class".

RESEARCH HYPOTHESES

Hypothesis 1: We expect an increase in the administration of tests and quizzes because of the growing promotion of assessment in education.

Hypothesis 2: We expect an increase in use of authentic, applied tasks, and working in small groups because these are indicators of 'new pedagogy'.

Hypothesis 3: We expect an increase in problem solving activities and a decrease in practicing computational skills because of the move towards reasoning and problem solving in math didactics.

Given the growing exchange among educational researchers and professionals worldwide, the promotion of mega-trends by international agencies, such as the OECD, and frequent policy borrowing between countries, we expect these changes to be almost universal. Although there are likely differences between countries, we do not propose any specific predictions regarding which country would have experienced more or less change. For example, it is well known that assessment methods have been more prominent and promoted earlier in the English-speaking world than in other countries, therefore it may be reasonable to assume that other countries followed, yet the gap between English-speaking and other countries may also have grown larger over the 12 years covered.

METHOD

Respondents were asked to report on how often the selected practices happened in their mathematics lessons, answering on a four-point scale, which remained the same for TIMSS 1995 and TIMSS 1999 (1 = almost always, 2 = pretty often, 3 = once in a while, and 4 = never for students; and 1 = every lesson, 2 = most lessons, 3 = some lessons, and 4 = never or almost never for teachers). The scales were changed for TIMSS 2003 (to 1 = every or almost every lesson, 2 = about half the lessons, 3 = some lessons, and 4 = never for both students and teachers) and remained unchanged for TIMSS 2007. No national adaptations were documented in any of the cycles for the chosen variables.

The wording of two items changed slightly:

- relating the mathematics content to daily life was in 1995/1999 worded as "We use things from everyday life in solving mathematics problems" and in 2003/2007 it was changed to "We relate what we are learning in mathematics to our daily lives"; and
- practicing computational skills was in 1995/1999 worded as "Practice computational skills" and in 2003/2007 it was changed to "Practice adding, subtracting, multiplying, and dividing without using a calculator".

Despite these changes, we decided to include these two items in our analyses, as we believe that the same idea is being assessed with both versions and they represent interesting and important aspects of mathematics instruction.

We used a design-based rank test (Lumley, & Scott, 2013; Lumley, 2016) to explore changes in frequency of teaching practices over time. Non-parametric estimators account for the ordered nature of the potential responses, and do not impose any restriction on the distance between two consecutive levels. Student variables were weighted using total student weight and teacher variables were weighted using mathematics teacher weight.

All the analyses were conducted on a country level basis. We first merged the student and the teacher data, and then we combined two successive cycles in one dataset. Three different datasets for each country were thus obtained. In the first dataset, TIMSS 1995 data was combined with TIMSS 1999 data. In the second dataset, TIMSS 1999 data was combined with the data from TIMSS 2003; here (and only here), the two middle response categories were collapsed because of the changed scale. In the third dataset, TIMSS 2003 data was combined with TIMSS 2007 data. In each dataset, the change between cycles was then observed.

We report only the effect size r of the change within a pair of cycles (the test statistic z was divided by the square root of the number of all observations over the two time points). A positive effect size indicates that there was an increased use of that particular practice in the second cycle compared to the first one, and negative values indicate that the frequency of use decreased for the particular practice in the second cycle in comparison to the first.

The present report on purpose refrains from reporting and comparing country means for any of the seven items. We know that most questionnaire scales using Likert-type answering options in cross-cultural studies are not scalar invariant. Therefore, scale means cannot be compared meaningfully between countries, *a fortiori* means on individual items should not be compared either. By reporting effect sizes for change within countries, however, we do interpret this data with the necessary degree of cautiousness.

RESULTS

Changes in educational assessment: using tests or quizzes

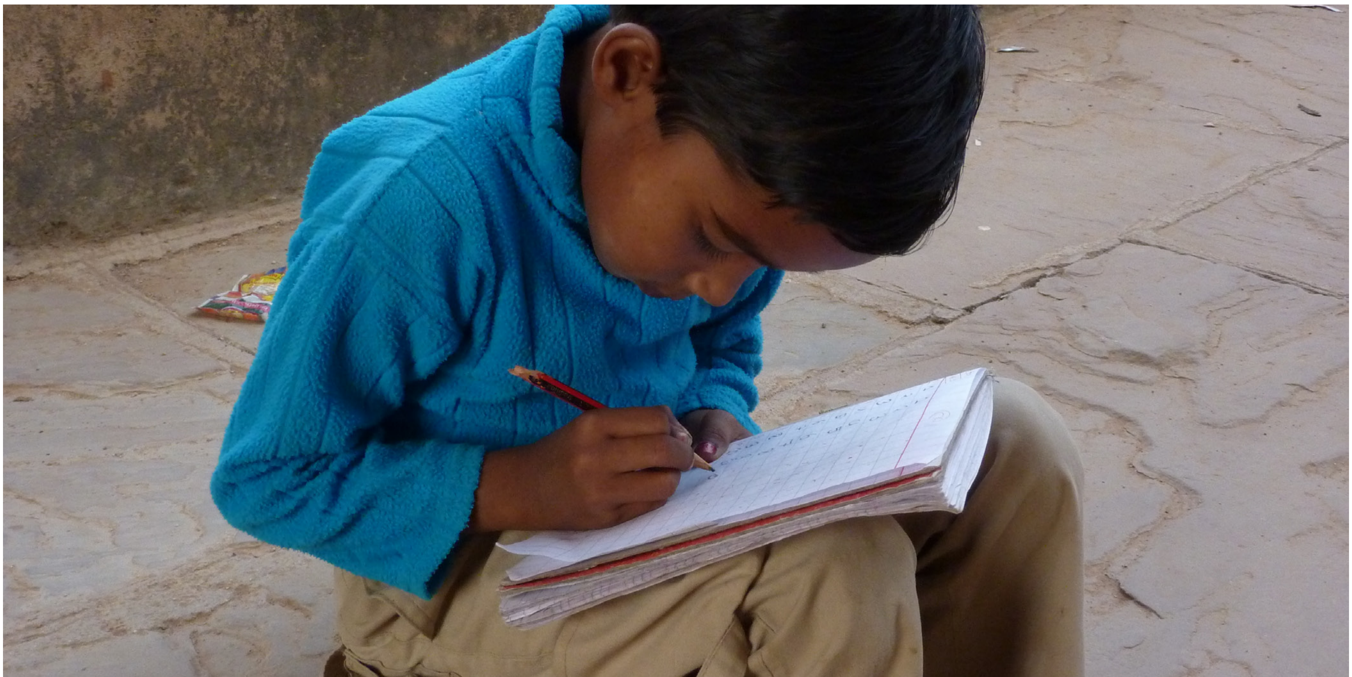
The strong "universal" hypothesis of increased use of assessment in classrooms is definitely not supported by the observed changes in frequency of having a quiz or test (Table 1), although there are consistent gains in Hungary, Israel, and the United States, and some strong increases for single periods in Cyprus, Iran, and Italy. Several countries (most notably Cyprus) show varying changes, while consistent decrease is observed for Romania only.

In conclusion, there seems to be a slight tendency towards increased use of assessment in classrooms on an international scale (especially over the time period 1995–1999), but the changes appear to be inconsistent both within and between countries.

Table 1: Effect size of the change in frequency of use of tests or quizzes between successive TIMSS cycles, as reported by students

Country	Having a quiz or test		
	1995–1999	1999–2003	2003–2007
Australia	0.04*	−0.08*	0.01
Canada (Ontario)	0.12*	−0.06*	−0.01
Canada (Quebec)	−0.04	−0.03	0.07*
Cyprus	0.03*	0.38*	−0.29*
England	0.10*	−0.16*	0.02
Hong Kong, SAR	−0.40*	−0.04*	0.09*
Hungary	0.09*	0.35*	0.07*
Iran, Islamic Republic of	−0.04*	0.61*	−0.26*
Israel	0.06*	0.01	0.00
Italy	0.00	0.28*	−0.08*
Japan	0.01	0.03*	−0.19*
Korea, Republic of	−0.02	0.09*	−0.12*
Lithuania	–	–	0.04*
Romania	−0.01	−0.21*	−0.02
Russian Federation	0.01	−0.04*	0.04*
Singapore	−0.07*	−0.05*	0.03*
Slovenia	0.10*	−0.14*	−0.05*
United States	0.01	0.07*	0.02
Median	0.01	−0.03	0.00
Number of countries with significant increase	7	7	6
Number of countries with significant decrease	3	8	6

Notes: *Statistically significant results ($p < 0.05$). Countries with consistent increases are marked in blue, countries with consistent decreases in yellow. No data were available for TIMSS 1999 for Lithuania.



Changes in surface methods: using calculators and doing homework

Instructional practices changed over different cycles for surface methods across the selected participating countries (Table 2). Interestingly, in contrast to the teaching practices that are presented later in this brief, there was a consistent decline in using these surface-level practices in a number of countries.

For "Beginning homework in the classroom", there is a clear tendency towards reduced frequency of use across the years. This trend may perhaps be interpreted as an improvement in the structure of classroom teaching, towards disentangling classroom learning and homework. Also, critiques directed against homework by researchers and professionals (see for example Corno, 1996) may have had an impact. Israel showed some increase during the initial time period (1995–1999), but then exhibited a strong decrease thereafter. The only country with a consistent (albeit modest) increase in this practice was Hungary.

In 1995, calculators were most rarely used in Iran, Japan, Korea, Romania and Slovenia[3]. Subsequent cycles of TIMSS revealed an increase in calculator usage in all of these countries, except for Romania. In Iran, the increase was steady, while Japan, Korea and Slovenia seem to be "late adopters," with some reduction in use in the middle period.

In 1995, calculators were most frequently used in some English-speaking countries (Quebec, England, Australia and Hong Kong, where education is also based on British influence). Some increase was subsequently observable in Quebec, while calculator use became less popular in England and, to some degree, in Australia. Frequency of use varied across cycles for Hong Kong.

Overall, the use of calculators changed quite remarkably. It dropped 1999–2003 and increased 2003–2007 in more than half of the countries. In addition to ongoing debates about the pros and cons of calculators in mathematics education (Ellington, 2015), the results may also indicate an increasing use of computers and other technology in classrooms. (We do not know whether respondents discriminated between "calculators" and other tools.) In general, the frequency of use of calculators is becoming more similar across countries.

Table 2: Effect size of the change in frequency of use of surface-level instructional practices between successive TIMSS cycles, as reported by students

Country	Begin homework in class			Using a calculator		
	1995–1999	1999–2003	2003–2007	1995–1999	1999–2003	2003–2007
Australia	0.08*	−0.10*	−0.03*	0.02*	−0.10*	0.02*
Canada (Ontario)	0.02	0.03*	−0.05*	0.09*	−0.04*	0.11*
Canada (Quebec)	0.08*	0.00	−0.06*	0.02	0.06*	0.01
Cyprus	−0.06*	−0.10*	0.00	−0.08*	−0.06*	−0.06*
England	−0.03	−0.04*	−0.07*	−0.17*	−0.17*	0.00
Hong Kong, SAR	−0.14*	0.03*	0.05*	−0.18*	0.13*	0.09*
Hungary	0.02	0.00	0.03*	−0.08*	0.05*	−0.06*
Iran, Islamic Republic of	0.05*	−0.08*	−0.06*	0.10*	0.13*	0.03*
Israel	0.10*	−0.20*	−0.01	0.03*	−0.19*	0.06*
Italy	0.02	−0.13*	−0.01	0.08*	−0.07*	0.03*
Japan	−0.11*	0.04*	0.05*	0.09*	−0.10*	0.94*
Korea, Republic of	0.08*	−0.03*	0.07*	0.10*	−0.08*	0.14*
Lithuania	–	–	−0.03*	–	–	0.04*
Romania	−0.05*	−0.17*	−0.02*	−0.11*	−0.02	0.00
Russian Federation	−0.08*	−0.22*	−0.02	−0.18*	−0.15*	−0.02
Singapore	0.04*	−0.13*	−0.01	0.08*	0.01	0.06*
Slovenia	0.07*	0.04*	−0.11*	0.02	−0.03	0.22*
United States	0.00	0.01	−0.07*	−0.01	−0.03*	−0.02*
Median	0.02	−0.04	−0.02	0.02	−0.04	0.03
Number of countries with significant increase	7	4	4	8	4	11
Number of countries with significant decrease	5	10	9	6	10	3

Notes: *Statistically significant results ($p < 0.05$). Countries with consistent increases are marked in blue, countries with consistent decreases in yellow. No data were available for TIMSS 1999 for Lithuania.

3 The table of frequencies is published in the TIMSS 1995 international report (Beaton, Mullis, Martin, Gonzalez, Kelly, & Smith, 1996, Table 5.16, p. 166).

Changes in pedagogy: applied tasks and small group work

The hypothesized "universal" change towards a new, constructivist pedagogy emerged most strongly during the most recent period observed, 2003–2007 (Table 3), when the vast majority of countries saw a significant increase both for applied tasks and for working in groups. A consistent increase was observed in quite a few countries, while no consistent decreases were identified. Japan was an outstanding case, where large gains in the use of both practices were observed in 2003–2007.

Regarding group work, the strongest increase was observed 2003–2007 (median effect size for 2003–2007 was 0.11) when all countries (but Korea and Iran) showed a significant increase and reached the highest effect size of + 1 in Japan. The increase is most prominent in Japan, Quebec, England, Hong Kong, the Russian Federation, and Singapore. In general, East Asian countries seem to have adopted collaborative learning over the years. Korea started as early as 1995–2003, Singapore also showed strong increases, despite some reduction in such practices during 1999–2003, while Japan and (to a lesser degree) Hong Kong followed these trends in 2003–2007.

Table 3: Effect size of the change in frequency of use of "constructivist" instructional practice between successive TIMSS cycles, as reported by students

Country	**Relating math to daily lives**			**Working in groups**		
	1995–1999	**1999–2003**	**2003–2007**	**1995–1999**	**1999–2003**	**2003–2007**
Australia	0.04*	–0.10*	0.06*	0.01	–0.11*	0.13*
Canada (Ontario)	0.02	–0.06*	0.05*	0.02	–0.06*	0.13*
Canada (Quebec)	0.02	–0.04*	0.02	0.16*	–0.09*	0.21*
Cyprus	0.00	0.05*	0.10*	0.04*	0.00	0.02*
England	0.03*	–0.20*	0.06*	0.00	–0.14*	0.20*
Hong Kong, SAR	0.07*	0.07*	0.02	0.04*	0.09*	0.18*
Hungary	0.02	0.22*	0.08*	–0.05*	0.06*	0.06*
Iran, Islamic Republic of	–0.23*	0.28*	–0.05*	0.07*	0.24*	–0.04*
Israel	0.05*	0.00	0.02	0.00	–0.22*	0.08*
Italy	0.01	0.10*	0.05*	0.00	–0.28*	0.03*
Japan	–0.01	0.01	0.26*	–0.12*	0.14*	1.04*
Korea, Republic of	–0.06*	0.05*	0.03*	0.28*	0.19*	0.01
Lithuania	–	–	0.08*	–	–	0.09*
Romania	–0.13*	–0.02	0.07*	–0.04*	–0.06*	0.03*
Russian Federation	0.06*	–0.08*	0.18*	0.01	–0.11*	0.18*
Singapore	0.12*	0.00	0.06*	0.23*	–0.11*	0.23*
Slovenia	0.06*	0.09*	0.12*	0.07*	–0.06*	0.13*
United States	0.02*	–0.06*	0.02*	0.02*	–0.05*	0.08*
Median	0.02	0.00	0.06	0.02	–0.06	0.11
Number of countries with significant increase	8	7	14	8	5	16
Number of countries with significant decrease	3	6	1	3	11	1

Notes: *Statistically significant results ($p < 0.05$). Countries with consistent increases are marked in blue. No data were available for TIMSS 1999 for Lithuania.

Changes in mathematics didactics: computational skills versus reasoning and problem solving

There were changes in different didactical approaches for the chosen countries (Table 4). These results contradicted the expectations of our working hypothesis. The frequency of working on problems increased early on, but decreased in 2003–2007. This decrease can be observed across all East Asian countries (Japan, Singapore, Korea and Hong Kong), and to a minor extent in Eastern and Central Europe (Lithuania, Romania and Hungary, with no significant change in Russia and an increase in Slovenia). However, practicing computations increased steadily. Practicing basic skills still seemed to be a very important activity in mathematics lessons in 2007. The *TIMSS 2007 Encyclopedia* (Mullis et al., 2008, p. 31) revealed that mastering basic skills and procedures received "some" emphasis in the intended mathematics curriculum only in Hong Kong, Quebec and Israel; in all other countries, this aspect was reported as receiving "a lot" of emphasis.

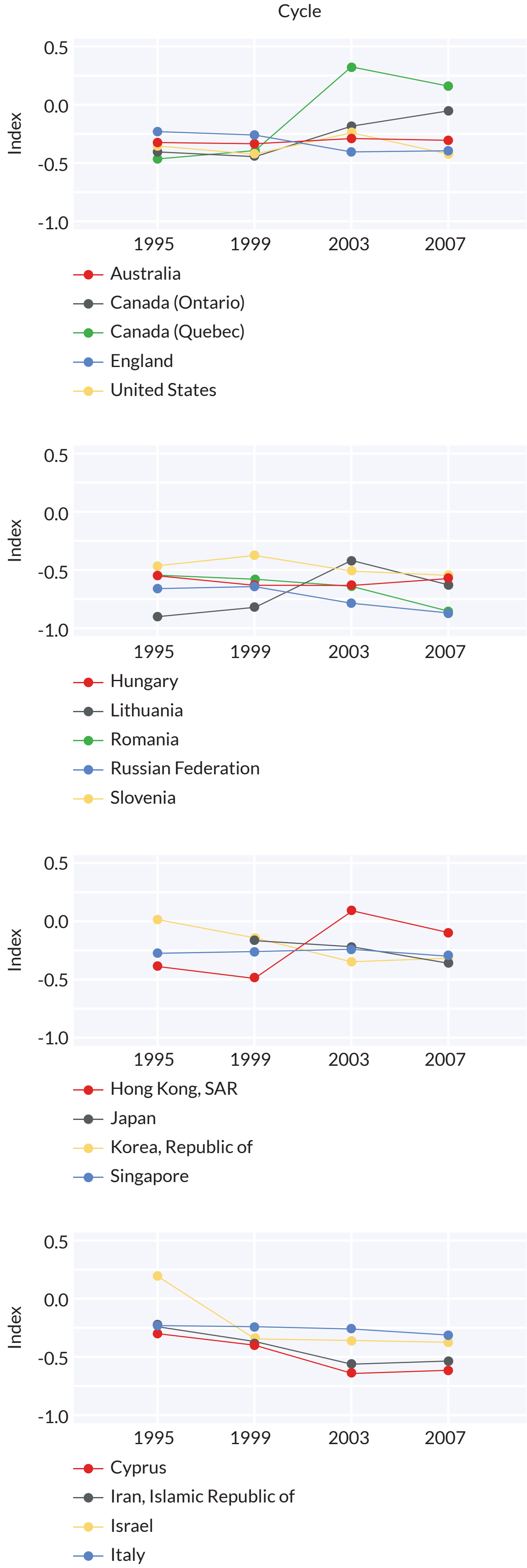
Thus, the move towards more reasoning, problem solving, and higher-order thinking in mathematics education does not seem to be a universal trend from 1995–2007, as would be expected based on publications on mathematics didactics. In some regions there was even an opposite trend during 2003–2007.

Table 4: Effect size of the change in frequency of using different didactical approaches between successive TIMSS cycles, as reported by teachers

Country	**Working on problems without obvious solutions**			**Practice computational skills**		
	1995–1999	**1999–2003**	**2003–2007**	**1995–1999**	**1999–2003**	**2003–2007**
Australia	0.02*	0.25*	−0.08*	0.03*	0.13*	0.02*
Canada (Ontario)	−0.04*	0.20*	0.08*	0.01	−0.01	−0.04*
Canada (Quebec)	0.11*	0.34*	−0.09*	0.01	−0.20*	0.05*
Cyprus	−0.07*	0.20*	0.16*	−0.01	0.41*	0.06*
England	0.03*	0.02	0.02	0.06*	0.25*	0.01
Hong Kong, SAR	0.19*	0.14*	−0.07*	0.22*	−0.48*	0.19*
Hungary	0.11*	−0.02	−0.04*	0.15*	0.10*	−0.07*
Iran, Islamic Republic of	0.06*	0.07*	−0.04*	0.24*	0.20*	0.00
Israel	−0.56*	0.12*	−0.02	0.16*	0.14*	0.01
Italy	0.21*	0.02	−0.01	0.10*	0.04*	0.06*
Japan	0.17*	0.01	−0.20*	–	0.05*	0.02*
Korea, Republic of	0.06*	0.11*	−0.12*	0.24*	0.33	−0.10*
Lithuania	0.13*	0.23*	−0.28*	−0.02	−0.22	0.08*
Romania	−0.13*	0.18*	−0.15*	−0.04*	0.29	0.11*
Russian Federation	0.15*	−0.10*	−0.02	0.04*	0.14*	0.09*
Singapore	0.09*	0.10*	−0.09*	0.06*	0.05*	0.01
Slovenia	0.03*	−0.04*	0.08*	−0.05*	0.09*	0.07*
United States	0.10*	0.14*	−0.07*	0.10*	−0.02*	0.11*
Median	0.08	0.11	-0.06	0.06	0.09	0.04
Number of countries with significant increase	14	12	3	11	13	11
Number of countries with significant decrease	4	2	11	2	4	3

Notes: *Statistically significant results ($p < 0.05$). Countries with consistent increases are marked in blue. No data were available for TIMSS 1995 for Japan.

Figure 1. Trends in the index working on problems versus practicing computational skills for the selected countries over the four TIMSS cycles



To investigate this further, we calculated an index that represented the frequency of working on problems in relation to the frequency of practicing computational skills. The ratio was calculated on an individual participant level and log transformed. The reported values represent the averaged log transformed ratio of the responses to these two items for a country in each TIMSS cycle. An index larger than 0 means that working on problems was more frequently used than practicing computational skills. If the index was smaller than 0, this indicates that practicing computational skills was more frequent than working on problems. According to our third hypothesis, we expected this index to increase over time. However, the hypothesis was largely rejected.

We found that, among all groups, the Central and Eastern European countries were the most computationally oriented (Figure 1). This situation remained unchanged over all four cycles of TIMSS. Russia and Romania demonstrated an increasing preference for computational skills; Lithuania initially moved towards problem solving, but the trend reversed in 2003–2007.

English-speaking and East Asian countries showed considerable variation. The expected trend towards a preference for problem solving was clearly visible in both the Canadian provinces and in Hong Kong. These three education systems, as well as Lithuania, saw a strong impetus towards problem solving between 1999 and 2003, while afterwards (2003–2007) there was less growth (see Ontario) or even a slight decline. In 2007, Quebec was the only education system where problem solving was more frequent than practicing computational skills.

In Cyprus, Iran, Korea, and most notably in Israel, the trend points towards less problem solving in comparison to practicing computational skills. The cases of Israel and Korea merit particular study. In both countries, teachers reported low levels of practicing for computation in 1995, with large increases over the next eight years. For Israel, the reported switch from "conceptual" to "computational" practices may be due to a change in target populations (Zuzovsky, 2008)[4]. In Korea, several curriculum reforms were rapidly introduced between the 1980s and the early 2000s (Lee, 2013), which might have caused confusion in classroom practices.

4 From 1999 on, Israel included Arabic-language schools in the target population for international studies. There, computational practices were both more popular and more effective than in Hebrew schools, as Zuzovsky (2008) reported. See also http://www.tau.ac.il/~danib/articles/0b-Zuzovsky.htm, retrieved on 29 October 2016.

CONCLUSIONS AND POLICY IMPLICATIONS

Between 1995 and 2007, reports from teachers and students as part of TIMSS documented considerable change in instructional practices at the country level. Among 367 pairwise comparisons (seven instructional practice items, three time intervals, and 17 or 18 countries), we identified 291 cases of significant change in reported frequency: 180 were positive (indicating increased use of a specific teaching practice), and 111 were negative. This indicates that national cultures of classroom teaching are not fixed and change over time. Thus, classroom teaching practices may be shaped by policies or changes in professional norms on a national level. However, the analysis presented here does not allow for any statement on what caused these changes.

Contrary to our expectation, there was limited support for universal change driven by "mega-trends" in educational research and policy. The use of tests in classrooms turned out to show relatively small counts of significant change in both directions, contrary to our expectation of a global rise in assessment-based teaching. Likewise, changes in subject-matter didactics, in our case: from computational to problem-solving approaches in mathematics education, were visible in a few countries only. The supposed trend was clearly visible in Canada, Hong Kong and Lithuania, but some countries demonstrated slight changes in the opposite direction.

Changes were not consistent across time periods. This is most obvious for constructivist practices such as relating mathematics to daily lives, and working in small groups. For these practices, we see mixed findings between 1995 and 2003. Finally, between 2003 and 2007, there was a "universal" boost in constructivist methods. The change towards more collaborative learning was most distinctive in East Asian countries. Probably, "mega-trends" in education, such as constructivism, require several years of debate and implementation before revealing measurable impact on everyday classroom practices.

The findings in this brief support a quest for the systematic development of teaching practices, classroom cultures, and teacher quality. To some extent, instruction seems to be shaped on the level of a country, a system, or a culture. Universal "mega-trends" seem to have limited impact. However, further research, based on international studies, is needed to answer the following questions: (1) To what extent can national patterns of instructional practice be shaped by national policies, such as reforms in curricula, standards, and teacher education? (2) How and how strongly do such changes impact on student outcomes?

REFERENCES

Bayer, S., Klieme, E., & Jude, N. (2016). Assessment and evaluation in educational contexts. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing Contexts of Learning, An International Perspective* (pp. 471–490). Cham: Springer.

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L., & Smith, T.A. (1996) *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssandpirls.bc.edu/timss1995i/MathB.html

Corno, L. (1996). Homework is a complicated thing. *Educational Researcher, 25*, 27–30.

Ellington, A. (2015). The effects of calculators on students' achievement and attitude levels in K–12. In E.A. Silver, & P.A. Kenney (Eds.), *More Lessons Learned from Research, Volume 1: Useful and Usable Research Related to Core Mathematical Practices* (pp. 181–188). Washington, DC: National Council of Teachers of Mathematics.

Hattie, J. (2009). *Visible Learning*. New York: Routledge.

Lee, J. (2013). History of mathematics curriculum in Korea. In J. Kim, I. Han, M. Park., & J. Lee (Eds.), *Mathematics Education in Korea. Volume 1: Curricular and Teaching and Learning Practices*. New Jersey: World Scientific Publishing

Lumley, T. (2016). *Survey: analysis of complex survey samples*. R package version 3.31–2.

Lumley, T., & Scott, A. J. (2013). Two-sample rank tests under complex sampling. *Biometrika*, *100*(4), 831–842.

Mullis, I.V.S., Martin, M.O., & Loveless, T. (2016). *20 Years of TIMSS. International Trends in Mathematics and Science Achievement, Curriculum, and Instruction*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Olson, J.F., Berger, D.R., Milne, D., & Stanco, G.M. (2008). *TIMSS 2007 Encyclopedia: A Guide to Mathematics and Science Education Around the World, Volume 1*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

NCTM. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, Virginia: National Council of Teachers for Mathematics.

Nilsen, T. & Gustafsson, J. E. (Eds.) (2016). *Teacher Quality, Instructional Quality and Student Outcomes*. IEA Research for Education, Volume 2. Cham: Springer International Publishing.

Schoenfeld, A.H. (2006). Mathematics teaching and learning. In P.A. Alexander, & I.H. Winne (Eds.), *Second Handbook of Educational Psychology*. London: Routledge.

Schwerdt, G., & Wuppermann, A.C. (2009). *Is Traditional Teaching really all that Bad? A Within-Student Between-Subject Approach.* Munich: CESifo Working Paper Series 2634, CESifo Group.

Seidel, T., & Shavelson, R.J. (2007). Teaching effectiveness research in the last decade: Role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454–499.

Teltemann, J., & Klieme, E. (2016). The impact of international testing projects on policy and practice. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (pp. 369–386). New York: Routledge.

Tobias, S., & Duffy, T.M. (Eds.) (2009). *Constructivist Instruction – Success or Failure?* New York: Taylor & Francis.

Zuzovsky, R. (2008). Capturing the dynamics that led to the narrowing achievement gap between Hebrew-speaking and Arabic-speaking schools in Israel: Findings from TIMSS 1999 and 2003. *Educational Research and Evaluation*, *14*(1), 47–71.

Zuzovsky, R. (2013). What works where? The relationship between instructional variables and schools' mean scores in mathematics and science in low-, medium-, and high-achieving countries. *Large-scale Assessments in Education, 1:2.* doi:10.1186/2196-0739-1-2. Retrieved from http://link.springer.com/article/10.1186/2196-0739-1-2.

IEA POLICY BRIEF

ABOUT THE IEA

The International Association for the Evaluation of Educational Achievement, known as the IEA, is an independent, international consortium of national research institutions and governmental agencies, with headquarters in Amsterdam. Its primary purpose is to conduct large-scale comparative studies of educational achievement with the the aim of gaining more in-depth understanding of the effects of policies and practices within and across systems of education.

ISSN: 2215-0196
Photo credits:
IEA
Copies of this publication can be obtained from:
IEA Amsterdam
Keizersgracht 311
1016 EE Amsterdam
The Netherlands
By email: secretariat@iea.nl
Website: www.iea.nl
Follow us @iea_education

Editors of the policy brief

Tom Loveless
Brookings Institute

David Rutkowski
Centre for Educational Measurement at the University of Oslo (CEMO)

Please cite this publication as:
Rožman, M., & Klieme, E. (2017, February). Exploring cross-national changes in instructional practices: evidence from four cycles of TIMSS (Policy Brief No. 13). Amsterdam, The Netherlands: IEA.