The roles of teaching styles and curriculum in Mathematics achievement: Analysis of TIMSS 2011

Nadir Zanini and **Tom Benton** Research Division

Introduction

In these times of rapid change fuelled by technological advances, the demand for improved mathematical knowledge is growing worldwide. Mathematical skills such as problem solving and inference are increasingly becoming part of both university access and labour market requirements. In some countries such as the United Kingdom, the importance of Mathematics has been recognised by policy-makers for individual progression, as well as for the economy and society at large (Oates, 2010; Department for Education [DfE], 2010). Evidence from the UK as well as comparative studies recognises Mathematics as a key subject (see, among others, Andrews, 2014). In particular, a wealth of research based on international benchmarking surveys, such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA), have investigated the connection between achievement in specific subjects and contextual factors (such as student and family background, school inputs and institutional background)[1]. However, much less is known about how aspects of a Mathematics curriculum are associated with Mathematics achievement.

This article provides empirical evidence about the link between Mathematics achievement, curriculum, teaching methods and resources used in the classroom. More specifically, this research explores common teaching styles and topics taught across countries with respect to their Mathematics achievement.[2] In order to do so, we make use of the fifth TIMSS survey, which provides a rich set of information regarding aspects of the curriculum (e.g., the emphasis on problem solving and interpreting data sets), resources used by teachers in the classroom (e.g., calculators and textbooks) and teaching styles (e.g., how often students are asked to take written tests, to work out problems individually rather than with teachers' guidance), along with measures of achievement in Mathematics gathered in 2011. Although TIMSS is administered to students and their teachers in both Grades 4 and 8 (Years 5 and 9 respectively, within England), analysis in this research is restricted to the Grade 8 students (aged 14). When analysing data aggregated at jurisdictional level, this allows us to explore relationships in the Mathematics achievement of 15 year-olds as measured by PISA 2012.

There are two main elements which connect with achievement that are investigated in this article. The first is the relationship between Mathematics performance and the prevalence of different teaching methods (also referred to as teaching styles or instructional practices) within different countries. It should be noted that in undertaking this analysis we are not assuming that countries will necessarily limit themselves to one dominant teaching style. Indeed, even at the level of the individual classroom we are aware that the same teacher will use a variety of styles and methods of teaching within the same lesson (Boaler 2008, Benbow & Faulkner 2008). This is likely to be even truer at the level of whole nations where different teachers within different schools are likely to have a different emphasis within their own style of teaching. The second element is the role of the curriculum in Mathematics performance measured by international assessment tests, such as TIMSS and PISA. In exploring the Mathematics curriculum, it is important to consider that it is possible that teachers do not teach the topics that the National Curriculum requires (Andrews, 2011). To overcome this issue, this research looks into the curriculum actually adopted by teachers and measured by means of their responses to a set of items which summarise the Mathematics topics they have taught in their classrooms up to Grade 8.

The underlying idea of the research in this article is to use international data from TIMSS to inform the debate about which methods of teaching Mathematics, which resources used in the classroom, and which topics taught may be most effective in terms of achievement in Mathematics measured by international tests. Since the 1990s, these tests have depicted East Asian students outperforming their Western counterparts (for the most recent results, see Mullis, Martin, Foy & Arora, 2012: Exhibit 1.1; Organisation for Economic Co-operation and Development [OECD], 2013: Figure I.2.14). This has led to the desire among policy-makers and education practitioners in countries such as the United States and the UK to emulate practices in use in high-performing jurisdictions, such as Singapore. As argued by Usiskin (2012), at both policy-level and school-level, the curriculum and the textbooks used have been identified as the main determinant of Singapore's Mathematics performance without any justification. Our aim is to check whether Mathematics achievement is linked to the curriculum and the use of textbooks as well as other resources and particular teaching styles.

In carrying out this analysis it is important to consider that a country's characteristics (such as geographical and economic conditions, and aspects of the education system) can have a strong influence on students' Mathematics achievement (Andrews, 2012; Usiskin, 2012). Therefore, drawing conclusions from such comparisons may be misleading (Gill & Benton, 2013). To help avoid this problem, as far as possible, the characteristics of students in each country that may affect achievement are allowed for in the analysis presented.

The results from the analysis should not be taken as a suggestion that the particular teaching methods, and the specific topics taught within high-performing jurisdictions, should be adopted elsewhere. Rather, we accept that establishing the causal factors behind attainment levels of different countries is problematic and that, at best, these studies can be used as a mirror to reflect upon teaching practices within our own country rather than a blueprint (Clarke, 2004; Elliott 2013). However, we feel that such reflections are best based upon detailed quantitative analysis of the type provided, rather than a simple 'eyeballing' of international league tables of achievement.

1. For a broad review of the contextual factors affecting achievement using data from international tests, see Hanushek and Woessmann (2011).
2. Results presented in this article are drawn from Zanini and Benton (2015) which also includes a more in-depth case study set within the UK.

Data

Our analysis makes use of data from TIMSS 2011[3]. The data focuses upon students in Grade 8 and their teachers. Data from a total of 42 countries participating in the TIMSS Mathematics assessments were included within the analysis. Four countries[4] which used the Grade 8 TIMSS instruments with students in grades other than Grade 8 were excluded. Non-country, benchmarking participants (such as individual states within Canada) were also excluded.

In some previous analyses of PISA (see Gill & Benton, 2013) the focus has deliberately been on countries identified as being sufficiently similar to each other. However, in this research we decided to take a different route as part of our aim was to attempt to group countries with regard to their curriculum and use of different teaching methods. For this element of analysis, the diversity of countries included was an advantage as cultural diversity may lead to differences in curriculum and teaching practice which may in turn be identifiable. On the other hand, including the entire, diverse array of countries within our analysis decreases the confidence with which we may be able to generalise any findings to the unique context within a single country.

Our analysis examines Mathematics achievement within different countries in two different ways. Firstly, achievement is defined using the overall Mathematics achievement scale as measured by TIMSS. This provides a measurement of students' achievement in relation to the kind of Mathematics curriculum that is generally taught internationally (Wu, 2009). It was also of interest to examine how the methods of teaching and the curricula within different countries might relate to students' abilities to apply Mathematics more broadly beyond the specific topics they have been taught. For this reason we also analysed the relationship between identified patterns of teaching from TIMSS, and Mathematics achievement as measured by PISA 2012 (published in OECD, 2013). PISA "does not just ascertain whether students can reproduce knowledge; it also examines how well students can extrapolate from what they have learned and apply that knowledge in unfamiliar settings." (OECD, 2013). As such, this provides a distinct concept of Mathematics achievement which was of interest within our research. Data on achievement in PISA 2012 was available for 29 of the 42 countries with relevant data for TIMSS 2011.

As we describe in the next section, in examining the relationship between Mathematics achievement and aspects of teaching, it was important to at least attempt to control for the impact of other background variables. This supplementary information was almost exclusively drawn from the school and student surveys collected as part of TIMSS 2011 itself. However, as pointed out by the OECD (2013):

> *The relative prosperity of some countries allows them to spend more on education, while other countries find themselves constrained by a lower national income. It is therefore important to keep the national income of countries in mind when comparing the performance of education systems across countries.* (OECD, 2013, p.24)

For this reason we also included data on the per capita Gross Domestic Product (GDP) of countries in 2011 in our analysis.[5]

3. This data is freely available for download from http://timss.bc.edu/timss2011/international-database.html. More technical details on questionnaires preparation, sampling and data collection are described in Martin and Mullis (2012).
4. Botswana, Honduras, South Africa and Yemen.
5. Additional data sources were used to retrieve GDP information for specific countries (see Zanini & Benton, 2015: Section 2).

Methods

Using latent class analysis to segment countries into groups

One approach to analysing the links between teachers' responses to questionnaire items and the achievement of students would be to consider each questionnaire item separately. However, assuming we wish to distinguish the effect of one teaching method from the effect of another, such an approach would immediately run into problems due to the number of questions we might wish to analyse coming close to, or possibly exceeding, the number of countries available for analysis. In order to estimate linear regression coefficients and to disentangle the effect of one teaching method from the effect of another, the number of variables considered must be less than the number of observations included in the analysis.

Given the above consideration, our analysis needed to condense the information contained in the questionnaire, avoid treating individual items as if they give a raw quantity of the way instructional time is used, and ensure that we examine the pattern of countries' responses across all items together. For this reason, our analysis focuses upon identifying groups of countries with similar patterns of answers to the questions of interest. This part of the analysis was achieved using *latent class analysis*.

In its theoretical formulation, latent class analysis attempts to explain the relationships between various measured variables in terms of respondents belonging to one of a number of discrete latent (or unmeasured) classes (or groups). A typical latent class analysis would assume that all of a respondent's questionnaire answers are independent of one another once we know the grouping they fit in to. In our own situation we are less concerned with these theoretical underpinnings than with using the technique as a convenient way to segment countries into groups with similar patterns of responses across various questionnaire items. Thus we are not assuming that any of the theoretical assumptions given as examples above are actually true. Instead, we use the associated software for the sake of convenience and can verify whether the method has been effective by examining whether the derived groupings of countries actually display relatively similar behaviour in terms of the questions of interest.

Analysis, was completed using *Mplus Version 7* (Muthén & Muthén, 2012). This software provides a convenient feature whereby data is analysed at teacher level and groupings are produced at country level. Separate latent class analyses were undertaken to produce groupings of countries in terms of:

1. How frequently students were required to do various tasks during Mathematics lessons and the importance of various types of teaching materials for instruction.
2. When and whether various Mathematics topics were taught to Grade 8 students.

The second of these analyses focuses on what is taught in different countries whilst the first focuses on some aspects of how it is taught. The groupings of countries derived for each of these two research areas are described further in the Results section.

Using meta-regression to account for the impact of background variables on achievement

Once country groupings were produced, we examined how these groupings related to students' achievement. In order to explore this it is necessary to account for the influence of background characteristics.

One possible method was to account for each country's GDP per capita in 2011. However, initial analysis revealed a correlation of only 0.45 between GDP per capita and achievement in TIMSS 2011, rising to only 0.60 once the country of Qatar was removed.

For this reason we instead attempted to aggregate the information collected within the background section of the TIMSS student questionnaire to derive a single number for each country capturing the majority of the important information from the background questionnaire. This was achieved using the following process. Firstly the full international student-level data set was restricted to those students with listwise complete data on the following background questions:

- Gender
- How often they speak the language of the test at home
- How many books they have in their home
- Whether they have a computer at home
- Whether they have a study table/desk at home
- Whether they have books of their own (not school books) at home
- Whether they have their own room at home
- Whether they have an internet connection at home
- Highest level of education completed by their mother ('Don't know' was a valid response)
- Highest level of education completed by their father ('Don't know' was a valid response)
- Whether their mother was born in the country they are living in
- Whether their father was born in the country they are living in
- When they moved to the country they are living in (Being born there was a valid response).

A student level regression was then performed using Mathematics achievement data upon each of the above background characteristics. Using the results of this regression, for each student in the full data set (i.e., including those with listwise *incomplete* data) their response to each of the above questions was replaced with the corresponding regression coefficient. For each of the questions above, the average level of these effects was calculated for each single country. For each country in turn, these average effects were added up across questions to generate an overall, aggregated measure of background characteristics[6]. This measure was linearly rescaled so that the country with the highest measure (Korea) was assigned 100 and the country with the lowest measure (Ghana) was assigned 0.[7] Unsurprisingly, given the way it was derived, this measure was found to have a very high correlation with achievement in TIMSS 2011 Mathematics (correlation=0.80). More encouragingly, this measure was also found to correlate very highly with countries' Mathematics achievement in PISA 2012, which was not used to help create the scale (correlation=0.72). As a result, it is clear that this measure provides a useful mechanism to help control for the differences between countries.

Once the groupings of countries had been derived, it was possible to analyse the relationship between these groupings and Mathematics achievement whilst taking account of the aggregated background measure and GDP per capita using meta-regression. In essence meta-regressions are simply country-level regressions of achievement on country groupings, the background measure and GDP per capita. However, in contrast to standard linear regression, meta-regression allows us to account for the fact that the outcome (in this case Mathematics achievement) is measured with error. As such, the technique allows for the fact that some of the variation between countries will be purely due to measurement error and that the magnitude of this error may vary between countries. For our analysis, the technique also allowed for the possibility that there may be variation between countries that is neither due to measurement error nor explicable in terms of the covariates included within the regressions.[8]

Results

Teaching styles: grouping and meta-regressions

To begin with, our analysis examined countries in terms of the teaching styles used in lessons and the resources used to support learning; that is, questions 19 and 20 from the TIMSS 2011 Mathematics teacher questionnaire. The first question asked teachers to report how often ('every or almost every lesson', 'about half the lessons', 'some lessons', and 'never') they ask students to do specific activities. The second question asked teachers whether they use a list of resources (textbooks, workbooks/worksheets, concrete objects that help, or computer software) and, if so, if they use it as a 'basis for instruction' or as a 'supplement' (see Appendix).

As we have described, the analysis was undertaken using latent class analysis with the aim of identifying a small number of country groupings where, within each group, the extent of teachers' reported use of various strategies and resources was relatively similar. The latent class analysis identified five groups of countries. The choice of the number of classes was driven partly by statistical indices such as the Bayesian Information Criterion (BIC), (see Nylund, Asparouhov & Muthén, 2007), and partly by a desire to ensure that each group contained a reasonable number of countries.[9]

A summary of the groupings, the countries they contain, an overall interpretation of the response patterns and a suggested label for each grouping are shown in Table 1. Briefly this suggests that the five groupings are:

1. *Simplified*. A set of mainly Nordic countries where teachers are the least likely to report using many of the techniques in every, or almost every, lesson.
2. *Learn, repeat, and check*. A set of East Asian and former Soviet countries with high frequencies of teacher demonstration and independent or routine work, but lower frequencies for explaining answers or relating to daily lives.
3. *Routine independence*. A set of English speaking countries with fairly low frequencies of teacher demonstrations and more frequent use of independent or group work.
4. *Restrained diversity*. A set of mainly European countries with a high frequency use of a range of techniques, but with a relatively high reliance on textbooks as the basis for instruction, and fairly infrequent use of written tests and quizzes.

6. The questionnaire in Israel did not include the question about whether students had a room of their own. For this reason, Israel was assigned the average effect (across countries) for this question.
7. The aggregated background measure assigned to each country is provided in Zanini and Benton (2015) Table 2.
8. For further details on meta-regression see Benton (2014). The meta-regression analyses were completed in R using the package metafor (Viechtbauer, 2010).
9. Further analysis aimed at validating the grouping revealed that the simple five class segmentation of countries captures nearly half of the overall variation in scores and that this was hardly improved by increasing the number of classes to six (see Zanini & Benton, 2015: Section 3.3).

5. *Test-centric diversity.* A diverse group of countries although many of them (10 out of 17) are Middle Eastern. Teachers in these countries report high frequency use of many different techniques, particularly tests and quizzes, indicating that a variety of methods may be used within the classroom.

Table 1 shows that many countries that share similar contextual (i.e., geographical, historical, societal and/or cultural) factors are grouped together. It is both reassuring and disappointing that the groupings of countries match relatively well with simple contextual descriptions of the countries. It is reassuring in that it is to be expected that countries with similar geographical, historical, societal and/or cultural backgrounds might be expected to share similar styles of teaching. However, this also means that the groupings in terms of teaching style are also strongly interrelated with other (possibly unmeasured) contextual factors. This means that any attempt to examine the impact of the different teaching styles over and above these other factors is problematic.

The performances in TIMSS 2011 and PISA 2012 of each country

Table 1: Descriptions of country groupings by teaching styles

Grouping	*Countries included*	*Notes*	*Label*
1	*Mainly Nordic:* Finland, Norway, Sweden, Slovenia	Teachers in these countries are least likely to indicate using any particular method 'every or almost every lesson'. This indicates that students are less likely to experience multiple teaching methods in every lesson. Lower frequency of written tests or quizzes being used than in other countries. These countries are also the most likely to use textbooks as the basis for instruction.	*Simplified*
2	*East Asian and Former Soviet:* Chinese Taipei, Hong Kong, Japan, Kazakhstan, Korea, Malaysia, Russian Federation, Ukraine	High frequency of listening to the teacher and watching demonstrations of problem solving. Also high frequency of independent working and applying procedures to solve routine problems. Lower frequencies than other countries for explaining answers, relating to learning to daily lives or finding own solutions to complex problems. Testing is used fairly frequently; possibly to check progress.	*Learn, repeat, and check*
3	*English Speaking:* Australia, New Zealand, Singapore, United States, England	Along with Group 1, these countries have the least frequent amount of listening to the teacher explain how to solve problems. Memorisation of rules, procedures and facts is also relatively low. High frequency of independent/group work. The lesser use of textbooks as the basis for instruction in these countries indicates that a variety of resources may be being used. High frequency of time addressing routine problems and comparatively little spent on complex problem solving. Written assessments and quizzes are used relatively infrequently.	*Routine independence*
4	*Mainly other European:* Armenia, Georgia, Hungary, Israel, Italy, Jordan, Lithuania, Macedonia	Similar to the final category but less extreme. However, a greater reliance on textbooks and less use of assessment in the form of quizzes and tests.	*Restrained diversity*
5	*Scattered geographical regions:* Bahrain, Chile, Palestinian National Authority, Ghana, Indonesia, Iran, Lebanon, Morocco, Oman, Qatar, Romania, Saudi Arabia, Syrian Arab Emirates, Tunisia, Turkey	For many techniques teachers were the most likely to indicate that they will incorporate these techniques into 'every or almost every lesson' indicating an intention from teachers to incorporate a range of techniques and resources in most lessons. It includes more than a third of teachers stating that they will use written tests or quizzes in 'every or almost every lesson' (and two-thirds in at least half of lessons). This group of countries are the second least likely to indicate that textbooks are used as the basis of instruction indicating that a range of materials are being used.	*Test-centric diversity*

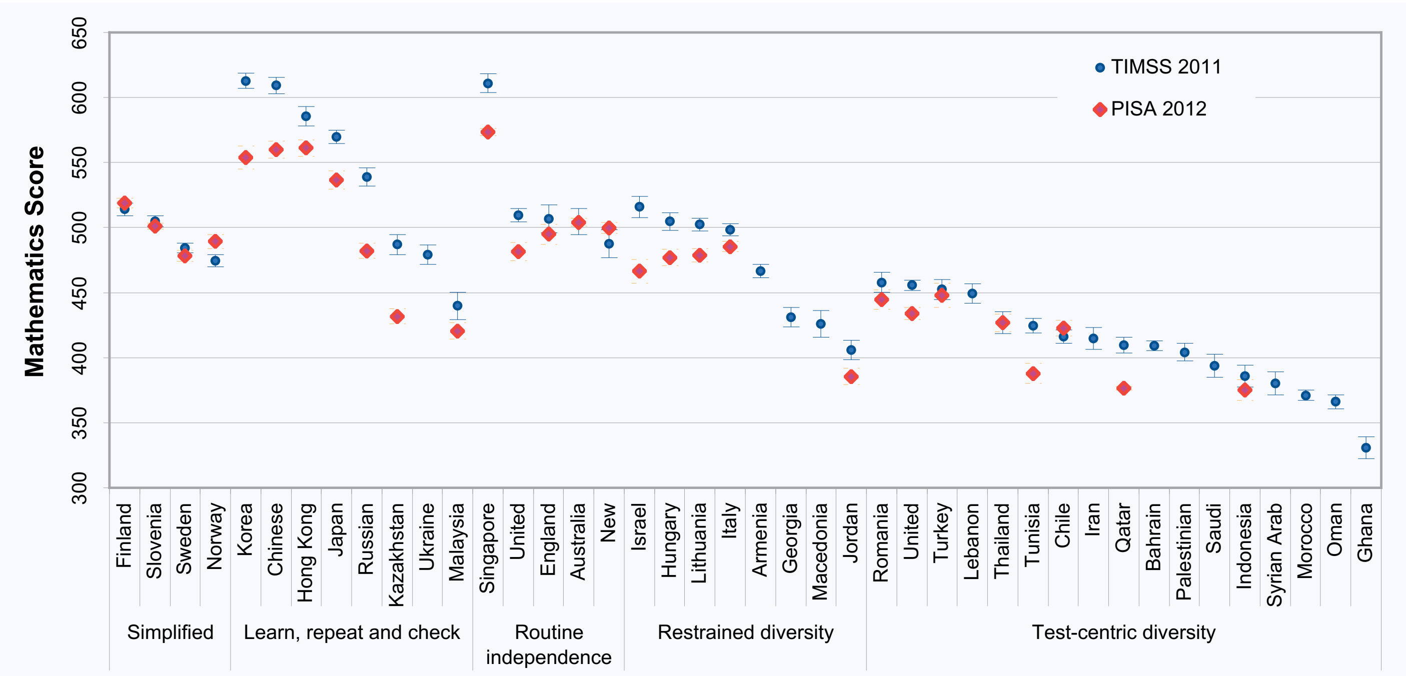

Figure 1: TIMSS 2011 Grade 8 and PISA 2012 Mathematics performance by countries within each teaching style group

within each grouping are shown in Figure 1. As can be seen, there is little obvious difference in the average performance of the different groups of countries with the exception of the generally lower performance of countries in the *Test-centric diversity* teaching category. However, this may itself be explained by the contextual nature of the countries within this grouping rather than the effect of this particular teaching style.

A more interesting picture emerges when we examine the difference between performance in PISA and performance in TIMSS. In order to explore this, performance in PISA is plotted against performance in TIMSS in Figure 2. Countries in each of the different groupings are identified separately and a regression line showing countries' expected performance in PISA 2012 given their performance in TIMSS 2011 is included. As can be seen, countries in the *Learn, repeat, and check* category tend to perform worse on PISA than might be expected from their performance in TIMSS. In contrast, countries in the *Simplified* category tend to perform better in PISA than might be expected from their TIMSS results. Such results might be explicable in terms of a *Learn, repeat, and check* approach being helpful in terms of learning the Mathematics associated with a particular curriculum but less helpful in enabling such knowledge to be applied in new situations. In contrast a *Simplified* approach may potentially allow teachers to provide more depth in their instruction so that students can understand how a particular piece of Mathematics may be applied in numerous situations.

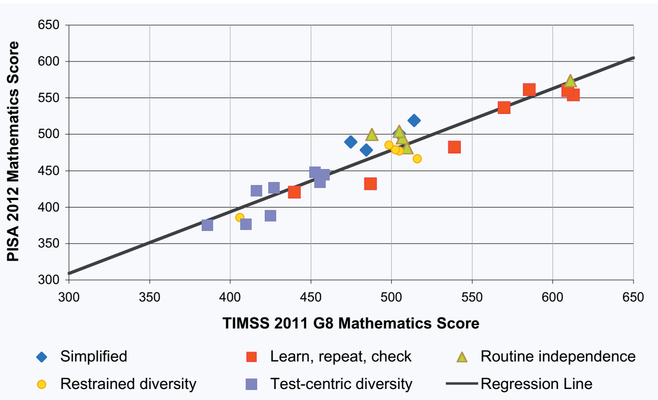

Figure 2: PISA performance against TIMSS performance for countries with different reported approaches to teaching Mathematics

The results in the figures do not account for the influence of other background factors upon results. In order to do this, performance in TIMSS is plotted against the newly derived aggregated background measure provided in Figure 3. Countries with different approaches to teaching are identified separately. A regression line is added to this graph showing the expected level of achievement in TIMSS for countries with different levels of the aggregated background measure. This graph suggests that countries with a *Learn, repeat, and check* approach to teaching tend to over-perform relative to their background characteristics (seven out of eight countries are above the line). In contrast countries with a *Simplified* or *Restrained diversity* approach to teaching tend to under-perform (four out of four and eight out of eight below the line respectively – albeit only slightly). Having said this, given the cultural differences between the groups of countries noted earlier, it is difficult to be confident that these results are due to the impact of teaching styles.

Performance in PISA is plotted against the aggregated background measure in Figure 4. When Mathematics attainment is quantified using this measure, differences between the groups of countries are far less evident. Countries with a *Restrained diversity* approach are still universally below the line of expected performance, but only slightly so. All other groups of countries are spread both above and below the line. This suggests that particular teaching styles may favour performance in TIMSS over performance in PISA (as seen in Figure 2), although once again this finding may be confounded by other cultural or societal factors.

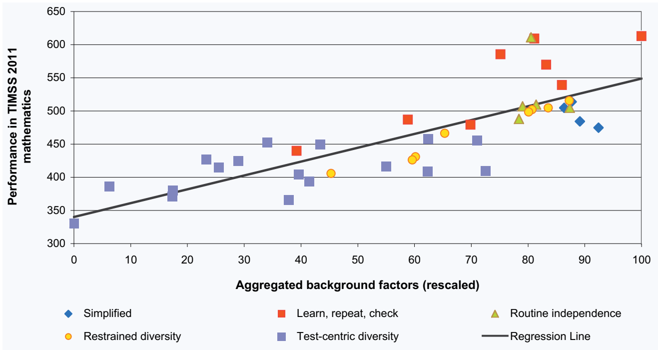

Figure 3: Performance in TIMSS by countries' aggregated background measures with different reported approaches to teaching Mathematics

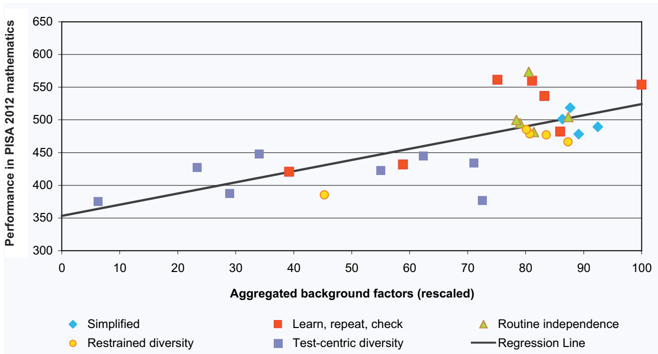

Figure 4: Performance in PISA by countries' aggregated background measures with different reported approaches to teaching Mathematics

Temporarily setting aside concerns over the potential confounding influence of other unmeasured variables, we further explore the statistical significance of the relationships suggested in Figures 2, 3 and 4 using meta-regression. Findings of meta-regression are shown in Table 2. The first set of results confirms that countries with a *Learn, repeat, and check* style of teaching tend to significantly outperform similar countries with other styles in TIMSS 2011. Specifically, this style of teaching is associated with countries achieving an additional 69 points on average compared to similar countries with a test-centric diverse approach. Other teaching styles do not appear to be associated with significantly greater performance, although countries within the *Routine independence* group come close.

The second set of results in Table 2 shows that the above results are not repeated within PISA 2012. Although countries within the *Learn, repeat, and check* group tend to outperform other similar countries within the diverse group, the extent of the difference is smaller and is no longer statistically significant. It should also be noted that the difference between this group and the *Simplified* group is substantially smaller than for the analysis of outcomes from TIMSS. Furthermore, the achievement of the *Learn, repeat and check* group is now slightly behind the achievement of similar countries in the *Routine independence* group, which may be largely attributed to the influence of Singapore on the latter group.

The meta-regression of performance in PISA whilst accounting for performance in TIMSS (and other background variables) confirms that

Table 2: **Results of meta-regression examining the relationships between Mathematics performance and teaching style**

	Performance in TIMSS 2011 (1)			*Performance in PISA 2012 (2)*			*Performance in PISA 2012 (3)*		
	Coeff.	*S.E.*	*P-val.*	*Coeff.*	*S.E.*	*P-val.*	*Coeff.*	*S.E.*	*P-val.*
Intercept	346.1	15.0	<.0001	356.0	22.4	<.0001	36.4	33.2	0.2850
Aggregated background measure	1.8	0.4	0.0002	1.5	0.5	0.0072	-0.1	0.3	0.8184
Per capita GDP	-0.2	0.4	0.6280	-0.4	0.5	0.3800	-0.1	0.2	0.7309
Country group (Ref.: Test-centric diversity):									
Simplified	1.9	25.2	0.9419	24.7	28.7	0.3996	28.5	12.2	0.0295
Learn, repeat & check	68.9	20.1	0.0016	47.1	23.8	0.0607	-12.8	11.8	0.2899
Routine independence	45.2	22.6	0.0535	51.0	25.6	0.0586	15.9	11.4	0.1795
Restrained diversity	2.6	20.4	0.9005	-3.8	26.4	0.8874	-4.0	11.3	0.7270
TIMMS 2011 score	-	-	-	-	-	-	0.9	0.1	<.0001

countries with a *Simplified* teaching style significantly over-performed in PISA 2012 relative to their performance in TIMSS 2011. Specifically they tended to achieve almost 30 points higher than would be expected in PISA given their performance in TIMSS and the background characteristics of their students. None of the other groups displayed any significant difference in their PISA performance once performance in TIMSS was accounted for.

Taken together, the results above suggest that particular styles of teaching may be more beneficial for the particular focus of TIMSS and that others may be more beneficial in terms of PISA. Specifically the results support the idea that a *Learn, repeat, and check* style may be helpful in boosting performance against curriculum-related Mathematics questions but that this advantage may not translate to the ability of students to apply Mathematics to real life situations. In contrast the *Simplified* approach appears to be disadvantageous (compared to certain other teaching styles) when assessed using a TIMSS style assessment, but this disadvantage disappears if students are assessed on their ability to apply their knowledge.

Although interesting, the results above must be treated with caution. As we have mentioned, the groupings by teaching style are closely related to particular geographical areas. As such the possibility that there may be other confounding cultural or societal factors influencing the results cannot be understated. For this reason, the results we present here should be seen as interesting results for reflection rather than definitive proof that particular styles of teaching can boost achievement in one or other of the international tests.

Curriculum: grouping and meta-regressions

Having examined teaching styles and resources employed by teachers in their lessons, we now turn our attention to the curriculum actually taught across countries as reported by teachers in question 23 of the TIMSS 2011 Mathematics questionnaire[10]. In the questionnaire, topics were grouped into four different domains: 'Number', 'Algebra', 'Geometry' and 'Data and Chance'. For each topic the possible answers were 'Mostly taught before this year', 'Mostly taught this year' or 'Not yet taught' (see Appendix). From an inspection of the data (see Zanini & Benton, 2015: Table 8) it emerges that most of the topics under the 'Number' domain seemed to be taught before Grade 8. Also, it is quite clear that some topics such as simultaneous equations and those linked to geometrical representations of shapes are considered quite advanced at international level, as they had not yet been taught by most of the teachers.

In order to identify a small number of country groupings where the topics taught by teachers were relatively similar across countries within each group, a latent class analysis was performed. Following exactly the same procedure employed for teaching styles, the latent class analysis on the curriculum identified five groups of countries. The choice of the number of classes was driven partly by the BIC statistical index and partly by the need to ensure that each group contained a reasonable number of countries. As the 'five-group solution' proved to fit the data better than any 'one to four classes' model and that increasing to six or seven classes led to very small groups (of which one constituted of one country), the five-class segmentation was adopted (for details see Zanini & Benton, 2015: Table 9).

From the careful inspection of the results of the grouping analysis it is possible to derive that the five groups are:

1. *Number and Algebra (N&A):* the smallest group, (only comprising of Chinese Taipei and Ukraine), where the Mathematics curriculum up to the year of interview was mainly focused on 'Number' and 'Algebra';
2. *Delayed introduction 1 (DI1):* a group of countries characterised for having delayed the introduction of most topics relative to other countries, even those fundamental to a Mathematics curriculum;
3. *Non-algebraic focus (NAF):* teachers in these countries were the least likely to report having already taught topics in 'Algebra' which are usually considered basics of the Mathematics curriculum;
4. *Delayed introduction 2 (DI2):* as for Group 2, it seems that most of the Mathematics curriculum was delayed relative to countries in other groups;
5. *Geometry and Data (G&D):* when compared to other countries, the Mathematics curriculum in these countries seems to be more focussed on 'Geometry' and 'Data and Chance', of which some topics can be quite advanced for Grade 8 students.

A more detailed summary of the response patterns and the list of countries in each grouping are shown in Table 3. In contrast to the results shown for teaching styles, groupings of countries by mathematical curriculum do not match with any simple ways of describing them. Even when countries sharing similar cultural backgrounds or from the same geographical area are in the same group, they are also mixed up with

10. Kazakhstan and Russian Federation are not considered in the analysis in this section because of the high number of missing values in the teachers' responses to the items in the TIMSS questionnaire related to the curriculum.

Table 3: Descriptions of country groupings by curriculum

Grouping	*Countries included*	*Notes*	*Label*
1	Chinese Taipei, Ukraine	In these countries teachers reported that 'Number' and 'Algebra' topics have been taught 'mostly before this year'. Teachers in these countries were among the least likely to report teaching topics in the other domains 'before this year' and among the most likely to indicate having 'not yet taught' these topics. This clearly suggests that for countries in this group the Mathematics curriculum up to the year of interview were mainly focused on Number' and 'Algebra'.	*Number and Algebra (N&A)*
2	Palestinian National Authority, Lebanon, Morocco, Norway, Sweden, Syrian Arab Republic, Tunisia	Teachers in these countries were among those least likely to have taught a number of topics under different domains 'mostly before this year'. When compared to other countries, not many Mathematics topics were introduced before Grade 8.	*Delayed introduction 1 (DI1)*
3	Finland, Georgia, Indonesia, Iran, Italy, Lithuania, Slovenia	Teachers in these countries were among the most likely to indicate having taught topics in the 'Number' and (partially) 'Geometry' domains 'before this year', but the least likely to report having already taught 'Algebra' topics, most of which are considered basics of a Mathematics curriculum.	*Non-algebraic focus (NAF)*
4	Australia, Chile, Ghana, Malaysia, New Zealand, Qatar, Saudi Arabia, Thailand, United Arab Emirates, England	All Mathematics topics were less likely to be reported to be taught by teachers 'mostly before this year'. This is particularly true in the domain of 'Number'. Along with Group 2, it seems that most of the introduction of the Mathematics curriculum was delayed with respect to countries in other groups.	*Delayed introduction 2 (DI2)*
5	Bahrain, Armenia, Hong Kong, Hungary, Israel, Japan, Jordan, Korea, Oman, Romania, Singapore, Turkey, Macedonia, United States	Teachers in these countries were the most likely to indicate having 'taught before this year' most of the topics related to the domains of 'Geometry' and 'Data and Chance'.	*Geometry and Data (G&D)*

countries with very different contextual factors. For example, it can be seen that in Group 2, Norway and Sweden are grouped with Morocco and Tunisia among others, but not with Finland which is grouped in Group 3 with Italy and Iran. This can be interpreted as an indication that the Mathematics curricula may be less connected to societal and contextual factors than teaching styles.

Figure 5 shows the performances in TIMSS 2011 and PISA 2012 respectively of each country within each grouping. It is straightforward to see that no striking difference in the average performance of the different groups of countries arises. However, it should be noted that the highest performance in both TIMSS and PISA tests was achieved by four countries (Korea, Singapore, Hong Kong and Japan) in Group 5 (i.e., those with the label 'Geometry and Data'), with only one exception, Chinese Taipei. In interpreting this evidence it has to be considered that the societal and contextual factors within these countries can be a confounding factor of the impact of the curriculum on performance. Similarly, no clear patterns arise from the analysis of the relationship between performance in PISA and performance in TIMSS, indicating that a grouping's performance in

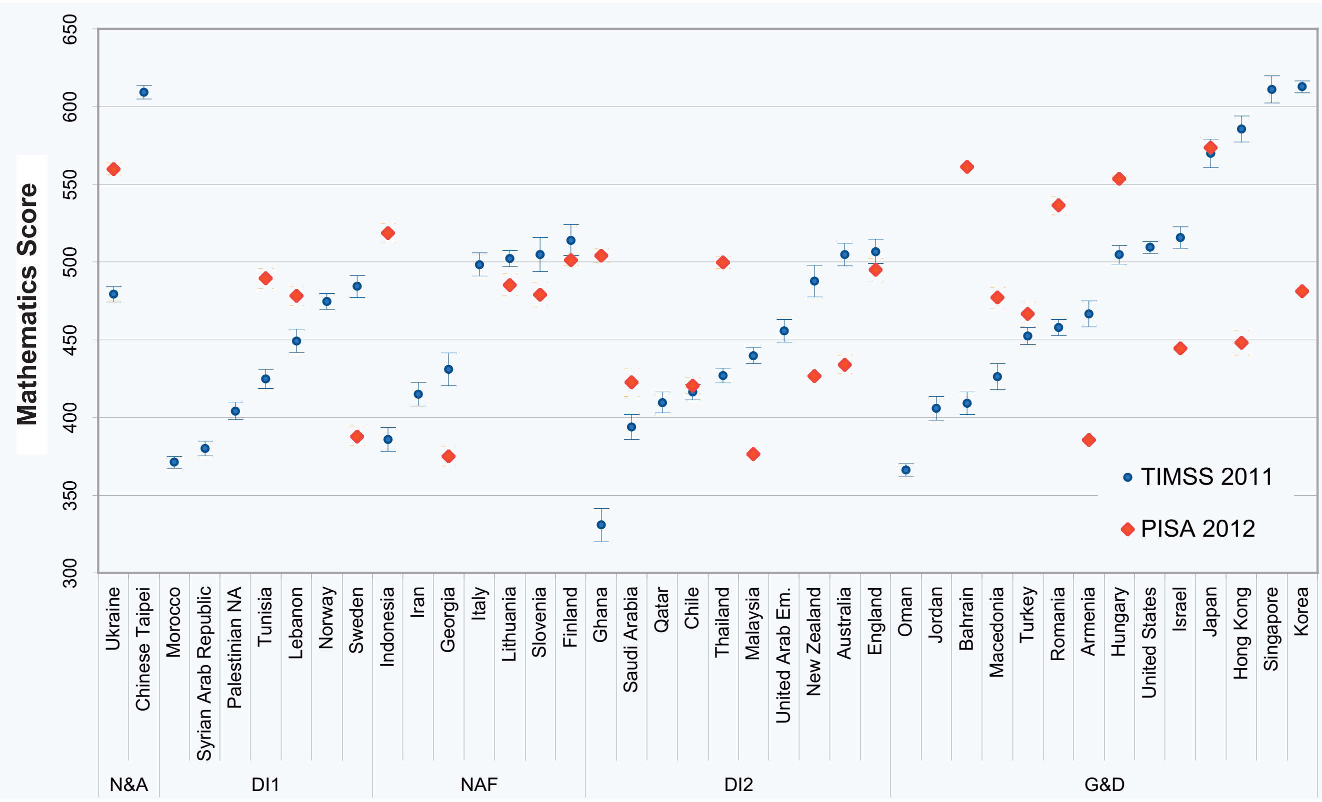

Figure 5: TIMSS 2011 Grade 8 and PISA 2012 Mathematics performance by countries within each curriculum group

Table 4: Results of meta-regression examining the relationships between Mathematics performance and curriculum

	Performance in TIMSS 2011 (1)			*Performance in PISA 2012 (2)*			*Performance in PISA 2012 (3)*		
	Coeff.	*S.E.*	*P-val.*	*Coeff.*	*S.E.*	*P-val.*	*Coeff.*	*S.E.*	*P-val.*
Intercept	366.13	23.13	<.0001	374.29	28.77	<.0001	6.86	40.35	0.87
Aggregated background measure	1.79	0.39	<.0001	1.68	0.46	0.00	0.15	0.26	0.55
Per capita GDP	0.00	0.00	0.75	0.00	0.00	0.76	0.00	0.00	0.73
Country group (ref.: Geometry and Data)									
Number and Algebra	39.59	32.01	0.22	55.65	41.86	0.20	-12.49	19.29	0.52
Delayed introduction 1	-26.28	20.71	0.21	-34.10	26.97	0.22	16.60	12.70	0.21
Non-algebraic focus	-13.92	19.57	0.48	-12.91	21.79	0.56	17.49	9.83	0.09
Delayed introduction 2	-32.06	19.50	0.11	-27.28	20.69	0.20	17.46	10.04	0.10
TIMMS 2011 score	-	-	-	-	-	-	0.91	0.10	<.0001

PISA 2012 was generally in line with their performance in TIMSS 2011.

These results suggest that the curriculum may be related to TIMSS and PISA performance in Mathematics. More specifically, it seems that delaying the introduction of topics like those related to 'Number' and Algebra' is negatively correlated to TIMSS performance. However, this finding may be confounded by other cultural and societal factors.

Setting aside the potential confounding influence of the cultural differences between countries, in order to further investigate the statistical significance of differences after accounting for the influence of the aggregated background factors, a meta-regression analysis was undertaken. More specifically, three different models were estimated and the results are shown in Table 4.

The first set of results (column 1) relate to the TIMSS 2011 performance. This confirms that, although the estimate of the coefficients of countries in both *Delayed introduction* groups and those in the *Non-algebraic focus* class are negative (which means that these countries tend to be out-performed by similar countries with different curricula), the differences are not statistically significant. It should be noted, however, that the magnitude of the coefficient associated to the *Delayed introduction 2* group is -32 (p-value = 0.11), suggesting that students in these countries tend to achieve 32 points lower than those in countries within the *Geometry and Data* group.

The results relating to PISA 2012 performance (column 2) are not particularly different from those relating to TIMSS 2011. The coefficients associated with the groupings are not significant, suggesting that curriculum differences across countries do not affect PISA Mathematics performance. The last set of results (column 3) indicates that Mathematics performance in PISA is not significantly affected by the groupings, which confirms the above results. However, it is worth mentioning that, for countries in the *Delayed introduction 2* group, the association with PISA performance is positive relative to achievement in TIMSS (column 3), whilst the association with TIMSS performance as a whole is negative (see column 1) - albeit the coefficients are not significant in both cases.

To summarise, these results provide no strong evidence that different curricula may be more beneficial for achievement in either PISA or TIMSS.

Conclusions and discussion

In this article we have investigated the relationship between Mathematics performance and the prevalence of different teaching styles, resources used in the classrooms, and curriculum using TIMSS 2011 Grade 8 (and partly PISA 2012) data. In undertaking this analysis, we relied on teachers' responses about the activities they asked their students to do, the resources they employed and the actual topics they taught.

The country-level analysis highlighted that countries were grouped differently by teaching styles than by curriculum. More specifically, grouping by teaching styles matched with contextual descriptions of the countries which include geographical, historical, societal and cultural factors. This suggests that, within countries with a similar background, teachers tended to share the same methods of teaching. Conversely, countries within the same groupings by topics taught did not share a common contextual description. This is also an indication that the Mathematics curriculum may be less influenced by countries' contextual characteristics than by teaching styles. On the other hand, the results of the meta-regression analysis suggested that teaching styles can be more connected to students' Mathematics performance on TIMSS and PISA than curriculum. In particular, our findings indicate that some teaching methods may be more beneficial in terms of PISA rather than TIMSS results and vice versa. Using the labels proposed in Table 1, a *Learn, repeat, and check* style appeared to be helpful in improving achievement measured by means of curriculum-related questions but not to apply Mathematics to real-life. In contrast, a simplified style appeared to be relatively disadvantageous in terms of TIMSS achievement, but positively associated to performance in PISA.

In interpreting these findings, it is important to stress that, since there may be other unobservable background factors affecting students' achievement, it is not possible to infer causal relationships from the results we have shown in this article. However, our research shows how data from international benchmarking studies can be exploited to provide empirical evidence about the link between teaching styles, curriculum, and Mathematics achievement. Although we cannot draw conclusions about which specific teaching practices and topics in the curriculum lead to better results in Mathematics achievement, our findings, based on detailed quantitative analysis, can be used to reflect upon Mathematics teaching styles and curriculum and their role in providing a more effective Mathematics education aimed at preparing students for their future lives and careers.

References

Andrews, P. (2011). The cultural location of teachers' mathematical knowledge: Another hidden variable in mathematics education research? In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching*, *50*, 99–118.

Andrews, P. (2012). Learning from others: Can PISA and TIMSS really inform curriculum development in mathematics? *Mathematical Gazette*, *96*(537), 386-407.

Andrews, P. (2014). European mathematics curricula and classroom practices. In P. Andrews & T. Rowland (Eds.), *Master Class in Mathematics Education*. London, Bloomsbury.

Benbow, C.P. & Faulkner, L.R. (2008). Rejoinder to the Critiques of the National Mathematics Advisory Panel Final Report, *Educational Researcher*, *37*(9), 645-648.

Benton, T. (2014). Using meta-regression to explore moderating effects in surveys of international achievement. *Practical Assessment Research and Evaluation*, *19*(3). Retrieved from: http://pareonline.net/pdf/v19n3.pdf.

Boaler, J. (2008). When politics took the place of inquiry: a response to the National Mathematics Advisory Panel's review of instructional practices. *Educational Researcher*, *37*(9), 588-594.

Clarke, D. (2004). Researching Classroom Learning and Learning Classroom Research, Guest Editorial in *The Mathematics Educator*, *14*(2), 2-6.

DfE (2010). *The Importance of Teaching. The Schools White Paper 2010*. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf.

Elliott, G. (2013). *The Cambridge approach to comparative evidence from other jurisdictions*. Cambridge Assessment Research Report. Cambridge: Cambridge Assessment.

Gill, T. & Benton, T. (2013). *Investigating the relationship between aspects of countries' assessment systems and achievement on the Programme for International Student Assessment (PISA) tests*. Cambridge Assessment Research Report. Cambridge: Cambridge Assessment.

Hanushek. E. A. & Woessmann, L. (2011). The Economics of International Differences in Educational Achievement. In E. A. Hanushek, S. Machin and L. Woessmann (Eds.) *Handbook of the Economics of Education*, Vol. 3. San Diego, CA: Elsevier.

International Association for the Evaluation of Educational Achievement (IEA), 2011. TIMSS *2011 Teacher Questionnaire Mathematics Grade 8*. TIMSS & PIRLS International Study Center, Boston College. Retrieved from: http://timssandpirls.bc.edu/timss2011/downloads/T11_TQM_8.pdf

Martin, M.O. & Mullis, I.V.S. (Eds.) (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Foy, P. & Arora, A. (2012). TIMSS 2011 *International Results in Mathematics*. Amsterdam: IEA.

Muthén, L.K. & Muthén, B.O. (2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Nylund, K.L., Asparouhov, T. & Muthén, B.O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study, *Structural Equation Modeling*, *14*(4), 535-569.

Oates, T. (2010). *Could do better: Using international comparisons to refine the National Curriculum in England*. Cambridge Assessment Research Report. Cambridge: Cambridge Assessment.

OECD (2013). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014)*. Paris: PISA, OECD Publishing.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3). Retrieved from: http://www.jstatsoft.org/v36/i03/paper.

Usiskin, Z. (2012). Misidentifying factors underlying Singapore's high test scores: A strong curriculum is not the sole reason for Singaporean students' success on international assessments. *Mathematics Teacher*, *105*(9), 666-670.

Wu, M. (2009). A comparison of PISA and TIMSS 2003 achievement results in mathematics. *Prospects*, *39*, 33-46.

Zanini, N. & Benton, T. (2015). *The Importance of Teaching Styles and Curriculum in Mathematics: Analysis of TIMSS 2011*. Cambridge Assessment Research Report. Cambridge: Cambridge Assessment.

APPENDIX

The reported questions 19, 20 and 23 from the TIMSS 2011 Teacher Questionnaire Mathematics Grade 8 (International Association for the Evaluation of Educational Achievement (IEA) (2011) are presented below:

QUESTION 19: When you teach mathematics to this class, how often do you usually ask students to do the following?

	Every or almost every lesson	*About half the lessons*	*Some lessons*	*Never*
a) Listen to me explain how to solve problems	O	O	O	O
b) Memorize rules, procedures, and facts	O	O	O	O
c) Work problems (individually or with peers) with my guidance	O	O	O	O
d) Work problems together with the whole class with direct guidance from me	O	O	O	O
e) Work problems (individually or with peers) while I am occupied by other tasks	O	O	O	O
f) Apply facts, concepts, and procedures to solve routine problems	O	O	O	O
g) Explain their answers	O	O	O	O
h) Relate what they are learning in mathematics to their daily lives	O	O	O	O
i) Decide on their own procedures for solving complex problems	O	O	O	O
j) Work on problems for which there is no immediately obvious method or solution	O	O	O	O
k) Take a written test or quiz	O	O	O	O

QUESTION 20: When you teach mathematics to this class, how often do you usually ask students to do the following?

	Basis for instruction	*Supplement*	*Not used*
a) Textbooks	O	O	O
b) Workbooks or worksheets	O	O	O
c) Concrete objects or materials that help students understand quantities or procedures	O	O	O
d) Computer software for mathematics instruction	O	O	O

QUESTION 23: The following list includes the main topics addressed by TIMSS mathematics test. Choose the response that best describes when the students in this class have been taught each topic

Domain	*Topic*	*Before this year*	*This year*	*Not yet*
Number	a) Computing, estimating, or approximating with whole numbers	O	O	O
	b) Concepts of fractions and computing with fractions	O	O	O
	c) Concepts of decimals and computing with decimals	O	O	O
	d) Representing, comparing, ordering, and computing with integers	O	O	O
	e) Problem solving involving per cents and proportions	O	O	O
Algebra	a) Numeric, algebraic, and geometric patterns or sequences (extension, missing terms, generalization of patterns)	O	O	O
	b) Simplifying and evaluating algebraic expressions	O	O	O
	c) Simple linear equations and inequalities	O	O	O
	d) Simultaneous (two variables) equations	O	O	O
	e) Representation of functions as ordered pairs, tables, graphs, words, or equations	O	O	O
Geometry	a) Geometric properties of angles and geometric shapes (triangles, quadrilaterals, and other common polygons)	O	O	O
	b) Congruent figures and similar triangles	O	O	O
	c) Relationship between three-dimensional shapes and their two-dimensional representations	O	O	O
	d) Using appropriate measurement formulas for perimeters, circumferences, areas, surface areas, and volumes	O	O	O
	e) Points on the Cartesian plane	O	O	O
	f) Translation, reflection, and rotation	O	O	O
Data and Chance	a) Reading and displaying data using tables, pictographs, bar graphs, pie charts, and line graphs	O	O	O
	b) Interpreting data sets (e.g., draw conclusions, make predictions, and estimate values between and beyond given data points)	O	O	O
	c) Judging, predicting, and determining the chances of possible outcomes	O	O	O