

Chapter 3

Methodology: Examining the Globalization of Science Curricula Using TIMSS

Abstract Twenty years of collected TIMSS data reveals interesting insights into the globalization of science curricula. To answer the research questions, three different methods were used to analyze the TIMSS dataset. First, changes in countries' intended science curricula were captured and coded over the course of three TIMSS cycles (1999, 2007 and 2015 for Grade 8, and 2003, 2007 and 2015 for Grade 4). Changes were identified using countries' responses to the TIMSS curriculum questionnaires. This approach tracks changes in national science curricula over time. Second, cluster and discriminant analyses of the curriculum questionnaire data were used to determine potential convergence of curricula; countries may be clustered into groups on the basis of the topics included or not included in their intended science curricula. Third, the TIMSS encyclopedias and TIMSS teacher questionnaires provide detailed information on additional features of the implemented science curricula, such as the mean time spent teaching science in each country, or the percentage of students taught the TIMSS science topics. Such information was carefully analyzed for a subsample of countries. However, inconsistencies in the way this information was collected and presented across different TIMSS cycles made comparisons across countries between cycles challenging. As each of the three methods has its advantages and limitations, investigating the three research questions from different angles and with different, yet complementary techniques offers the most comprehensive analysis of the available TIMSS data.

Keywords Cluster analysis • Curriculum coding • Curriculum convergence
Discriminant analysis • International large-scale assessment • Science education
Trends in Mathematics and Science Study (TIMSS) • TIMSS curriculum
questionnaire

A review of the existing literature identified several methods have been used previously to investigate globalization of curricula The research methodology we used to investigate our three research questions can be separated into three distinct strands.

- (1) We coded responses to the TIMSS science curriculum questionnaire data, in order to track changes in national science curricula over time.
- (2) We undertook cluster and discriminant analyses, which grouped countries into distinct groups on the basis of the TIMSS science topics included in their intended science curricula. (The reasons and rationale for using cluster and discriminant analyses over other statistical techniques such as latent class analysis will be discussed later in the chapter.)
- (3) We coded additional science curriculum features for a smaller sample of countries.

All three methods made use of data collected by the TIMSS curriculum questionnaire, which provides information on which TIMSS science topics are included in the intended science curricula of participating countries. In each case, we elaborate further on the limitations of the data sources and methods.

3.1 Coding of Curriculum Questionnaire Data

The TIMSS curriculum questionnaires provide a rich array of data on the science curricula and science teaching in participating countries at Grades 4 (ages 9–10) and 8 (ages 13–14). One section of the curriculum questionnaires pertains to science topics covered at each particular grade (either Grade 4 or 8). The questionnaire asks respondents (typically representatives from the Ministry of Education or equivalent) to indicate whether each TIMSS science topic is taught to “all or almost all students” in the grade, “only the more able students”, or whether it is “not included in the curriculum” through to that grade (see Fig. 3.1 for an example from the questionnaire).

The responses to the questionnaires reveal the TIMSS science topics included in a country’s science curriculum for a particular grade in a particular year. To address our first research question on whether there have been changes in countries’ intended science curricula since 1999, individual countries’ responses to the curriculum questionnaires between 1999 and 2015 were coded.

By comparing individual countries’ responses to the curriculum questionnaire between different TIMSS cycles, changes in the science topics included in countries’ intended science curricula over time can be detected.

Because the TIMSS curriculum questionnaire was first administered in 1999 for Grade 8 and 2003 for Grade 4, it was not feasible to code responses from every cycle of TIMSS; TIMSS 1999 was used as the baseline for Grade 8 and TIMSS 2003 as the baseline for Grade 4 (Table 3.1). For both grades, TIMSS 2015 was the most recent time point that could be considered. In addition to the earliest and most recent TIMSS cycles with available curriculum questionnaires, an intermediate TIMSS cycle (2007) was also included in this coding exercise. The inclusion of an intermediate cycle allowed changes in a country’s science curriculum to be detected at more than one time point. This is important as countries may have changed the

Eighth Grade Science Topics Covered

This science module refers to the national curriculum that was in effect for the eighth grade students assessed in TIMSS 2015—the curriculum that covers science instruction at the eighth grade of formal schooling for the majority of students. If you do not have a national curriculum, please summarize for your state or provincial curricula.

S8. (i) According to the national science curriculum, what proportion of grade 8 students should have been taught each of the following topics or skills by the end of grade 8?

Be sure to include curriculum expectations for all grades up to and including grade 8. Grades represent years of formal schooling. For example, if “Year 9” in your country corresponds to the eighth year of formal schooling, please choose grade 8.

(ii) Across grades from preprimary through upper secondary education, at what grade(s) are the topics primarily intended to be taught?

If there are not any specifications to this detail, please indicate national expectations to the best of your ability. If part of a topic does not apply [e.g., energy flow in part A topic (f)], please explain in the comment field.

	(i) Proportion of grade 8 students expected to be taught topic			(ii) Grade(s) topic is expected to be taught preprimary (PP) through the end of upper secondary (G12)												
	Check one circle for each line.			Check the corresponding grade(s) for each topic												
	All or almost all students	Only the more able students	Not included in the curriculum through grade 8	PP	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
A. Biology																
a) Differences among major taxonomic groups of organisms (plants, animals, fungi, mammals, birds, reptiles, fish, amphibians)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Major organs and organ systems in humans and other organisms (structure/function, life processes that maintain stable bodily conditions)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Cells, their structure and functions, including respiration and photosynthesis as cellular processes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Life cycles, sexual reproduction, and heredity (passing on of traits, inherited versus acquired/learned characteristics)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Role of variation and adaptation in survival/extinction of species in a changing environment (including fossil evidence for changes in life on Earth over time)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Interdependence of populations of organisms in an ecosystem (e.g., energy flow, food webs, competition, predation) and factors affecting population size in an ecosystem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Human health (causes of infectious diseases, methods of infection, prevention, immunity) and the importance of diet and exercise in maintaining health	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fig. 3.1 Extract from curriculum questionnaire for Grade 8 science TIMSS 2015. *Source* TIMSS 2015 Grade 8 curriculum questionnaire © 2015 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College

Table 3.1 TIMSS cycles used in the analysis

Grade	TIMSS cycles used
4	2003, 2007, 2015
8	1999, 2007, 2015

topics in their science curriculum more than once between the baseline and 2015 cycles. If a science topic was added between 1999 and 2007 and then removed between 2007 and 2015, the change would not be detected if only the baseline and 2015 cycles were considered.

The science topics included in the curriculum questionnaire varied slightly for each cycle of TIMSS. However, in order to map changes in the participating countries' curricula consistently across the TIMSS cycles, it was necessary to consider the same set of science topics in each cycle. Therefore, we first mapped the science topics across the TIMSS cycles to identify equivalent topics in the curriculum questionnaires of each TIMSS cycle.

We used the TIMSS 2015 science topics as the basis for this mapping exercise, matching science topics in the older TIMSS curriculum questionnaires to the 2015 TIMSS science topics. In some cases, there was a direct match in science topics across cycles, while, in other cases, there was no match. We found there were science topics in the TIMSS 2015 curriculum questionnaire that were not included in earlier questionnaires and, similarly, there were topics from earlier cycles that were not included in the 2015 curriculum questionnaire. As an illustration, the 1999 questionnaires included topics on the nature of science and scientific inquiry, but there were no equivalent science topics in the 2015 questionnaire. Only science topics that could be matched across all three TIMSS cycles were included in the coding exercise. For Grade 8 there were 20 science topics that mapped across the three TIMSS cycles and for Grade 4 there were 21 such topics (for a full list of these topics and the results of this mapping, see Chap. 4, Sect. 4.1.5).

To identify the nature of the curriculum change, we developed a coding framework (Table 3.2).

Table 3.2 Codes used in the curriculum mapping exercise

Code	Description of code
-2	Science topic has been removed from the curriculum at that grade
-1	Reduced emphasis on the science topic in the curriculum at that grade (topic was previously taught to most or all students in grade, but is now only taught to the more able students in the grade)
0	No change to science topic
1	Greater emphasis on the science topic in the curriculum at that grade (topic was previously taught only to more able at that grade but is now taught to most or all students in the grade)
2	Science topic has been introduced to the curriculum at that grade

In order to complete the coding exercise, the responses of individual countries to the curriculum questionnaire for each TIMSS cycle considered were obtained from the IEA data repository (data and documentation files from completed IEA studies are freely available for research purposes at <http://www.iea.nl/data>). For each science topic that mapped across the three TIMSS cycles, a country's response to whether the science topic was intended to be taught to "all or most students", "only the more able students", or was "not included in the curriculum" through to that grade was compared between cycles. For example, in Iran in 2003 at Grade 4, the curriculum questionnaire response to the life sciences topic "the characteristics of living things and the major groups of living things" indicated that the topic was taught to "all or most students". However, in 2007, the curriculum questionnaire response for this topic indicated that it was "not included in the curriculum through Grade 4". Therefore between 2003 and 2007 this topic was coded as a -2 (science topic removed from the curriculum at that grade) for Iran. This process was repeated for each country that had participated in at least two of the TIMSS cycles under consideration. Pairwise comparisons were made, for example at Grade 4, comparing the curriculum questionnaire responses in 2003–2007, 2007–2015, and 2003–2015.

This method addresses our first research question concerning whether there have been changes in countries' intended science curricula, as well as providing information about the nature of these changes; for example, whether countries were increasing the breadth of their curriculum by adding additional topics. This approach also partially addresses our second research question about whether science curricula are becoming increasingly similar across countries over time, as the coding indicates whether these changes are leading to an increased convergence in the intended science curricula across countries over time. The analysis also provides some evidence for our third research question on the potential existence of an international core science curriculum, by identifying science topics that are included in the intended curricula of all or the majority of countries participating in TIMSS.

3.2 Cluster Analysis and Discriminant Analysis

Our second research question asked whether science curricula taught in different countries have been becoming more similar over the last 20 years. We investigated an array of complementary statistical techniques, and opted for a cluster analysis followed by a discriminant analysis.

We applied the cluster analysis to the first TIMSS cycle (1999 for Grade 8 and 2003 for Grade 4) in order to identify clusters of countries that were similar in terms of their intended science curricula. The cluster analysis was conducted on countries' responses to the section of the curriculum questionnaire detailing which science topics were included in the intended curriculum at each grade. Only the science topics which could be mapped across all three TIMSS cycles were included in the analysis to ensure comparability. The second step, the discriminant analysis, was

again performed on the first cycle of TIMSS, in order to produce a model that was then used to classify countries in the following TIMSS cycles.

The advantage of this two-step technique over a simple cluster analysis repeated on the three cycles of TIMSS considered is that, while the cluster analysis classifies countries based on the specific information of what is taught in each of the three cycles using the exact same model every time, the discriminant analysis uses a completely separate algorithm to classify countries in each of the three cycles.

In this investigation, the comparison of clusters of countries over time is not meaningful as the clusters will have been created using different criteria each time. Furthermore, the set of countries participating in each cycle is not constant, with new countries joining and some countries which participated in the early TIMSS cycles leaving. This issue is partially addressed by discriminant analysis. This is because new entrants to the more recent TIMSS cycles do not have an effect on the clustering rule, which is fixed. As a result, new entrants are simply classified as belonging to one cluster group based on their responses to the TIMSS curriculum questionnaire.

That said, the clustering algorithm itself is influenced by the specific set of countries that were included in the first TIMSS cycle. A straightforward solution to this problem would be to restrict the analysis to countries that have participated in all three TIMSS cycles. We chose not to pursue this approach as we considered the advantages to be more than offset by the disadvantages.

One of the main drawbacks to this approach is that it would narrow the analysis to a relatively small set of countries (21 for Grade 8 and 18 for Grade 4) and, as a consequence, would limit the conclusions we would be able to make from the analysis. This, in turn, would give a very partial response to our research question, as we would be testing the hypothesis of convergence in science curricula on a subsample of countries. What is going on in this specific subsample might be not representative of what is happening in the overall population. With this caveat in mind, we chose to prioritize a more complete analysis and exploit all the information available by running our model on the full sample.

A limitation of our approach is that the number of clusters is identified using the first TIMSS cycle and is therefore fixed for the following two cycles. If science curricula are becoming more dissimilar, this would result in a higher number of clusters in more recent TIMSS cycles. We would not be able to detect this phenomenon using our approach. However, given the relative stability of science curricula, this event is unlikely. To check this assumption, when we conducted three separate cluster analyses on the three cycles of TIMSS, the ideal number of clusters identified by the algorithm was always the same (two in each case).¹

¹The results of the three separate cluster analyses are not reported due to concerns regarding the misinterpretation of the outcome. Nevertheless, the finding that the number of groups identified by the algorithm is constant in all cycles is relevant in establishing that the science curricula of countries are not diverging to such an extent to justify the identification of an increasing number of clusters over time.

Having outlined our statistical approach and the rationale behind it we now give a more thorough account of the statistical techniques used in this investigation.

Cluster analysis is a statistical technique that allows the grouping of a set of observations according to certain characteristics, in such a way that observations in the same cluster are more similar to each other than to observations in other groups. In the context of our investigation, the observations that are clustered are countries, and the variables used to group them are the responses given to the TIMSS curriculum questionnaire items on the science topics that are included in countries' intended science curricula. Our analysis aims to investigate the existence of convergence in science curricula, with convergence signaled by the tendency of a group to expand at the expense of others.

When running the analysis on the TIMSS dataset we chose to implement a two-step cluster analysis. The first step of this procedure consists of pre-clustering the records into small sub-clusters that are then consolidated into a smaller number of groups in the second step. The advantage of this methodology over other clustering algorithms is that, when the number of groups is unknown, the algorithm automatically returns the optimal number of clusters based on the Bayesian information criterion (BIC).² Observations are grouped together based on a distance measure that, in this case, reflects how different countries are from each other in terms of science topics covered in their curricula. In order to cluster the observations, we adopt the log-likelihood criterion that is appropriate when grouping observations using continuous as well as categorical variables. This distance measure works best when all variables are independent and categorical variables have a multinomial distribution.³ The log-likelihood is a probability-based distance. The distance between two clusters is related to the decrease in log-likelihood as they are combined.⁴

Our original intention was to perform a latent class analysis on the data. We performed a cluster analysis followed by a discriminant analysis due to the reduced size of the sample. Given the large number of parameters estimated, latent class analysis requires a considerable number of observations. (In statistics, a model cannot be identified unless the sample size is one more than the number of predictors and, as a rule of thumb, we need around ten observations per parameter to estimate a model with reasonable precision.)

²The Bayesian information criterion (BIC) is used to select between different models based on how well they fit the data and the number of parameters they use. The BIC is an increasing function of both the number of parameters (the BIC penalizes the complexity of the model) and the error variance (a measure of how well the model fits the data). The algorithm chooses the model with the lowest BIC that implies either fewer explanatory variables, better fit, or both.

³In practice, these conditions are rarely satisfied, but the algorithm is thought to behave reasonably well when the assumptions are not met.

⁴The goal of the clustering algorithm is to maximize the probability (or likelihood) of obtaining the figures observed in the dataset, given the clusters that are identified at the final stage. Hence, given that our aim is to maximize the likelihood function, the bigger the drop that results from merging two clusters, the more dissimilar the observations therein contained.

Previous studies which have used TIMSS data to categorize countries' curricula into particular groups using latent class analysis, such as Zanini and Benton (2015), made use of teacher questionnaire responses as opposed to curriculum questionnaire responses. There are a far greater number of observations in the teacher questionnaire than there are in the curriculum questionnaire, making the latent class analysis approach viable for studies using teacher questionnaire data. It is not possible to use latent class analysis in this study as there is only one observation available for each participating country, and this study uses the curriculum questionnaire as opposed to the teacher questionnaire because of time and resource constraints. That said, the purpose of cluster analysis is the same as latent class analysis but, instead of using a model to group observations, it employs distance measures that are less demanding in terms of number of observations needed.

Discriminant analysis is, in a sense, used as a tool to reverse what has been implemented with cluster analysis. It is a statistical technique used to determine which variables discriminate between two or more groups (Ayinla and Adekunle 2015). This procedure was performed on the groups of countries that had been previously identified by the cluster analysis. It created a model which identified the most important science topics in determining the grouping of countries.

To illustrate how this works in practice, if a particular science topic, for example "forces", is taught in all countries in group A and in none of the countries in group B, given group A and B,⁵ the algorithm that performs discriminant analysis is able to detect that this science topic is an important element of discrimination between the two groups. The algorithm infers that countries teaching that specific science topic will be highly likely to belong to group A, while countries that do not teach the topic will be highly likely to belong to group B.

In reality, the algorithm performing discriminant analysis is slightly more complex, as it takes into account all the science topics at once and, for each of them, estimates a coefficient that reflects the probability of belonging to a given group based on the answer given to that particular question (for example, whether students at that grade are expected to be taught that science topic or not). When all the coefficients are fitted in a linear model, for each observation (country), we are able to predict the group identity based on the answers given in the curriculum questionnaire.

The model can then be used to predict the clustering of countries in subsequent years. If convergence of science curricula has occurred in the last 20 years, we should observe a tendency for countries to concentrate in a single group.

As indicated above, this approach has advantages compared to cluster analysis alone but it is not immune from criticism. Firstly, the approach enables the identification of convergence (when one group of countries grows over time at the expense of others). However, as already mentioned, it does not enable the detection of divergence. This shortfall arises because, by construction, the number of clusters in cycles following the baseline year will be equal to or lower than the number of

⁵In this case the two groups were identified by means of cluster analysis.

clusters detected in that baseline year. We have already provided explanations for why this is not a major concern in our specific dataset. Moreover, we should point out that the scope of this study is to assess whether convergence in science curricula has occurred rather than divergence.

A second concern is that, although this procedure allows the detection of whether science curricula are getting more similar over time, it also focuses on features that made the science curricula of countries different in the first place. This means that the science topics that are particularly important in determining the group a country belongs to in the baseline year will keep on playing a primary role in subsequent cycles. Additionally, science topics that do not discriminate strongly between groups in the baseline cycle will only play a marginal role in the classification of countries in subsequent years.

Continuing with the earlier example, if, in the baseline year, all countries in group A are teaching a given science topic and no countries in group B are teaching that topic, but, in the second cycle, all countries include the science topic in their curriculum, the algorithm would, in this case, predict that all observations in the second cycle belong to group A. In light of this result, we could be tempted to conclude that curricula are converging but this might not be the case. If there is another science topic in the questionnaire and this topic is taught by all countries in the first cycle, the algorithm would completely ignore this item as it is not able to discriminate between observations (countries). However, if in the following cycle, some countries stop teaching that topic so that part of the sample has it in their science curricula and part of the sample does not, the model will not be able to exploit this information because the topic is not included in the predictive model.⁶ In this example, one group grows at the expense of the other, but convergence is not supported. The curricula are becoming more similar in some dimensions, but are diverging in other aspects that our model is not taking into account.

This is only a hypothetical example and, in reality, the situation is more complicated as the algorithm considers multiple science topics at a time. Moreover, it is unlikely to come across science topics with zero predictive power.⁷ More often we will have science topics with low predictive power (i.e. not able to discriminate between groups with high precision). In this case, the information will not be completely disregarded but the algorithm will instead attach a low coefficient to the science topic. As a consequence the answer given to the question related to that science topic will affect the probability of belonging to a certain group only by a very small degree.

In order to resolve the issue related to the dependence of our results on the predictive power of science topics in the initial cycle of the questionnaire, the same process was repeated in reverse. This means that, starting from the most recent TIMSS

⁶We could consider this topic to be included in the model with a zero coefficient, reflecting the fact that its predictive power was null in the first cycle.

⁷This occurs if, as in our example, all countries in the sample give the same answer to a certain question.

cycle available (2015 for both Grade 4 and Grade 8), the data were clustered and a discriminant analysis performed based on the resulting groups. Using the model that emerges from the discriminant analysis, countries were classified according to the answers given in earlier cycles of the questionnaire. This reverse approach was used to assess differences in curricula based on science topics that were homogeneously taught (or not taught) in the first TIMSS cycle and hence have diverged.

Using the previous example, with this new approach, the algorithm will now cluster countries based on the answers given in the final cycle of the questionnaire. A science topic that did not discriminate in the first cycle (as it was taught by all countries) is now an important predictor for clustering observations in the final cycle, and the science topic that was a good discriminant in first place, is no longer relevant. Applying the reverse approach to this case, the cluster analysis conducted on the final cycle will group observations into two separate groups while assigning all observations to a single group in the initial cycle (because all countries gave the same answer to the item in the earlier cycle). Given that one of the clusters was bigger in the first cycle than in the end-point questionnaire, we can conclude that curricula have diverged along the specific dimension (science topic) we are analyzing. The interpretation will be complicated by the fact that in reality our model considers multiple science topics at a time.

It is important to note that discriminant analysis is a linear technique and we have applied it to a dataset in which we have categorical variables. This approach is not inappropriate provided the underlying assumptions are understood. The challenge of applying a linear technique to categorical variables is that, for each item, the model will treat the categories as continuous variables and will return a single coefficient that has the same interpretation as the beta in a linear regression.⁸ Given that our variables are measured on a three-point scale (ranging from 1 to 3), for a given science topic, the probability of belonging to a given group will be two times larger/smaller for a country with a value of 2 (topic taught to only the more able students) compared to a country with a value of 1 (topic taught to all or most students). Similarly, given the linear nature of the coefficients, the probability of belonging to a given group will be three times larger/smaller for a country with a value of 3 (topic not taught in a given grade) compared to a country with a value of 1 (topic taught to all or most students). This is a limitation of our analysis but, as long as we are willing to assume that individuals in the middle of the scale (i.e. teaching to the more able students only) are equally different from individuals at the two extremes (teaching to all or teaching to nobody), then the use of a linear model is appropriate. For our purpose the assumption is not unreasonable considering the fact that the limited amount of data prevents us from estimating a non-linear model.⁹

⁸In a linear regression the beta coefficient represents the difference in the predicted value of Y for each one-unit difference in a certain variable, keeping all other variables constant.

⁹Here we encounter a similar problem to the one we have already discussed for latent class analysis. A non-linear model (such as latent class analysis) would require the estimation of a coefficient for each of the values taken by each categorical variable. In our case, for example at Grade 8, we have 20 variables with three possible outcomes each, implying 40 coefficients to estimate. This is impossible given the sample size.

From this short description of our statistical analysis it is evident that assessing convergence in science curricula is not a straightforward task and that different techniques are likely to result in different outcomes. Given that each approach has advantages and disadvantages, investigating the research question from different angles and with different yet complementary techniques offers the most robust solution.

Finally, readers should be aware that cluster analysis does not involve hypothesis testing and the computation of significance levels. As a result, the appropriateness of a solution can only be assessed by the researcher. This should be evaluated through the critical observation of the elements composing the clusters and the difference in the pattern of answers given by countries belonging to different groups.

3.3 Additional Analysis of a Sample of Countries

Both the coding of curriculum questionnaires and statistical methods outlined in Sects. 3.1 and 3.2 make exclusive use of responses from the TIMSS curriculum questionnaire. This allows a rigorous comparison of the intended science curricula of countries participating in TIMSS over time. This dataset, however, gives little information about the implemented science curriculum of a country. To complement these two earlier approaches, a number of other features of the implemented science curriculum were captured for a selection of countries that have participated in at least two of the TIMSS cycles considered in this report. Changes in these features were recorded to investigate whether there is convergence over time.

We collected information about various aspects of the implemented science curricula, such as the mean time spent teaching science in each country, the percentage of students taught the TIMSS science topics, the organization and structuring of the science curriculum, and science assessment. This information came from the TIMSS encyclopedias and from the TIMSS teacher questionnaires for the selected countries (Martin et al. 2000, 2004, 2008; Mullis et al. 2016). For this part of our study, we selected 15 countries at Grade 4 (Table 3.3) and 16 countries at Grade 8 (Table 3.4) on the basis of a number of different criteria. Firstly, the majority of these countries have appeared in all three TIMSS cycles under study at both Grade 4 and 8. The countries that have not, such as Qatar, have appeared in at least two cycles and were included to widen the geographic distribution and achievement profile of the sample. In a few cases, a country that participated in all three TIMSS cycles was not included. This was either because the country appeared in three cycles for one grade but not the other, or because there was no data available from the curriculum questionnaire on the topics included in the intended science curriculum in a particular grade. This was the case, for example, for the Netherlands. Such countries were discounted from the selection. Countries which participated in only one relevant TIMSS cycle were also excluded from this analysis, thus ensuring that the amount of information that can be collected on each country was maximized.

Table 3.3 Countries included in additional analysis at Grade 4

Country	Average TIMSS science score in 2003	Average TIMSS science score in 2007	Average TIMSS science score in 2015
Morocco	304 (6.7)	297 (5.9)	352 (4.7)
Iran	414 (4.1)	436 (4.3)	421 (4.0)
Qatar	N/A	294 (2.6)	436 (4.1)
New Zealand	520 (2.5)	504 (2.6)	506 (2.7)
Italy	516 (3.8)	535 (3.2)	516 (2.6)
Australia	521 (4.2)	527 (3.3)	524 (2.9)
Lithuania	512 (2.6)	514 (2.4)	528 (2.5)
England	540 (3.6)	542 (2.9)	536 (2.4)
Hungary	530 (3.0)	536 (3.3)	542 (3.3)
Slovenia	490 (2.5)	518 (1.9)	543 (2.4)
United States	536 (2.5)	539 (2.7)	546 (2.2)
Chinese Taipei	551 (1.7)	557 (2.0)	555 (1.8)
Hong Kong	542 (3.1)	554 (3.5)	557 (2.9)
Japan	543 (1.5)	548 (2.1)	569 (1.8)
Singapore	565 (5.5)	587 (4.1)	590 (3.7)

Note Standard errors are provided in brackets after the average achievement scores. *N/A* not applicable

In our selection of countries, we endeavored to reflect the geographic scope of TIMSS and also represent the full range of achievement on the TIMSS assessment. This means that the sample includes high achievers such as Singapore and Japan, more moderate achievers such as Italy and New Zealand, and countries that have performed less well, such as Morocco, Qatar and Iran. Furthermore, our sample includes Middle Eastern countries, East Asian countries, European countries, southern hemisphere nations and African nations. The selection also includes some less developed countries, such as Iran, and takes into account the outcomes of the cluster analysis by including countries from each cluster identified by the analysis.

The same 15 countries were selected for both Grade 4 and Grade 8 as this represented the most efficient strategy for collecting information on additional aspects of science curricula, such as time allocated to science and percentage of students taught the TIMSS science topics. For Grade 8, Israel was included as a sixteenth country as the literature review identified it as a country that had made extensive changes to its science curriculum as a result of participation in TIMSS (Klieger 2015). As a consequence, exploring the science curriculum in Israel seems highly relevant to this study.

Table 3.4 Countries included in additional analysis at Grade 8

Country	Average TIMSS science score in 1999	Average TIMSS science score in 2007	Average TIMSS science score in 2015
Morocco	323 (4.3)	402 (2.9)	393 (2.5)
Iran	448 (3.8)	459 (3.6)	456 (4.0)
Qatar	N/A	319 (1.7)	457 (3.0)
Italy	493 (3.9)	495 (2.8)	499 (2.4)
Israel	468 (4.9)	468 (4.3)	507 (3.9)
Australia	540 (4.4)	515 (3.6)	512 (2.7)
New Zealand	510 (4.9)	N/A	513 (3.1)
Lithuania	488 (4.1)	519 (2.5)	519 (2.8)
Hungary	552 (3.7)	539 (2.9)	527 (3.4)
United States	515 (4.6)	520 (2.9)	530 (2.8)
England	538 (4.8)	542 (4.5)	537 (3.8)
Hong Kong	530 (3.7)	530 (4.9)	546 (3.9)
Slovenia	533 (3.2)	538 (2.2)	551 (2.4)
Chinese Taipei	569 (4.4)	561 (3.7)	569 (2.1)
Japan	550 (2.2)	554 (1.9)	571 (1.8)
Singapore	568 (8.0)	567 (4.4)	597 (3.2)

Note Standard errors are provided in brackets after the average achievement scores. *N/A* not applicable

3.4 Methodological Limitations

There are a number of inherent limitations in our methodological approach and these should be taken into account when interpreting the results.

Firstly, the primary data source used for coding changes in the intended science curriculum, and for the cluster analysis, is responses to the section of the curriculum questionnaire on science topics intended to be taught. We have assumed that these questionnaires have been completed accurately but have no way of testing our assumption. There are, however, likely to be sources of error in the completion of the questionnaires. For example, the science topics in each country's curriculum are unlikely to correspond word for word to the science topics listed in the TIMSS curriculum questionnaire. Therefore, it is likely that there is some degree of subjectivity in how the curriculum questionnaires are interpreted and completed in each country.

Secondly, we have assumed that the curriculum questionnaire responses for each country accurately reflect the intended curriculum for the whole country. Whilst this is likely to be the case for countries with highly prescribed statutory national science curricula, such as England, this may not be the case in all countries, such as those with a federal structure like Germany, or those with a decentralized curriculum.

A third limitation is that coding of changes in the curriculum approach can only be used in countries which have taken part in at least two of the TIMSS cycles considered. This excludes countries that only started participating in TIMSS in more recent cycles and also excludes the many developing countries that do not participate at all. This means that developed countries are over-represented in this analysis, which limits the generalizability of the conclusions.

Fourthly, the TIMSS science topics included in the curriculum questionnaires in each cycle change slightly. As noted earlier, in the TIMSS 1999 curriculum questionnaire there was a section on science topics based on the nature of science and scientific inquiry skills. There is no equivalent section in the 2015 curriculum questionnaire. It is consequently not possible to code for changes to these areas of the curriculum and it is likely that, as a result, our coding will not capture all the changes in the content and intended science curricula of countries over time.

Another limitation is that some of the science topics included in the curriculum questionnaire are likely to be covered in subjects other than science in some countries. For example, in England, some of the Earth science topics are taught in geography. Another limitation is that the time frame considered differs for Grade 4 and Grade 8. Ideally the same time frame would have been used for both grades, but this was not possible as TIMSS was not administered at Grade 4 in 1999. As a result, and to maximize the time frames explored in the study, different starting points were used for Grades 4 and 8.

There is the further challenge that in each TIMSS cycle there are different numbers of participating countries, with some countries dropping out and new countries entering. This makes direct comparisons between different cycles less straightforward and the interpretation of results more complex. Furthermore, the scope of this study does not allow us to explore the reasons why countries leave the TIMSS study or choose to enter it.

In addition, by using TIMSS curriculum questionnaire data, this study is restricted to considering globalization in science curricula from 1999 to 2015 for Grade 8 and 2003–2015 for Grade 4. This means that changes in science curricula prior to 1999 that may have led to increased global alignment in science curricula across countries will be missed. Other studies investigating globalization of curricula have encountered similar issues. Rutkowski and Rutkowski (2009) suggested that using data from the 1990s as a baseline to measure curricular change and global alignment may be too late to detect globalization in national curricula. They argued that globalization was already exerting an influence on curricula well before this point in time.

Despite these limitations we consider the curriculum questionnaire to be a very useful and rich source of information on the intended curriculum in countries participating in TIMSS. These datasets represent the largest comparative body of evidence on science curricula recorded across both a wide range of countries and over an extended time period. By combining analyses with in-depth examination of the features of the science curriculum for a subset of countries from other TIMSS sources, we also consider some aspects of countries' implemented science curricula, uncovering additional evidence for or against convergence over time.

References

- Alayande, S., & Adekunle, B. (2015). An overview and application of discriminant analysis in data analysis. *IOSR Journal of Mathematics*, 11(1), 12–15. Retrieved from <http://www.iosrjournals.org/iosr-jm/papers/Vol11-issue1/Version-5/B011151215.pdf>.
- Klieger, A. (2015). Between two science curricula: The influence of international surveys on the Israeli science curriculum. *The Curriculum Journal*, 26(3), 404–424.
- Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., et al. (2000). *TIMSS 1999 international science report, findings from IEA's repeat of the third international mathematics and science study at the eighth grade*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Boston College. Retrieved from https://timssandpirls.bc.edu/timss1999i/pdf/T99i_Sci_All.pdf.
- Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., & Chrostowski, S. J. (2004). *Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Boston College. Retrieved from <https://timss.bc.edu/timss2003i/scienceD.html>.
- Martin, M. O., Mullis, I. V. S., Foy, P., Olson, J. F., Erberber, E., Preuschoff, C., et al. (2008). *TIMSS 2007 international science report*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Boston College. Retrieved from https://timss.bc.edu/TIMSS2007/PDF/TIMSS2007_InternationalScienceReport.pdf.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Boston College. Retrieved from <http://timssandpirls.bc.edu/timss2015/international-results/>.
- Rutkowski, L., & Rutkowski, D. (2009). Trends in TIMSS responses over time: Evidence of global forces in education? *Educational Research and Evaluation*, 15(2), 137–152.
- Zanini, N., & Benton, T. (2015). The roles of teaching styles and curriculum in mathematics achievement: Analysis of TIMSS 2011. *Research Matters*, 20, 35–44.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapters are included in the chapter's Creative Commons license, unless indicated otherwise in the credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

