An empirical look at globalization in education: An example with TIMSS mathematics data

Leslie Rutkowski, leslie.rutkowski@iea-dpc.de
David Rutkowski, david.rutkowski@iea-dpc.de
IEA Data Processing and Research Center
Hamburg, Germany

Abstract

Although globalization is a term in flux, it is broadly used to describe a world of increased interconnectedness and interdependence. While educationalists have long theorized about the possible effects of globalization, only recently has there been serious debate on how an increased connectedness among the political, economic and cultural has affected education. In this article, the influenced of global processes on international mathematics curricula as evidenced by item responses to three Trends in International Mathematics and Science Study (TIMSS) administrations (1995, 1999, and 2003) is considered. It was hypothesized that a harmonization of educational policy (and thus curriculum) should be noticeable as increased similarity within educational outcomes. Using hierarchical cluster analysis and non-linear principal component analysis, this paper attempts to address the following research questions: to what extent have students' responses on TIMSS become more similar over time; and, what do these changes suggest about globalization literature in education? Findings indicate that policy and curriculum may be impacted more by regional forces than global forces. In particular, results imply that there is little evidence that proposed global educational policy conversion is observable in student responses to assessment questions used in this study.

Keywords: *globalization, cluster analysis, non-linear canonical correlation, large-scale assessment, TIMSS*

An empirical look at globalization in education: An example with TIMSS mathematics data

Although globalization is a term in flux, often defined by the branch of social science from which you hail (notwithstanding vigorous debates within branches), it is broadly used to describe a world of increased interconnectedness and interdependence (yet not necessarily amalgamation). Further, how globalization may impact educational systems around the world is a contentious issue, with no clear resolution, especially in terms of empirical evidence. Dale (2000) provides a comparison and contrast of two theories of the impacts of global forces on education. First, a so-called *Common World Educational Culture* (CWEC) theory, "developed over a number of years by John Meyer and colleagues and students at Stanford University" (p. 420), argues that a universal model of education is being developed primarily by global institutions; while on the other hand, Dale's *Globally Structured Agenda for Education* (GSAE) theory suggests that capitalization is the driving force behind globalization and this force is negotiated by the nation-state. Dale argues that globalization's impacts on education are not universal but are the result of capitalism and global competition, at most resulting in economic or socio-political regions of influence. Regardless of the theory to which you subscribe, theories of globalization and education have received scant empirical analysis.

According to Dale (2000), there exists a GSAE where educational systems are being mediated through two systems, supra-national agencies (e.g. World Bank) and the nation-state. In line with this argument, we believe that a harmonization of educational policy (and thus, conceivably, curriculum) should be noticeable as increased similarity within educational outcomes. Dale posits that his GSAE theory is a turn from the CWEC theory whose proponents "argue that the development of national educational systems and curricular categories is to be explained by universal models of education, state, and society, rather than by distinctive national factors" (p. 420).

According to Dale, proponents of the CWEC are attempting to use national educational systems as a resource for explaining a "hypothesized world culture" (p. 420). Dale differs in that curricular homogenization is the product of the influence of capitalism and a deep desire for countries to compete in a global economy. In particular, Dale argues for three regions of harmonized curricula– Europe, Asia, and America. Further, Dales argues that the impact of globalization is realized through indirect processes that can be traced. It is within this line of reasoning that we can begin to empirically test some of these theories.

Meyer, Kamens, and Benavot (1992) attempted to measure a hypothesized global educational curriculum empirically through a set of indicators derived from national curriculum statements; however, Dale (2000) identifies weakness in both the data and the analysis, leaving an area ripe for investigation. The current paper seeks to partially fill this hole with an internationally recognized robust data set and an appropriate yet unique analysis for empirically investigating the relationship between extra-national processes and education. Finally, Dale also offers a means for understanding how globalization may influence national educational policies and practices in the context of three key issues: *grasping the nature and force of extra-national effects*, *specifying what may be affected*, and *exploring how that effect might occur*.

Inspired by Dale's (2000) discussion and a paucity of empirical research in the field, this paper attempts to empirically investigate some of the effects of globalization on education. To do so, we consider how international mathematics curriculum (*what may be affected*) may be influenced by global processes (*the nature and force of extra-national effects)* as evidenced by item responses (*how that effect might occur*) to the mathematics portion of three TIMSS (Trends in International Mathematics and Science Study) administrations. The effects that we aim to measure are those most hypothesized by globalization scholars in education – most notably an increased interconnectedness and interdependence of cross-national educational policies. While the question of *how* globalization influences all aspects of education (and *who* controls these forces) is multidimensional and not completely testable, there appears to be some theories of globalization as it relates to education that can be empirically examined.

The phenomenon of international assessments in education has enabled researchers to examine what Dale (1999) explained as the widely acknowledged notion that globalization affects national policies in a range of areas to include educational outcomes. If in fact nations are borrowing educational policies, comparing educational systems, and setting educational benchmarks based on recommendations from an international agenda, we may reasonably see an increased similarity in students' responses on international educational assessments over time. In other words, an increase in similar curriculum should produce an enhanced similarity of how students respond to assessment items that are intend to measure curriculum.

In the last four decades a number of internationally comparative large scale assessments (LSAs) have been administered; however, two organizations currently lead the international LSA market: the International Association for the Evaluation of Educational Achievement (IEA) and the Organisation for Economic Co-operation and Development (OECD). The IEA's two main studies include the TIMSS, which measures mathematics and science achievement of 4^{th} and 8^{th} graders from more than 60 countries, and Progress in International Reading Literacy Study (PIRLS). PIRLS measures reading achievement of 4^{th} graders internationally. TIMSS is assessed every four years, while PIRLS is assessed every five years. The OECD's analogous assessment, Programme for International Student Assessment (PISA) is designed to measure mathematics, science and reading and is administered every three years. The most recent assessment administrations from both of these organizations include PIRLS 2006, TIMSS 2007, and PISA 2006.

While both organizations administer internationally comparable assessments, from a conceptual perspective *what* each organization aims to measure on its respective assessment is different. For instance, OECD/PISA (2007) describes their assessment as "measuring how well students can apply the knowledge and skills they have learned at school to real-life challenges." Further, OECD/PISA indicates that, "PISA does not test how well a student has mastered a school's specific curriculum" (online). In contrast, the IEA's assessments are based on a curriculum model, where the intended curriculum, the implemented curriculum, and the attained curriculum are considered. For the current study, it is worth reiterating the crucial difference in the intent of each assessment: PISA is designed to measure work-force knowledge, while IEA studies are designed to measure curriculum mastery. Given that our investigation involves the expected effects of globalization on educational policy and curriculum, both the TIMSS and PIRLS assessments are the most suitable for our needs. For the current study, we focus on TIMSS responses.

Following Dale's (2000) argument, it is reasonable to expect that what students know (based on curriculum) is one type of outcome for which an increase in similarity across countries may be expected. A finding to the affirmative suggests that extra-national forces may be rapidly harmonizing educational outcomes. It is also reasonable to expect that Dale's three hypothesized "regions," as defined above, may be intact. Both of these hypotheses are, to a limited extent, measurable. Employing the techniques explained in more detail below, this paper attempts to address the following research questions: To what extent have students' responses on TIMSS become more similar over time? Further, do we find empirical support for either the GSAE or the CWEC theories of world education harmonization? To this end, cross-national studies of educational achievement provide a rich source for empirically examining theories of globalization in education.

Methods

Data

In 2003 and 1995, TIMSS assessed students at both the 4th grade level and the 8th grade level (Martin, 2004). In 1999, TIMSS assessed students at the 8th grade level only. To allow for a consistent analysis across all three time points, the current study used responses to the 8th grade mathematics portion of the 1995, 1999 and 2003 TIMSS for those countries that participated in all three cycles. For the analysis, this leaves a pool of 16 economically developed and transitional countries and educational systems: Flemish Belgium (BFL), Cyprus (CYP), England (ENG), Hong Kong (HGK), Hungary (HUN), Iran (IRN), Japan (JPN), Korea (KOR), Latvia (LVA), Lithuania (LTU), New Zealand (NZL), Romania (ROM), the Russian Federation (RUS), Singapore (SGP), Slovak Republic (SVK), and the United States (USA). Excluded from these analyses were any countries that had been identified as having any type of sampling problems (as noted in Mullis, Martin, Gonzalez, & Chrostowski, 2004).

Analysis

We considered a number of methods for measuring similarity between countries with respect to their mathematics attainment. A simple comparison of scale scores over time was rejected as this would limit the investigation to asking *how much* respondents from a country know on a given topic. Instead, we chose an item-level analysis that allows for a closer examination of *what* respondents in a given country know and how this compares to other countries. This is not to say that our results avoid significant aggregation, similar to traditional methods of comparing scale scores. In fact, the methods we used, discussed below, subject the data to intense aggregation – going from the item level to a single measure of similarity for a given country. However, we adopt a slightly different perspective by generating country profiles and measuring the similarity between these profiles over time. Further, this type of analysis makes later item level comparisons feasible and possible.

For this analysis, we chose to use hierarchical cluster analysis and non-linear principal component analysis (NLPCA) via optimal scaling on TIMSS item responses for the three administration cycles mentioned. We chose cluster analysis given its applicability to our research question and the ready interpretation provided by the graphical representations available in cluster analysis, namely dendrograms. To compare and validate these findings we also chose to use an NLPCA, which optimally assigns continuous values to ordinal or qualitative data. Given that the measures we use, discussed in detail later, may suffer from a ceiling or floor effect, we reasoned that an approach that could relax the multivariate normality assumption necessary for many techniques would be useful.

Of interest for the current analysis are the item response patterns for each country on the mathematics items administered during three TIMSS cycles. To this end, we started with the percentage or proportion correct for each country on each item. However, as pointed out by other authors (Lie & Roe, 2003; Olsen, 2005; Zabulionis, 2001), using a raw proportion correct statistic as the basis for a measure of similarity would result in clusters based on overall performance on an item. That is, high-achieving countries would group with other high-achieving countries and low-achieving countries would group with other low-achieving countries. Based on a modification of the Linn and Harnisch (1981) difference score for detecting differentially functioning items, we used a statistic that cancels out the country effect and the item effect of a given item (Olsen, 2005). This statistic is discussed in detail below.

Following basic housekeeping of the data, the individual item responses were used to calculate item difficulty (classical proportion correct) for the multiple-choice items and a *credit* *proportion* for the constructed response items. This calculation was done separately for each of the TIMSS cycles (1995, 1999, 2003). The result was three matrices of item difficulties (rows) by country (columns) for each of the TIMSS administrations. After difficulties and credit proportions were calculated, each entry in each of the three matrices was then standardized to remove the item effect and the country effect. This was based on an ANOVA type model and was used to calculate the item residual for each item-by-country. The formula used for each item-by-country combination was of the form:

$$\varphi_{res(ic)} = \varphi_{ic} - \overline{\varphi} - (\overline{\varphi}_c - \overline{\varphi}) - (\overline{\varphi}_i - \overline{\varphi}) \quad (1)$$

Where $\overline{\varphi}$ = the mean over all countries and all items; φ_{ic} = the overall difficulty for country *c* on item *i*; $\overline{\varphi}_c$= the overall average item difficulty for country *c* over all items; $\overline{\varphi}_i$= the overall item difficulty for item *i* over all countries. The result for each country and administration was a vector of residuals that provides a sort of *fingerprint* for their achievement on all of the items during a given administration. For any country, the vector of residuals indicates how much better (positive residuals) or worse (negative residuals) than expected that country performed on each of the administered items. These country vectors combine to form a matrix of item-by-country interactions that is the basis for the subsequent cluster analysis. According to Olsen (2005), "these data are measures of how much the achievement for a country on an item deviates from what could be expected given the overall achievement of the country and the overall difficulty of the item" (p. 2).

Figure 1 illustrates the basic form of the item-by-country residual matrix used in later cluster analyses. Here, each column represents a given country and each row represents a given item. Each cell entry represents the residual on a particular item for a country. In other words, each entry indicates how much better or worse than expected a country performed on a particular item. Looking down a particular column, we can see if a country did generally worse or better than expected. Similarly, the entries across a row indicate how the set of countries performed on an item.

Given the large dimensions of the matrix, it is a difficult enterprise to make sense of these data by eyeballing them; however, cluster analysis provides a visually simple and interpretable method for understanding the multivariate relationships in the data.

Insert Figure 1 about here

For the cluster analysis, each of the item-by-country residual matrices produced for all three TIMSS administrations were used to generate a correlation matrix across the columns. This correlation matrix is then used as the basis for the cluster analyses. For all 16 countries at each of three time points, separate dendrograms were created for mathematics responses. This process resulted in one dendrogram for each of the TIMSS cycles (1995, 1999 and 2003). Producing dendrograms in this way allowed for a comparison of response patterns across subjects and over time. Of interest were the intermediate groupings on each time point. Those countries in a given cluster were considered to be more similar to each other and, at the same time, distinct from countries outside of the given cluster.

Next, we optimally scaled (Meulman, Van der Kooij & Heiser, 2004) the item-by-country residuals. Given that we are working with residuals derived from proportion correct, the process of optimal scaling assigns to each residual a value that allows us to use a traditional principal component analysis. This method is typically used for categorical or ordinal variables; however, given the restricted range of our measures, we reasoned that this was a conservative approach to validating the findings from the cluster analysis. Findings from this analysis are presented as scatter plots of the principal components. Similar to dendrograms, this type of graphical representation provides an accessible interpretation to otherwise strictly numerical results.

For a given administration on a particular topic, we expect that cultural, linguistic, or geographic patterns would emerge (Gronmo et al., 2004; Kjaernsli & Lie, 2004; Lie & Roe, 2003; Olsen, 2005; Zabulionis, 2001). Further, if globalization is having theorized effects on educational policy, we should see that the number of clusters will decrease over time as countries grow more homogeneous. Alternatively or additionally, we may expect to see a *tightening* of existing clusters. That is, the node heights in a given dendrogram may become shorter over time indicating an increase in similarity within a cluster structure.

Results

Results for both types of analyses follow. We first present the findings from the cluster analysis, followed by the NLPCA results. The results of both types of analyses were essentially the same – the same *groups* of countries clustered together regardless of the approach used.

Cluster Analysis

For those countries who participated in all three cycles of TIMSS, the cluster results are included in Figures 2, 3 and 4. Figure 2 deals with the 1995 administration of TIMSS mathematics items, Figure 3 is for the 1999 administration, and Figure 4 is the dendrogram that relates item responses on the 2003 TIMSS mathematics administration. In Figure 2, we can see that response patterns for these mathematics items follow a distinct geographic, cultural and linguistic pattern. For instance, based on an R-squared value of 0.50 (the amount of variance accounted for by this cluster structure), evidence suggests four distinct clusters. From the top of the figure to the bottom, the first cluster, comprised of Cyprus, Romania, Hungary, Latvia, Lithuania, the Russian Federation, and Slovakia, suggests an Eastern European cluster. In this grouping, Latvia and Lithuania is the most similar pair of countries followed by Russia and Slovakia. Next is a single country cluster that contains Iran. No other educational systems are similar to Iran until much higher up the dendrogram, where Iran joins the Eastern European countries.

The next distinct cluster in this dendrogram includes Hong Kong, Japan, Korea, Singapore and Belgium (Flemish). This cluster is primarily an East Asian group with the addition of Flemish Belgium. Finally, an English speaking cluster emerges, which contains the United States, New Zealand and England. Notable here is the extreme similarity of New Zealand and England, who comprise the most similar country pair in this administration of TIMSS mathematics items.

Insert Figure 2 about here

In Figure 3, where the analysis includes all mathematics items administered in 1999, the same general patter emerges; however, we see a notable change from the 1995 administration—Iran joins with Cyprus to form a distinct cluster. At an R-squared value of 0.50, we see evidence of four distinct clusters. From the top of the dendrogram to the bottom, the clusters are as follows. Next, the Eastern European cluster remains intact; however, the most similar country pair is now Latvia and the Russian Federation. Slovakia, who in 1995 was most similar to Russia, was most similar to Hungary in 1999. Further, with Hungary and Flemish Belgium, it occupies an entirely different sub-cluster than the rest of the Eastern European clusters. This indicates that while these Eastern European countries all join to form a cluster later in the dendrogram, Hungary and Slovakia demonstrate a relatively high-degree of external isolation from their other Eastern European neighbors.

Again, New Zealand and England join with the United States to form an English-speaking cluster. Also, New Zealand and England maintain their status as the most similar pair in the 1999 administration of mathematics items. Finally, the East Asian cluster remains intact with some slight changes. There is some *tightening* of the sub-cluster that contains Japan, Korea and Hong Kong. In particular, Hong Kong has grown more similar to Japan and Korea in terms of mathematics item response patterns. As in the 1995 administration, Singapore remains relatively distinct from its Asian neighbors.

Insert Figure 3 about here

In the most recent administration of TIMSS, the cultural, linguistic and geographic clusters continue to hang together as in the previous two administrations. A few notable points, discussed in detail, result from this analysis. At an R-squared value of 0.50, the following clusters are evident from top to bottom in Figure 3. First, we see that Romania joined the cluster occupied by Iran and Cyprus. These countries then join the rest of the Eastern European countries to form an Eastern European cluster plus Iran and Cyprus. As in 1999, Latvia and Russia form the most similar pair of countries followed by the pair that includes Hungary and Slovakia. In the 2003 analysis, the English-speaking cluster absorbed Flemish Belgium; however, this addition did not change the marked similarity between New Zealand and England. Finally, the Asian cluster remained intact; however, the node for this cluster is precisely at the R-squared value of 0.50. This suggests that the pair including Hong Kong and Singapore may distinguish itself in terms of mathematics item response patterns from the pair that includes Japan and Korea. Further, Japan and Korea are less similar in terms of their mathematics response patterns in 2003 than was the case in 1999.

Insert Figure 4 about here

Non-Linear Principal Component Analysis

We fit a number of NLPCA models to the data and found that a three-component solution was the most interpretable while accounting for a reasonable amount of variance. For those countries analyzed, plots of the NLPCA results are included in Figures 5 through 7. Figure 5 contains principal component plots of the 1995 administration of TIMSS mathematics items, Figure 6 is for the 1999 administration, and Figure 7 contains the graphic representations that relate item responses on the 2003 TIMSS mathematics administration. While all three principal component plots are informative, the most important (based on variance accounted for, by definition) and interpretable are scatter plots of principal components 1 and 2.

Regarding the 1995 NLPCA of TIMSS item-by-country interactions, we see that these results are largely in line with the cluster analysis findings. In particular, the plot of principal components one and two in Figure 5 illustrates the three groupings that were evident in the earlier analysis. Namely, we found an English speaking group, a largely Eastern European group and an Asian group. The other two principal component plots also largely substantiate these findings; however, there was some marginal interspersion of the Eastern European countries with the English-language group and the Asian group.

Insert Figure 5 about here

The NLPCA of the 1999 TIMSS data resulted in slightly different findings than the 1995 NLPCA. As illustrated in the plot of principal component one and two in Figure 6, the Eastern European group is more dispersed than in the cluster analysis. Additionally, the United States drifted from its fellow English speaking countries and instead grouped with Iran and Cyprus. The other two plots provide evidence for a persistent English-speaking group and we see that the Asian group remained cohesive in this analysis.

Insert Figure 6 about here

Finally, the NLPCA results for the 2003 data are consistent with the cluster analysis findings. In Figure 7, there is strong evidence for an Asian group, an English-speaking group and a fairly cohesive Eastern European group. In this wave of the analysis, Cyprus and Iran remain largely independent from any of the other groups. This finding is slightly different than what we found in the cluster analysis; however, it is notable that the overall similarity of Cyprus and Iran remained intact for both analyses that used 2003 data.

Insert Figure 7 about here

Discussion

This paper attempted to investigate similarities of student responses as they pertain to educational curriculum evidenced by item responses to the mathematics portion of three TIMSS administrations. The effects that we hoped to measure were the extent to which students' responses on TIMSS have become more similar over time. Using an item level analysis, versus comparing achievement means across countries, we were able to illustrate possible trends over time in terms of global curriculum harmonization as hypothesized by CWEC scholars. Using hierarchical cluster analysis and NLPCA on item-by-country residuals of mathematics item responses, we tentatively show that student responses tend to demonstrate regional groupings. Due to fairly stable regional groupings, results imply that there is little evidence that the CWECs theory of proposed global educational curriculum convergence is observable in student responses to assessment questions used in this study.

As varied as globalization and education literature is, there is some consensus that global processes are affecting national educational policy and that policy borrowing is taking place at an increasing rate. In addition, we assumed that increased educational assessment and comparison would tend to homogenize national educational systems. Of course, global processes do not operate in a vacuum and the same global pressure can be interpreted differently from country to country or person to person. This idea was summarized nicely by Dale and Robertson (2002) when they wrote, " globalization represents a complex, overlapping set of forces, operating differently at different levels, each of which was separately set in motion intentionally, though their collective outcomes were not uniform, intended, or predicted" (p. 11). This point is also crucial to Dale's GSAE approach.

Our analysis found that regional groups[1] existed at the base year of 1995 and continued over 1999 and 2003. In fact, these are quite similar to regional groupings hypothesized by Dale's GSAE theory. Rather than students' responses becoming more similar over time, regional clusters remained intact over the three administrations. With only small changes over time, an English speaking cluster, an Asian cluster, and an Eastern European cluster were consistently found on all analyses. Our research question addressed global similarities of educational curriculum and further assumed that global forces are affecting curriculum which can either be explained by the CWEC or the GSAE. The existence of three clusters in 1995 and the relative similarities of these clusters over time appear to support Dale's GSAE theory.

[1] Regions as defined here do not directly correspond to a geographical space. Rather, regions are defined as a set of commonalities to include geographical space, language, and culture.

It is not certain that regional groupings as defined here are in opposition to the CWEC's overall argument that institutions are the forces affecting national curriculum only that these global institution are not changing curriculum across the board. For example, a unifying shift towards world regionalization can be seen as a global process. However, the lack of change in students' similarities between and within regions suggests that there is little evidence within this data set to indicate a movement towards any single set of common factors. Dale and Robertson (2002) stated that when examining globalization "it is necessary to ask whether the transnationally initiated effects…are leading in the direction of a convergence of education systems and/or policies and/or practices" (p. 36). The evidence from TIMSS seems to imply that if in fact curriculum is changing around the world what students know in line with curriculum is similar at the regional level not the global level.

Limitations

Limitations to this study are as follows. First, the data used for this study are not causal in nature; instead, we are limited to an observational view of trends over time. Also, while we believe that this was the best method to answer the research question of interest, the intense aggregation to which the data must be subjected may result in other findings being obscured. To put this in perspective, consider that we began with a data set comprised of several hundred thousand observations on hundreds of items. Using the methods discussed earlier, we reduced this large data set to a similarity matrix of the size 16 x 16 in the case of the 3-cycle trend countries. One way to alleviate the degree of aggregation and a possible area for further research may be to reanalyze these data by the TIMSS content domain subscales or the cognitive domain subscales. Interested readers are advised to consult the TIMSS 2003 Technical Report (Martin, Mullis & Chrostowski, 2004) for a detailed discussion of these subscales. This may shed light on the possibility that certain areas of learning are receiving more global scrutiny and importance than others.

Further, given that we selected countries who participated in the TIMSS study since 1995, we were limited to a reduced number of countries that excluded some geographic areas of interest. Finally, two contrasting factors may limit this study. A longer time period may be ideal for identifying the effects of policy harmonization on curriculum and changes to curriculum on item responses. Or it may be the case that the first year of the study, 1995, may have already been too late to adequately detect changes that arose due to global forces. The Internet, international trade, and global travel, while all experiencing marked advances in the past 12 years, were well established by the first TIMSS administration and were fast becoming a dominant feature of the global landscape. Also, events with significant world impact such as the North American Free Trade Agreement and the fall of communism in Eastern Europe were well underway by 1995. Given these limitations, it is worth noting that the robustness of the data and the consistency of the results over time suggest that our findings are valid and worth further investigation.

Areas for Further Research

As discussed in the introduction, TIMSS is designed to measure an international curriculum while PISA is designed to measure work-force knowledge. Given that globalization is said to strongly affect international economic processes, including the labor market, it may be that items intended to measure the knowledge of future workers would be more susceptible to the forces of globalization. As such, it may be useful to repeat the present analysis with responses to mathematics items on the PISA assessment.

An additional area for consideration deals with individual items across countries. For instance, we see that over time, an English-speaking cluster remains intact. It may be worth considering what types of items are similar for the countries that grouped together. Are there cognitive domains or content domains that drive the similarities within cluster groups? This could be answered using a cluster analysis on items (instead of countries) for sets of countries that grouped together. Finally, it may also be of interest to consider item response patterns on the PIRLS assessments.

Conclusion

Educational processes and systems are complex and any attempt to measure them, especially at this level of aggregation can only lead to broad and general discussions. However, we contend that this discussion is necessary. Further, attempting to measure global processes in education may provide another puzzle piece to theoreticians as well as national and local policy makers, who are working at understanding and, in the latter case, steering educational systems. We recognize that this paper has worked at a very abstract and general level; however, we contend that the methods and data lend themselves well to a fruitful discussion of global processes' effects on education.

References

Dale, R. (1999). Specifying globalization effects on national policy: A focus on the mechanisms. *Journal of Education Policy, 14*(1), 1-17.

Dale, R. (2000). Globalization and education: Demonstrating a 'common world educational culture' or locating a 'globally structured educational agenda'? *Educational Theory, 50*(4), 427–448.

Dale, R., & Robertson, S. L. (2002). The varying effects of regional organizations as subjects of globalization of education. *Comparative Education Review, 46*(1), 10.

Green, A. (1999). Education and globalization in Europe and East Asia: Convergent and divergent trends. *Journal of Education Policy*, *14*, 55–71.

Gronmo, L. S., Kjaernsli, M., & Lie, S. (2004). Looking for cultural and geographical factors in patterns of responses to TIMSS items. In C. Papanastasiou (Ed.), *Proceedings of the IRC-2004 TIMMS* (Vol. 1, pp. 99-112). Lefkosia: Cyprus University Press.

Kjaernsli, M., & Lie, S. (2004). PISA and scientific literacy: Similarities and differences between the Nordic countries. *Scandinavian Journal of Educational Research, 48*(3), 271-286.

Lie, S., & Roe, A. (2003). Unity and diversity of reading literacy profiles. In P. L. Svein Lie, and Astrid Roe (Eds.), *Northern lights on PISA: Unity and diversity in the Nordic countries in PISA 2000*. Oslo: University of Oslo, Norway.

Linn, R. L., & Harnisch, D. L. (1981). Interactions between item content and group membership on achievement test items. *Journal of Educational Measurement, 18*(2), 109-118.

Martin, M. O. (Ed.) (2004). *TIMSS 2003 user guide for the international database.* Boston: TIMSS & PIRLS International Study Center, Boston College.

Martin, M. O., Mullis, I. V. S, Chrostowski, S. J. (Eds.) (2004). *TIMSS 2003 technical report.* Boston: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Boston: TIMSS & PIRLS International Study Center, Boston College.

Meulman, J. J., Van der Kooij, A. J., & Heiser, W. J. (2004). Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data. In: D. Kaplan (ed.), *Handbook of quantitative methods in the social sciences*, (pp. 49-70). Newbury Park, CA: Sage Publications.

Meyer, J. W., Kamens, D., & Benavot, A. (1992). *School Knowledge for the Masses: World Models and National Primary Curricular Categories in the Twentieth Century*. London: Falmer Press.

OECD/PISA. (2007). *FAQ: OECD/PISA*. Retrieved 08/07, 2007, from http://www.pisa.oecd.org/document/53/0,3343,en_32252351_32235731_38262901_1_1_1_1,00.html.

Olsen, R. V. (2005). An exploration of cluster structure in scientific literacy in PISA: Evidence for a Nordic dimension? *NorDiNa, 1*(1), 81-94.

Zabulionis, A. (2001). Similarity of mathematics and science achievement of various nations. *Education Policy Analysis Archives, 9*(33).

Figure 1. Example of an item-by-country residual matrix.

	$cntry_1$	$cntry_2$	$\dots$	$cntry_n$
i_1	β_{11}	β_{12}	$\dots$	β_{1n}
i_2	β_{21}	β_{22}	$\dots$	β_{2n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i_m	β_{m1}	β_{m2}	$\dots$	β_{mn}

Figure 2. Dendrogram for TIMSS 1995 mathematics items.

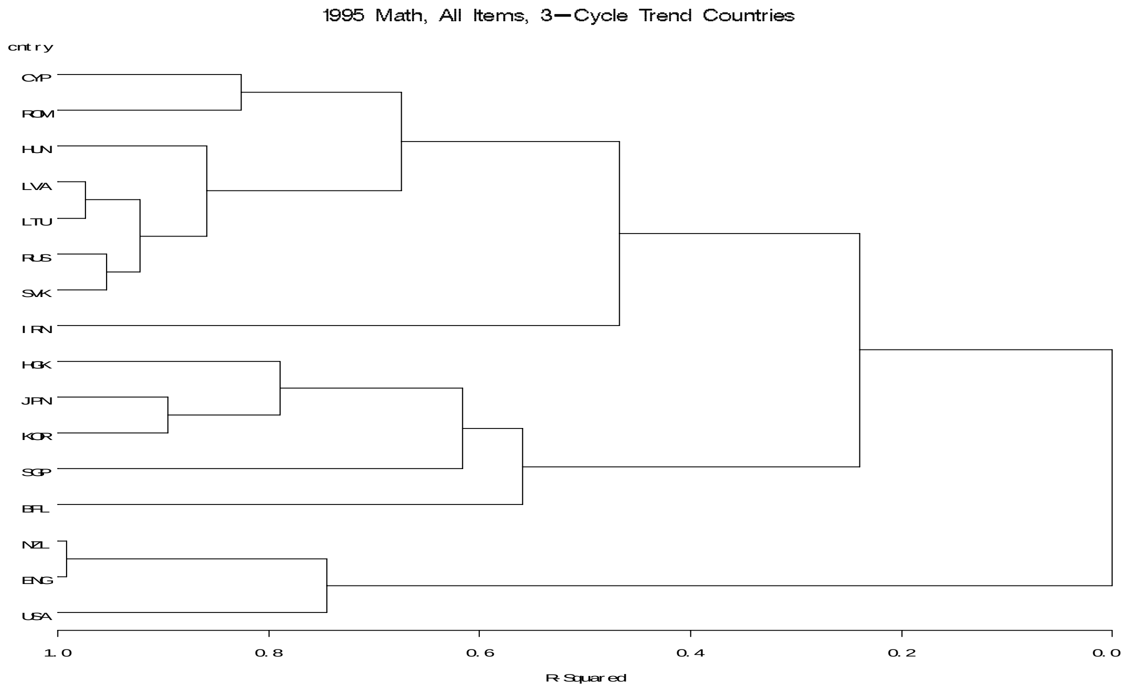
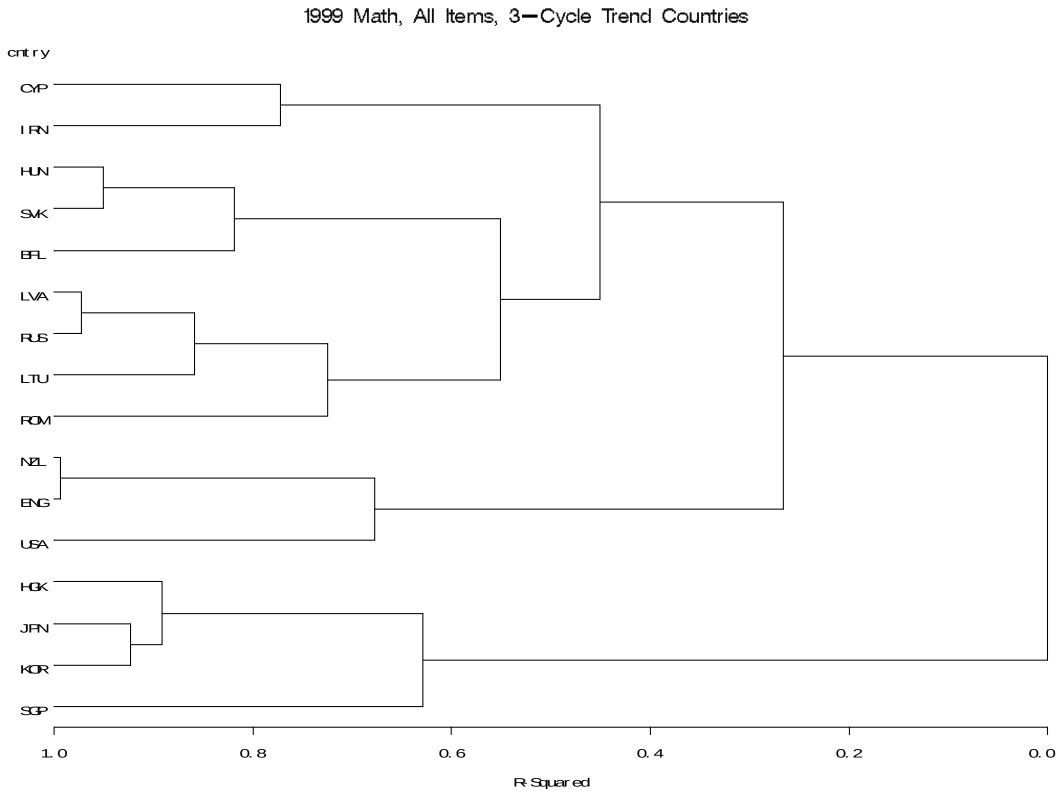

1995 Math, All Items, 3–Cycle Trend Countries
cntry
CYP
ROM
HUN
LVA
LTU
RUS
SVK
IRN
HGK
JPN
KOR
SGP
BFL
NZL
ENG
USA
1.0
0.8
0.6
0.4
0.2
0.0
R-Squared

Figure 3. Dendrogram for TIMSS 1999 mathematics items.



1999 Math, All Items, 3−Cycle Trend Countries
cntry
CYP
IRN
HUN
SVK
BFL
LVA
RUS
LTU
ROM
NZL
ENG
USA
HGK
JPN
KOR
SGP
1.0
0.8
0.6
0.4
0.2
0.0
R-Squared

Figure 4. Dendrogram for TIMSS 2003 mathematics items.

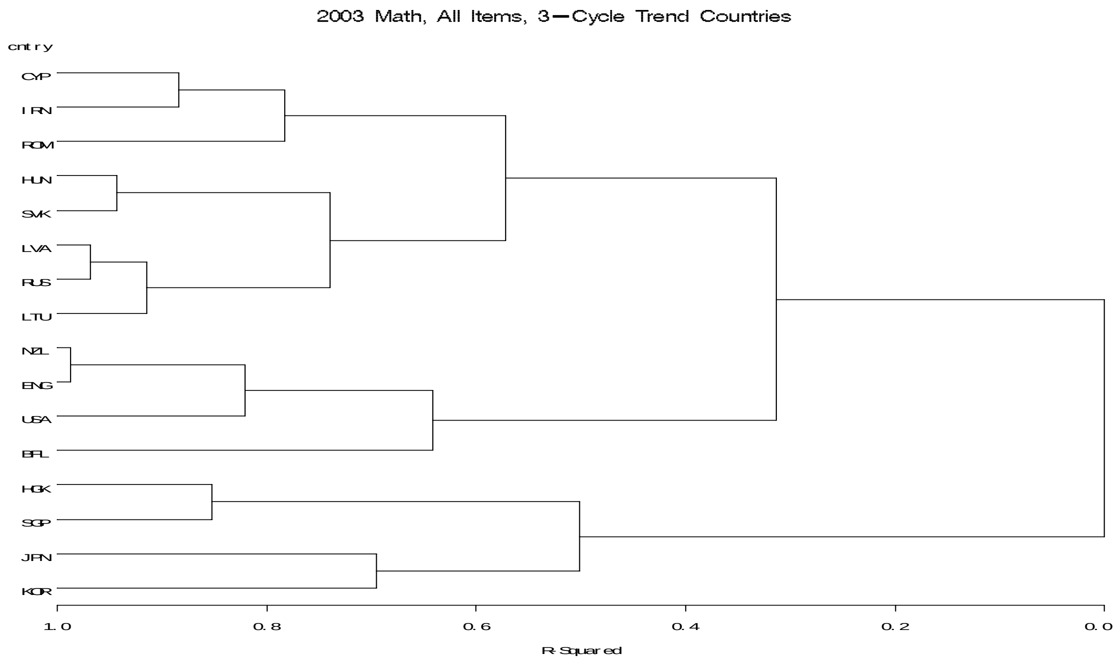
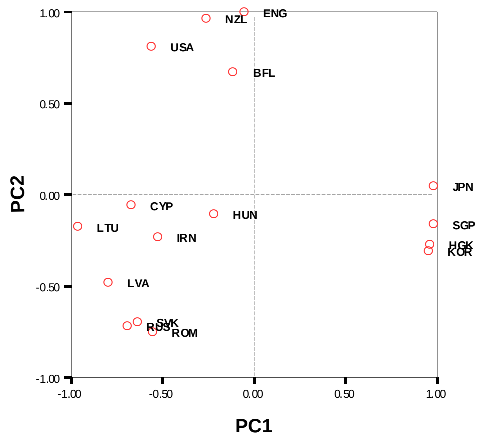
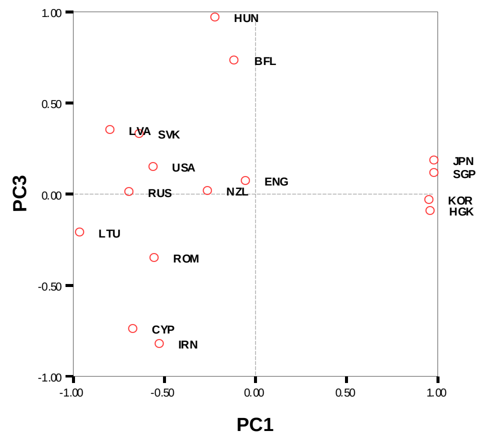

2003 Math, All Items, 3—Cycle Trend Countries
cntry
CYP
IRN
ROM
HUN
SVK
LVA
RUS
LTU
NZL
ENG
USA
BFL
HGK
SGP
JPN
KOR
1.0
0.8
0.6
0.4
0.2
0.0
R-Squared

Figure 5. Plot of principal components for 1995 analysis.


PC2
PC1
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
ENG
NZL
USA
BFL
JPN
CYP
HUN
LTU
IRN
SGP
HGK
KOR
LVA
SVK
RUS
ROM


PC3
PC1
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
HUN
BFL
LVA
SVK
USA
ENG
JPN
SGP
RUS
NZL
KOR
HGK
LTU
ROM
CYP
IRN

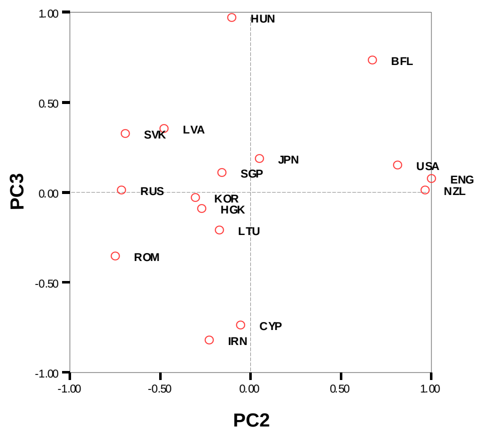
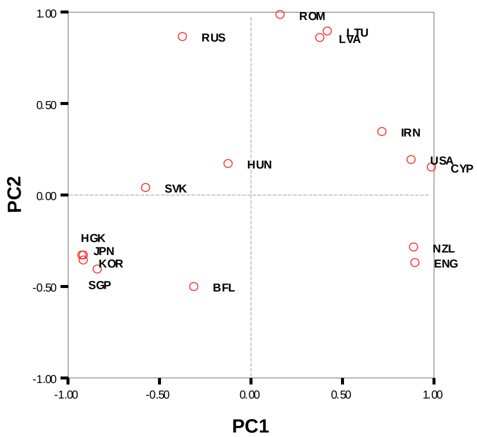
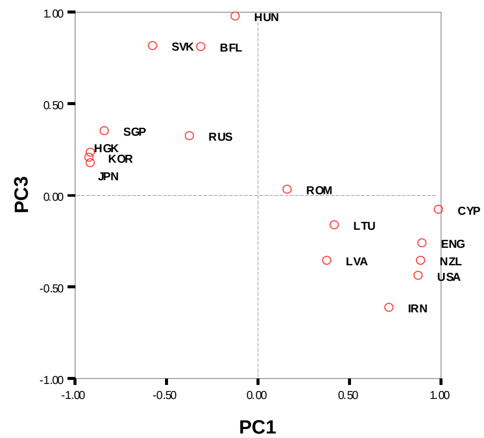

PC3
PC2
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
HUN
BFL
SVK
LVA
JPN
SGP
USA
ENG
NZL
RUS
KOR
HGK
LTU
ROM
CYP
IRN

Figure 6. Plot of principal components for 1999 analysis.


PC2
PC1
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
ROM
RUS
LTU
LVA
IRN
HUN
USA
CYP
SVK
HGK
JPN
KOR
SGP
BFL
NZL
ENG


PC3
PC1
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
HUN
SVK
BFL
SGP
RUS
HGK
KOR
JPN
ROM
CYP
LTU
ENG
LVA
NZL
USA
IRN

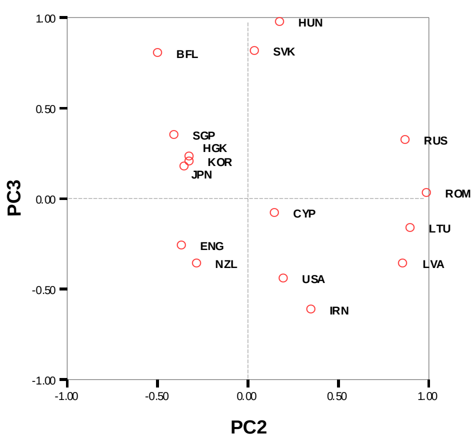
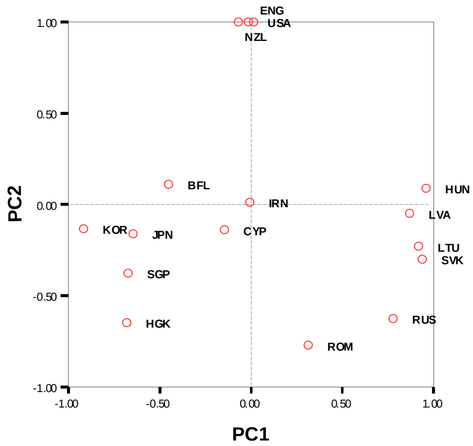
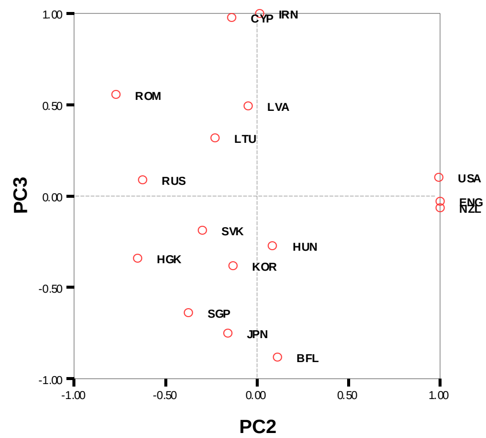
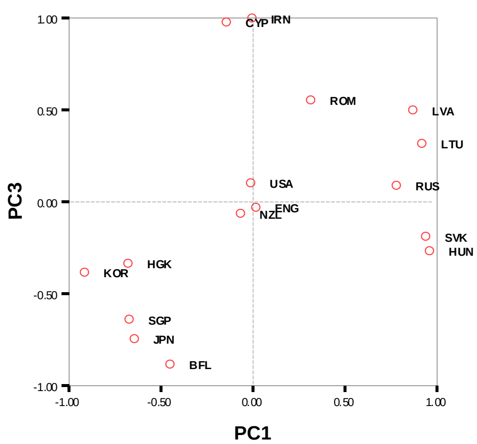

PC3
PC2
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
HUN
BFL
SVK
SGP
HGK
KOR
JPN
RUS
ROM
CYP
LTU
ENG
NZL
LVA
USA
IRN

Figure 7. Plot of principal components for 2003 analysis.


ENG
USA
NZL
BFL
IRN
HUN
LVA
KOR
JPN
CYP
LTU
SVK
SGP
HGK
RUS
ROM
PC2
PC1
1.00
0.50
0.00
-0.50
-1.00


CYP
IRN
ROM
LVA
LTU
RUS
USA
ENG
NZL
SVK
HUN
HGK
KOR
SGP
JPN
BFL
PC3
PC2
1.00
0.50
0.00
-0.50
-1.00


CYP
IRN
ROM
LVA
LTU
USA
RUS
NZL
ENG
SVK
HUN
HGK
KOR
SGP
JPN
BFL
PC3
PC1
1.00
0.50
0.00
-0.50
-1.00