



IARTEM *e-Journal* Volume 4 No 2

Textbook approval systems and the Program for International Assessment (PISA) results: A preliminary analysis

Hendrienne J. Wilkens
Utrecht University
Netherlands

Abstract

This paper examines the textbook approval systems of various countries in relation to educational outcomes. This investigation is based on research in the Netherlands that aimed to gain an overview of textbook approval systems in use across the world. The study also looked at the Program for International Assessment PISA results, in order to determine which countries have high and low PISA results. The researchers also sought to determine whether there was a relationship between state influence, as reflected and formalised in textbook approval systems, and student achievement educational outcomes.

It is hoped this investigation will be a starting point for ongoing discussion about the usefulness of comparing textbook approval systems, and the possible relationships between national approval systems and educational outcomes.

Key words

Textbooks, (Textbook approval, approval systems, PISA, state influence on textbook systems, educational resources)

PISA and policy

PISA (Programme for International Student Assessment) is carried out every three years by the Organisation for Economic Cooperation and Development (OECD).

The general purpose of PISA is to assess the extent to which 15 year old students in OECD (and other) economies have acquired the appropriate literacies in reading, mathematics and science to make a significant contribution to their societies. PISA assessment takes place every three years with approximately 400,000 students internationally. PISA attempts to measure the essential knowledge and skills that 15 year old students require to meet 'the challenges of our society.' The 15 year old age group was chosen for PISA because 15 year olds are close to completing compulsory secondary education. The PISA assessment is large scale, complex and growing. Originally administered in OECD countries, the first PISA in 2000 involved 32 participating countries, rising to 57 countries in 2006 and 65 countries in 2009. Increasingly, countries outside OECD membership wish to participate in PISA as a benchmarking exercise to gain an indication of the strengths and weaknesses of their educational systems' performance.

The aim of PISA is to identify factors that influence student learning and achievement at the student, family and school levels. These factors can then be analysed extensively and reported on in the international and national education communities. Once the PISA test results have been scored, they are correlated with student and school data. PISA uses advanced correlation and multivariate statistical procedures to establish relationships between student and school characteristics and student achievement on PISA tests. After each PISA evaluation (2000, 2003, 2006, 2009, 2012) the OECD releases a major survey of international performance on PISA. The OECD also publishes major research reports on the characteristics of the students, schools and education systems that contribute to student achievement and learning; and also country reports where each nation's performance is also analysed in relation to its educational structures and policy settings. Finland, Korea and Japan are the leading OECD performers in PISA 2000, 2003, 2006. In 2009 Hong Kong, Shanghai and Singapore joined Finland and South Korea as the leading PISA performers.

Publication and release of PISA three yearly results (2000, 2003, 2006, 2009, 2012) are much anticipated events, nationally and internationally. Interpretations of the meaning of results are employed to

- assess and compare education systems' performances;
- directly influence policy directions; often with a sense of urgency (as the results and their interpretations are often accompanied by solutions and potential policy directions);
- provide public commentary on current and potential educational policies, as a result consequence of the high media profile given to international comparisons; and
- provide a discourse that frames policy and educational debate with an emphasis on benchmarking, comparison and quality through testing, ranking and rating.

Research question

The research question for this study was: Is there a relationship between the degree of formal state influence on textbooks, and PISA results?

While this question focuses on the relationship between state influence and PISA-results, it is also important to have some ideas about which variables are responsible for this relationship. These issues were discussed during the Eleventh Biannual Conference of the International Association for Research on Textbooks and Educational Media (IARTEM) in Kaunas, Lithuania (2011). Generally this research question explores the relationships between educational resources (and therefore funding) and student achievement. However, there are consistent findings which indicate that material education resources, have at best a relatively small effect on the kind of standardised test performance that PISA represents (Hanushek, 2003, 2006) For example the top PISA performers in 2009, Shanghai, Hong Kong, Singapore and South Korea have large class sizes and spend less on education per student than other nations undertaking the PISA test. Finland and Japan also in the top PISA performers, also spend considerably less per student than countries such as the US and the Netherlands (Horsley 2011).

One of the problems in conducting correlation type research that explores this relationship is that we have very imperfect knowledge of what actually works in promoting student achievement. Most of the research based on analysing PISA results ignores teaching and learning resources such as textbooks.

Grubb (2008) has argued that money may be necessary to provide a certain level of school resources, but resource use is actually constructed in a school by many different individuals. This research has highlighted the number of variables related to pedagogical practices and teacher attitudes that seem to influence student achievement scores (Hattie, 2009). According to Grubb (2008), the most effective teaching and learning resources are compound, complex and abstract, some of them due to combinations of resources, some of them (such as school climate) embedded in a web of expectation and personal relations within schools, and some of them (such as streaming and different levels of curriculum) reflecting a complicated mixture of self-selection, curricular and pedagogical practices and teacher demands. Print and digital textbooks and other teaching and learning resources are modified, adapted and customised by teachers to produce classroom teaching and learning materials. This process of resource construction, (Grubb 2008), depends on both the level of classroom teaching and learning resource inputs and the ways that teachers turn these inputs into classroom resources that support learning.

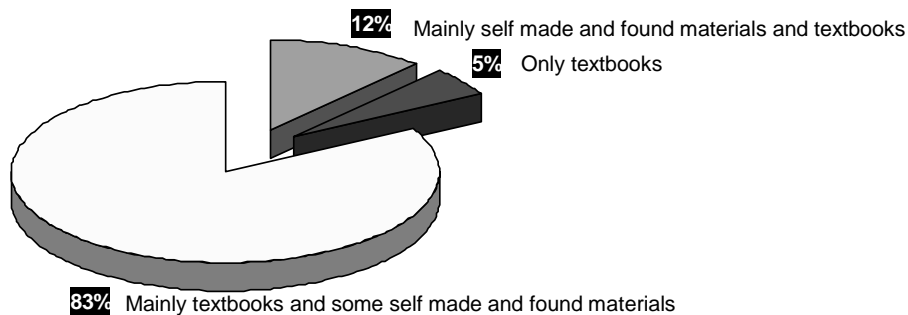
Input resources must be *constructed* by school leaders and teachers working together, by the kinds of sustained staff development that are the only ways to improve teaching practices. It may be that textbook approval systems play a critical role in this process by collecting data about the way that textbooks can afford learning from teachers and incorporating this data into textbook design and development. More recently Kennedy (2010) has argued that in considering teaching quality, it is best to explore the interaction between the way that teacher characteristics and the school, classroom and situation characteristics interact to form and structure such practices. In particular, textbook approval systems may

increase the amount of teacher planning time, the quality and quality of the teaching and learning resources and teachers' workloads all to positively influence teachers' practices and student learning.

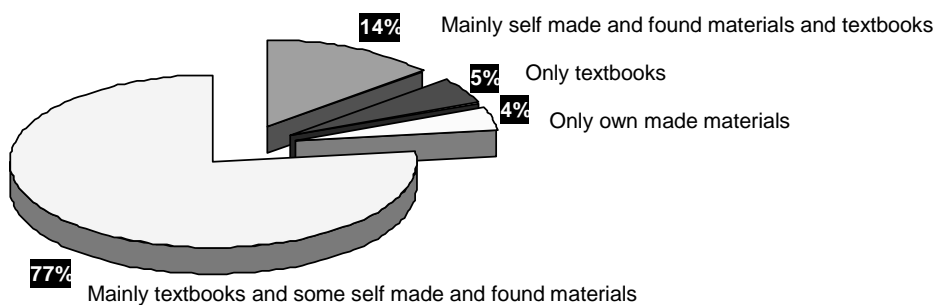
Why study textbook approval systems?

The reason textbook approval systems need to be studied is that in most countries textbooks are very important in schools. In the Netherlands, for example, it can be seen that most teachers mainly use textbooks, complemented with materials they find on the Internet (or elsewhere) or self-made materials (see figure 1). In secondary school:

- 77 % of the teachers mainly use textbooks in combination with own materials
- 5 % of the teachers use only textbooks
- 14 % use mainly self found and self made materials and textbooks
- Only 4 % of the teachers use only their own learning materials



Primary education



Secondary education

Source: SLO, 2010-2011

Figure 1: The use of textbooks in primary and secondary school

These same percentages can be seen in various countries. However, the reasons for investigating textbook approval systems do not only concern quantitative data about usage: qualitative arguments play an important role. In our view, textbooks are at least as important as teachers when it comes to the support of learning. Some supporting arguments follow:

- By the nature of teaching, teachers tend to be more activity-oriented than goal-oriented; aspects of curriculum design and structure can therefore be met better by textbooks than by teachers.
- Teachers planning time is limited and high quality resources which have been quality assured may provide more planning time for teachers to plan lessons – compared to nations where teachers prepare their own teaching and learning resources (Kennedy 2010)
- It is impossible for teachers with 25 or 30 pupils in their classrooms to adapt lessons to the needs of each of these 25 or 30 individuals, while digital materials at least have the potential to do so (Horsley and Walker 2012).
- The pedagogical challenges for teachers are increasing, which makes an appropriate pedagogical approach by teachers more necessary. Teachers nowadays are more and more involved in a coach-like role; it is therefore very important that textbooks at least partly take over the typical didactical and instructional role of teachers.
- Teachers are not always certified (at least in the Netherlands), capable, or present when pupils work with their textbooks (for example, at home, or when they have been ill and have to make up for unattended lessons). Horsley and Walker have claimed that there is very little research on the teaching and learning resources that are the basis of homework (2012). Sikorova found out that textbooks hardly play a substantial role for pupils in doing their homework (2011).

Perhaps most importantly, the quality of textbooks can be subjected to quality assurance and be developed to strengthen their affordances, and limit their constraints in contributing to students learning.

Why do approval systems differ?

Why do some countries have a textbook approval system and others not? In some countries approval systems are used referring to the right of parents, or the society as a whole, to a good education for their children, and that the government must guarantee this (Ninomiya, 2012 in press). This has become a more important priority as a result of PISA analysis which distinguishes between high and low achievement nations with high and low social equity. Exploring social equity measures is seen as an important aspect of PISA analysis. In other countries a textbook approval system aims to ensure textbooks cover the National Curriculum (Sousa & Lourdes Dionísio 2012, in press).

Countries without a formalised approval system legitimise this absence by arguing that teachers are in the best position to decide what is good for their pupils and what is not – including the selection, adoption and use of textbooks. It seems evident that, in this case, capable teachers are required.

It seems that the reasons why some countries have an approval system and others do not involves broader aspects of the educational and political systems of these countries.

Typology of approval systems

Figure 2 shows the typology of approval systems developed by Repoussi and Tutiaux-Guillon (2010). The typology is based along several dimensions:

- whether or not there is any formal state textbook approval
- if the approval has a prescriptive or a recommending status
- if there is any fixed number of different textbooks per subject that schools are allowed to choose from

On the basis of these three dimensions they distinguish the following five different textbook approval systems or models:

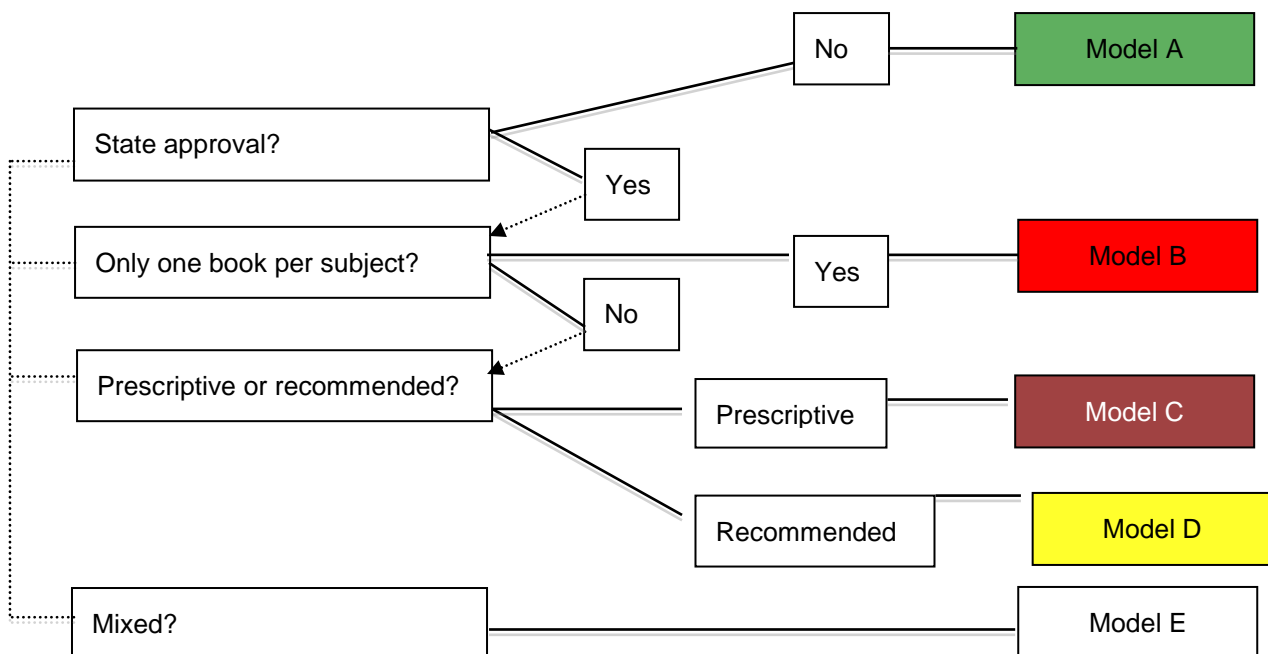


Figure 2: Typology of approval systems (Repoussi & Tutiaux-Guillon, 2010)

The first question to be asked in this typology is: Is there state approval? If no, the country belongs to model A.

If yes, the next question to be asked is: Is there only one book per subject approved? If yes, the country belongs to model B.

If no, the question is asked: Are the books are prescribed or recommended? If prescribed, the country belongs to model C. If recommended, in model D.

In some countries there is a mixed form, mostly due to regional differences. These countries belong to model E.

Why correlate approval systems with PISA-results?

This investigation will focus on the Program for International Student Assessment (PISA). PISA is undertaken by the Organisation for Economic Co-operation and Development (OECD). The assessment takes place in 65 countries every three years, with 15-year-old students. The students are assessed in reading, maths and science. On each occasion, the emphasis is on a different subject: maths in 2003, science in 2006, reading in 2009.

What are the arguments for studying the relationship between textbook approval systems and PISA?

The literature suggests a relationship between the two. Some researchers claim that a strong state influence does not sufficiently take into account the differences between pupils (Khutorskoi, 2006; Ogan-Bekiroglu, 2007). However, it is important to take these differences into account: to motivate students, to adapt to their different learning styles, to challenge students, and to adjust to slow and fast learners (Entwistle, 1975; Jonassen and Grabowski, 1993; Pratt, 2002; et. al.).

Other researchers argue that a strong state influence hampers the creativity of the teacher. Creative teachers are important because they feel more responsibility and ownership for their lessons (Valverde et. al., 2002). It is also claimed that PISA assessments reflect more innovative didactical ideas than the stiff criteria of the textbook approval: therefore, PISA assessments can motivate teachers to use new didactical ideas in daily classroom life (Pingel, 2010).

So it may be concluded that a strong state influence may hamper effective learning.

It therefore seems logical to determine if there is a relationship between approval systems and PISA results. This leads to the hypothesis that Textbook Approval Systems reflecting a strong state influence are connected to lower PISA results than Textbook Approval Systems without any official state influence.

This leads to the following research questions:

- Is there a difference in PISA results between models A (no influence) and B, C, D, E (with more or less influence)?
- Is there a difference in PISA results between models B (strong influence) and C, D, E (moderate and weak influence)?
- Is there a difference in PISA results between models B (strong influence) and C (moderate influence)?
- Is there a difference between PISA-results in reading, maths and science for each model?

Methodology

A study was conducted of publications and other documentation. In addition, representatives of ten countries were asked to complete an online questionnaire about the approval system of their country. PISA-results for 2009 were also examined.

The degree of state influence can be seen from Table 1 below¹:

Model A (no influence)	Model B (strong influence)	Model C (moderate influence)	Model D (weak influence)	Model E (mixed influence)
Australia Denmark England Estonia Finland Ireland Italy Netherlands Norway Sweden	(Serbia)	Austria Czech Republic Chile China Croatia France Germany Hungary Japan Latvia Lithuania Poland (Serbia) Singapore Slovenia Slovakia Turkey	Canada	Azerbaijan Brazil New Zealand USA

Table 1: The degree of state influence

From this study, there are three groups to compare: Countries without formal state influence (model A); countries with moderate state influence (model C) and countries with a mixed influence (model E). Model B has no countries and Model D has only one country, Canada. Before disregarding Canada (model D) for the purposes of this study, the results of the mean PISA scores will be shown.

Results (mean PISA scores)

The PISA assessment 2009 was conducted on 450,000 students in 65 countries.

As can be seen from Table 2, mean scores for all subjects are higher in model A (no state influence) and model D (weak influence).

¹ It will be noted that Serbia is mentioned twice. This is due to the fact that until 2004 there was a strong state influence over textbooks; however, this has become more moderate. It should also be noted that at the time of this investigation.

Model	N	Reading scores	Math scores	Science scores	Total scores
A	10	503,10	505,00	513,60	507,23
B	-	-	-	-	-
C	17	491,24	494,12	499,76	495,04
D	1	524,00	527,00	529,00	526,67
E	4	448,75	455,75	453,00	452,50

Table 2: Mean PISA scores for models A, B, C, D and E

For an analysis of the results between groups, an Analysis of Variance (ANOVA) was conducted, and model B removed. The results (see Table 3) show a significant difference between groups for reading at the 5% level. This means that the chance that these results are based on coincidence is less than 5%.

Subject	Sum of Squares	DF	Mean Square	F	P
Reading	9690,510	3	3230,170	3,053	.045*
Math	8149,485	3	2716,495	1,768	.176
Science	11504,760	3	3834,920	2,755	.061

* Significant at the 5% level

Table 3: Difference between groups (ANOVA)

For maths and science there is no significant difference between models A, C, D and E. Model D, however, is represented by only one country, which lead to the question: What would happen if model D was removed? As Table 4 shows, after removing model D there was a significant difference in science as well:

Subject	Sum of Squares	DF	Mean Square	F	P
Reading	8542,840	2	4271,420	4,037	.029*
Math	7008,259	2	3504,130	2,280	.121
Science	10585,380	2	5292,690	3,803	.035*

* Significant at the 5% level

Table 4: Difference between groups (ANOVA), model D removed

There is, therefore, a significant difference between countries in the case of reading and science. To determine if there was a significant difference between specific models in reading, math and science, a post hoc Bonferroni test was conducted (see Table 5 below). With the Bonferroni-test it is possible to test several hypotheses simultaneously.

Subject	Independent variable	Dependent variable	Mean difference	Significance
Reading	A	C	11,865	1.000
		E	54,350	.026*
	C	A	-11,865	1.000
		E	42,485	.078
	E	A	-54,350	.026*
		C	-42,485	.078
Math	A	C	10,882	1.000
		E	49,250	.128
	C	A	-10,882	1.000
		E	38,368	.267
	E	A	-49,250	.128
		C	-38,368	.267
Science	A	C	13,835	1.000
		E	60,600	.031*
	C	A	-13,835	1.000
		E	46,765	.096
	E	A	-60,600	.031*
		C	-46,765	.096

* Significant at the 5% level

Table 5: Results between models: Bonferroni test

As Table 5 shows, there is now a significant difference between models A and E for reading, as well as a significant difference between models A and E for science. However, what does this say about these models? There may be noise (or bias), because it is not exactly clear how models A and E differ.

Conclusions

This investigation includes only half of the countries with PISA assessments. For a complete picture the approval system data of the other countries would be required. PISA might not be the only or most appropriate measure for educational outcomes; some countries seem to manipulate the scores to get a higher ranking. For example, in Portugal in 2000, the government set up an approval system as a result of declining PISA-results, changed the exam programmes for the main subjects, approved textbooks in accordance with these programmes and subsequently the PISA-scores were raised.

There are differences between models A and E concerning reading and science; however, the statistical variance of countries in models E (and C) is far greater than countries in model A. Therefore, further investigation is required of the variables responsible, for example: teacher qualifications, approval criteria, use of textbooks, amount of money spent per student, etc. (see also McEwan and Marshall, 2004). Concerning teacher qualifications, it could be argued that the importance of textbooks increases in countries with less qualified teachers, than in countries where

teachers are tertiary educated, as in Finland, the highest ranked PISA country in Europe.

As to approval criteria, there is arguably a difference if the emphasis is on criteria with respect to content (curriculum coverage) or on criteria with respect to the pedagogical approach. It is also important to consider the perspective from which these criteria are formulated: from the publisher's perspective, the ministry's perspective, the teacher's perspective, the parent's perspective, or the learner's perspective.

As to the use of textbooks, this of course differs if textbooks are the main source for teachers' lesson plans, as in the Netherlands, or if textbooks are just one of the many resources teachers use, as in Australia.

In conclusion, however, we believe textbooks are very important to teachers and education; and would emphasise the need for high quality textbooks – paper and digital). It could be argued, therefore, that approval systems are necessary in achieving this. TIMSS (Third International Mathematics and Science Study – a four year cycle of international student evaluations 1995-2015) analysis has collected data on textbook use in grade 4 and grade 8 as part of the TIMSS analysis of student achievement in mathematics, science and reading student. A further development of this research is to conduct similar analysis on TIMSS results.

References

Grubb, W.N. (2008). Multiple resources, multiple outcomes: Testing the “improved” school finance with NELS88. *American Educational Research Journal*. Vol. 45, No. 1, 104-144.

Hanushek, E. (2003). The failure of input-based schooling policies. *Economic Journal* 113, 64-98.

Hanushek, E. (2006). School resources. In: Hanushek, E. & Welch, F. (Eds). *Handbook of the Economics of Education*. Elsevier. Amsterdam, Vol 2, 865 -908.

Hattie, J. (2009). *Visible Learning*. Routledge. Oxford.

Horsley, M. (2011). *Investment in classroom teaching and learning materials: equity and access in providing classroom teaching and learning materials in Australian schools*. Australian Publishers Association. Sydney

Horsley, M. & Walker, R. A. (2012, in press). *Reforming homework: Practices, learning and policies*. Palgrave Macmillan.

Entwistle, N. (1975). How students learn: information processing, intellectual development and confrontation. *Higher Education Bulletin*, Vol. 3, 129-148

Jonassen, D.H. & Grabowski, B.L. (1993). *Handbook of Individual Differences, Learning and Instruction*., Lawrence Erlbaum. NJ.

Kennedy, M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher*. Vol. 39, No. 8, 591-598.

Khutoriskoi, A.V. (2006). The place of the textbook in the didactic system. *Russian Education and Society*, Vol. 48, No. 3, 78-93.

McEwan, P.J. & Marshall, J.H. (2004). Why does academic achievement vary across countries? Evidence from Cuba and Mexico. *Education Economics*. Vol. 12, No. 3, 205-217

Ninomiya, Akira (2012, in press). Textbooks selection criteria in Japan: how these criteria improve the quality.

Ogan-Bekiroglu, F. (2007). To what degree do currently used physics textbooks meet expectations? *Journal of Science Teacher Education*, Vol 18, 599- 628.

Pingel, F. (2010). *UNESCO Guidebook on textbook research and textbook revision*. Paris, Braunschweig.

Pratt, D.D. (2002). Good teaching: one size fits all? *New directions for adult and continuing education*. Vol. 93, 5-15

Repoussi, M., & Tutiaux-Guillon, N. (2010). New trends in history textbook research: Issues and methodologies toward a school historiography. *Journal of Educational Media, Memory and Society*, Vol. 2, No. 1, 154-170.

Sikarova, Zuzana (2011). The role of textbooks in lower secondary schools in the Czech Republic. *IARTEM e-Journal*. Vol .4, No. 2, 1 – 22

Sousa, Joana & Lourdes Dionísio, Maria de (2012 in press). Evaluation and selection of textbooks in Portugal: perceptions from mother tongue teachers.

Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002) *According to the Book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Dordrecht. Kluwer Academic.

Hendrianne Wilkens

Hendrianne Wilkens is senior researcher and consultant at CLU, Centre for Textbook Development and Evaluation, Utrecht University. She does research and gives training courses in aspects of quality of textbooks. She is also responsible for coaching students who are doing their master thesis on these aspects. Hendrianne has been for 20 years a former teacher in primary education.

Contact address:

Hendrianne Wilkens
CLU. Utrecht University
Heidelberglaan 2
3584 CS Utrecht

h.wilkens@clu.nl