Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.7, No.18, 2016

Using Comparative International Studies for Modeling Educational Effectiveness: A secondary analysis of PISA-2009 study

John Bosco Azigwe
PhD Candidate, Education Department, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus

Abstract

International surveys of student achievement are becoming increasingly popular with governments around the world, as they try to measure the performance of their country's education system. The main reason for this trend is the shared opinion that countries will need to be able to compete in the 'knowledge economy' to assure the economic wellbeing of their citizens. This paper argues that secondary analyses of international comparative studies can help the development of the theoretical framework of educational effectiveness research. Toward this end, we present the results of a secondary analysis of PISA-2009 study which investigates the validity of the Dynamic model of Educational Effectiveness Research especially with regards the school level factors. Across-country multilevel analysis of reading achievement revealed the importance of contextual factors and student level factors included in the dynamic model, especially student motivation, opportunity to learn, and school climate factors. Additionally, a comparative analysis with six countries seems to suggest that the model holds for individual countries as well especially at the student level. Based on these results, we draw implications for the design of comparative studies aiming to contribute to the development of evidence-based reform policies in education.
Key words: International comparative studies, theory, and educational effectiveness

1. Introduction

This paper is concerned with the major benefits of testing the validity of theoretical models of Educational Effectiveness Research (EER) by conducting secondary analyses of comparative international studies. Researchers have two options when investigating the generalizability of models of EER (Creemers, Kyriakides, & Sammons, 2010); they could either conduct meta-analyses of national studies and/or secondary analyses of international comparative studies. There are advantages in using comparative studies such as those conducted by the International Association for the Evaluation of Educational Achievement (IEA) and the Organization for Economic Cooperation and Development (OECD) to examine the validity of models of EER rather than meta-analyses of national studies (Creemers et al., 2010). The IEA and OECD studies deploy a common research design in all countries and all variables are operationalized in English language and translated to the respective national languages using strict translation procedures (OECD, 2011). Moreover, when creating a value-added based classroom/school effect, this effect is possible to be constructed statistically in the same way in all countries.

Additionally, two major benefits of international comparative studies are also raised (Kyriakides, 2006). First, cross-national studies (i.e. IEA and OECD) reveal cross-national variation on achievement scores. This implies that only international studies can tap the full range of variation in school and classroom quality, and therefore, in potential school and classroom effects. It is also likely that the existing estimates of the size of educational influences (i.e., schools and classrooms/teachers together) upon student outcomes are potentially influenced the studies' lack of school and classroom variation. Thus, the actual power of school and classroom variables can only be identified by cross-cultural and comparative work on international samples (Kyriakides, 2006). The second major benefit of conducting comparative studies in educational effectiveness is concerned with the need of understanding much more about why some variables explain effectiveness across countries while others do not (Reynolds, Creemers, Stringfield, Teddlie, & Schaffer, 2002). Results of meta-analyses of national studies on educational effectiveness indicate that the size of school effects differs across countries and so do the effects of their antecedent conditions (Kyriakides, Creemers, Antoniou, & Demetriou, 2010; Scheerens & Bosker, 1997). In these respects, international studies may help us identify school and teacher effectiveness factors that are present in different educational contexts (Creemers, 2006).

At the same time, the main disadvantage of conducting secondary analyses of comparative studies to test and develop models of educational effectiveness is that these studies were not necessarily designed to identify or test educational effectiveness factors (Kyriakides & Charalambous, 2005). Nevertheless, the ultimate goal of the international comparative studies has been to isolate those factors related to student learning which could be manipulated through policy changes in curriculum, resource allocation or instructional practice (e.g., Martin, 1996; OECD, 2002; Yang, 2003). It has been expected that information that arises from such investigations could

help policy-makers, curriculum specialists and researchers better understand the performance of their educational systems (Campbell & Kyriakides, 2000; Mullis et al., 2000). When taken into account that identifying factors at different educational levels that have an effect on student achievement is among the ultimate goals of EER, it can be claimed that secondary analyses of international studies could also help in testing the validity of models of EER.

In light of these arguments, this paper examined the extent to which the dynamic model of educational effectiveness (Creemers & Kyriakides, 2006) considered as one of the most influential theoretical models in the field (Reynolds et al., 2011), could help in identifying variables associated with student achievement in the Program for International Student Assessment (PISA) 2009 study. The empirical studies that were conducted to test the dynamic model (e.g., Creemers & Kyriakides, 2010a, 2010b; Kyriakides & Creemers, 2008, 2009) used value-added techniques and examined students' progress rather than final outcomes. Information gathered from value-added assessment is more valid in exploring the effectiveness of a school unit or an educational system than using outcome data only. Nevertheless, these national studies cannot help us identify the extent to which the proposed factors of the model can be considered as generic. It was therefore anticipated that, results of this analysis may provide further support and development of the dynamic model. Consequently they may also reveal the importance of designing international effectiveness studies based on models of EER.

The remainder of the paper is organized as follows: - First, a brief review of the key characteristics of the dynamic model is presented to inform readers on how variables from the PISA 2009 data set were selected for the ensuing analysis. Then, we present the findings of the study, first for the entire sample, and then for selected countries. In the concluding section, suggestions on how the dynamic model could be used for developing theory and policy are presented.

1.2. Key Characteristics of the Dynamic Model of Educational Effectiveness

Educational effectiveness studies conducted in several countries reveal that the influences on student achievement are multilevel (Teddlie & Reynolds, 2000). In this respect, the dynamic model is multilevel in nature and refers to factors associated with student learning outcomes that are situated at four different levels: student, classroom, school, and educational system (Creemers & Kyriakides, 2008). The theory behind the model is comprehensive in nature and looks simultaneously at all the different levels of the educational system (Reynolds et al., 2011). The teaching and learning situation is emphasized and the roles of the two main actors (i.e., teacher and student) are analyzed (Creemers & Kyriakides, 2008). Above these two levels, the model also refers to school-level factors. It is expected that school-level factors influence the teaching-learning situation by developing and evaluating the school policy on teaching and the policy on creating a learning environment at the school. The final level refers to the influence of the educational system through a more formal way, especially through developing and evaluating the educational policy at the national/regional level. A major strength of the model is that it is established in a way that helps policy makers and practitioners to improve educational practice by taking rational decisions concerning the optimal fit of the factors within the model and the present situation in the schools or educational systems (Creemers & Kyriakides, 2008).

1.2.1. Student background factors

The model refers to two main categories of student background factors operating at the student level which can influence the effectiveness of education, namely, socio cultural, economic and student engagement in learning (Creemers & Kyriakides, 2008). A distinction is made among the factors by referring to factors which are not directly within the control of the school and are unlikely to change (e.g., gender, SES, ethnicity, personality) and factors that may change over time (e.g., aptitude, motivation, expectations, personality, and thinking style).

1.2.2. Classroom-level factors

One of the key findings from decades of EER is the importance of the classroom level as a predictor of pupil outcomes (Scheerens & Bosker, 1997). Research has consistently shown not only that the classroom level can explain more of the variance in pupil outcomes than the school level, but that a large proportion of this classroom level variance can be explained by what teachers do in the classroom (Creemers & Kyriakides, 2008). Based on these arguments, the model refers to factors which describe teachers' instructional role and are associated with student outcomes. These factors refer to observable instructional behavior of teachers in the classroom rather than to factors that may explain such behavior (e.g., teacher beliefs and knowledge and interpersonal competences). The eight factors included in the model are as follows: orientation, structuring, questioning, teaching-modeling, applications, management of time, teacher role in making classroom a learning environment, and classroom assessment.

1.2.3. School-level factors

In addition to student and classroom level factors, the dynamic model acknowledges the importance of the school climate. The assumption of the model is that school factors are expected to influence classroom-level factors, especially the teaching practice. An emphasis is given to the two main aspects of the school policy which affect learning at both the level of teachers and students. Thus, the following four overarching factors at the school level are included in the model: (1) school policy for teaching and actions taken for improving teaching practice; (2) evaluation of school policy for teaching and of actions taken to improve teaching; (3) policy for creating a school level environment (SLE) and actions taken for improving the SLE; and (4) evaluation of the SLE.

1.2.4. Educational system level

The final level refers to the influence of the educational system through a more formal way, especially through developing and evaluating the educational policy at the national/regional level (Creemers & Kyriakides, 2006). It also is taken into account that the teaching and learning situation is influenced by the wider educational context in which students, teachers, and schools are expected to operate. Factors such as the values of the society for learning and the importance attached to education play an important role both in shaping teacher and student expectations as well as in the development of the perceptions of various stakeholders about effective teaching practice. Thus, the following factors at the educational system level are included in the model: (1) national/ regional policy for education; (2) evaluation of national/regional policy for education; (3) the wider educational context (i.e. support provided to schools by different stakeholders, expectation of schools about learning and learning outcomes by different stakeholders).

2. A Secondary Analysis of the Program for International Student Assessment (PISA) 2009 Study Taking into Account the Dynamic Model of EER

PISA is a significant source of international assessments which uncover variables and models of interest to educational quality and effectiveness for both industrialized and developing countries (Riddell, 2008). PISA is coordinated by the OECD, an intergovernmental organization of 34 member countries and measures the performance of 15-year-olds in reading literacy, mathematics literacy, and science literacy every three (3) years (OECD, 2009). The motivation of PISA at inception was to fill the perceived gap in the extensive set of indicator-based information on education systems that the Organization for Economic Cooperation and Development (OECD) had been providing through other national assessment programmes that pre-existed it (Baird, Isaacs, Johnson, Stobart, Yu, Sprague, & Daugherty, 2011). Although PISA began as a joint survey of OECD member countries, it has developed to involve non-member countries throughout the world in providing a global perspective on educational policy and reform (Lingard & Grek, 2008). Thus, PISA is distinguished by its connection with the OECD which offers an extensive network and machinery that guarantees a greater presence on the world stage (Murphy, 2010).

The aim of PISA through these surveys is to inform parents, students, the public and managers of education systems about whether young people reaching the end of compulsory education have acquired the necessary skills and knowledge to meet the challenges of present-day society (OECD, 2009). It also aims to provide a new basis for policy dialogue and collaboration in defining and implementing educational goals, in innovative ways that reflect judgments about the skills that are relevant to adult life (OECD, 2009). As a consequence, PISA assessments are forward-looking - rather than focusing on the extent to which participating students have mastered a specific school curriculum, it looks at their ability to use their knowledge and skills to meet real-life challenges (Baird et al., 2011). In these respect, PISA assesses how far students have acquired some of the knowledge and skills essential for live. In addition, questionnaires are administered to students and school principals in order to collect data that could be used in constructing indicators pointing to social, cultural, economic and educational factors that are associated with student performance (OECD, 2009).

PISA encompasses a great deal of potential with the presentation of empirical material of worth in educational outcomes at both the national and international level (Hopmann, Brinek, & Retzl, 2007). It offers a range of evidence with which to monitor the performance of educational systems by providing valid and reliable data which can be used by governments, schools, teachers and other educational professionals to debate the strengths and weaknesses of their education systems in relation to other countries (Mortimore, 2009; Willms, 2003). At the same time, PISA's suffers some limitations (i.e. Dolin, 2007; Goldstein, 2004, 2008; Mortimore, 2009; Postlethwaite, 2006; Smithers, 2004). The limitations notwithstanding, PISA has great potential for developing

theory and knowledge about the ways in which different structures, policies and practices lead to different educational outcomes (Perry, 2008).

2.1. Identifying variables associated with student achievement in reading literacy

In this section we present results of a two stage analysis of PISA 2009 data by taking into account the Dynamic Model of Educational Effectiveness (Creemers & Kyriakides, 2008) using multilevel modeling techniques. The first stage involved the entire sample of countries participating in the PISA 2009 study. In the second stage, using the same dataset, we conducted separate multilevel analyses of six selected countries (i.e., Canada, Tunisia, Japan, Korea, New Zealand, and United Kingdom) which are considered to have remarkable differences both in the way in which support is provided to schools in order to implement curriculum and in the core beliefs of their societies concerning appropriate educational aims and best practices. The selection of these countries is based on the claim that in some countries (i.e., the East Asian countries) virtually all educational professionals adopt the same values about what should happen in a classroom or a school, whereas in other countries (i.e., the English-speaking countries) there is huge variation in what is seen as appropriate or at least acceptable teaching practice which might reflect unresolved value debates at a national level (Alexander, 2000).

3. Research Questions

In this context, this paper attempts to validate the dynamic model of educational effectiveness, focusing, in particular, on school-level factors. Specifically, the paper aims to answer the following two research questions:

I. To what extent do the PISA 2009 data on student learning in reading literacy corroborate the dynamic model of educational effectiveness, especially with regards to school climate factors?

II. Provided that the PISA 2009 data as a whole empirically support the dynamic model of educational effectiveness, to what extent does the model also hold for individual countries?

In essence, the second research question serves as some sort of sensitivity analysis, given that it examines the tenability of the model for countries having different educational systems.

4. Methods

4.1. The sample

The PISA 2009 sample is estimated to involve 470 000 students in the schools of the 65 countries (OECD, 2009). The first stage of the analysis involved selecting all countries and (schools and students) within these countries. In the second stage, six countries were selected for the analysis. Table 1 below presents the number of schools and students in the respective countries selected for the second stage of the analysis.

Table 1: Selected countries for second stage analysis

Country	Number of schools	Number of students
United Kingdom	482	12,153
Canada	939	23,207
New Zealand	153	4,643
Japan	186	5,818
Korea	157	5,029
Tunisia	165	4,957
Total	2082	55,807

4.2. Identifying explanatory variables by taking into account the Dynamic Model of Educational Effectiveness. At this stage, reference is made to the dynamic model. We were able to select variables at the student school and educational context level was selected. However, it was not possible to select variables at the teacher/classroom level, because such data are not included in PISA 2009. The following variables were included in the analyses.

4.2.1. Response variable

PISA's major focus in 2009 was on reading literacy. The assessment on reading literacy is built on three major task characteristics: "situation – the range of broad contexts or purposes for which reading takes place; text – the "range" of material that is read; and "aspect" – the cognitive approach that determines how readers engage with

a text (OECD 2009).

4.2.2. Explanatory variables

Student background variables: Student background variables taken into account had to do with student gender (i.e., 0=female: 1=male) and age, and with the educational status and occupation of their parents. Specifically, the variable AGE is calculated as the difference between the middle month and the year in which students were assessed and their month and year of birth, expressed in years and months. The educational level of parents is classified using ISCED (OECD, 1999).

Opportunity to learn - Diversity of reading materials: The index of diversity of reading materials (DIVREAD) was derived from the frequency with which students read magazines, comic books, fiction, non-fiction books and newspapers. A higher value on this index indicates higher diversity in reading. This index was treated as a proxy measure of the opportunity to learn factor concerned with reading.

Subject motivation - Joy like reading activities (JOYREAD): The index of enjoyment of reading activities (ENJOY) was derived from students' level of agreement with statements in the student questionnaires (i.e. I read only if I have to; *ii)* reading is one of my favorite hobbies; *iii)* I like talking about books with other people etc.).

Metacognition strategies-understanding and remembering: The index of understanding and remembering (UNDREM) was derived from students' reports on the usefulness of their strategies for understanding and memorizing the text.

Metacognition strategies-summarizing: The index of summarizing (METASUM) was derived from students' reports on the usefulness of their strategies for writing a long summary of a text. Higher values on this index indicate greater students' perception of usefulness of this strategy.

Use of control strategies: The index of how students approach learning was based on student responses measured through the use of control strategies (CSTRAT) among others. Higher values on this index indicated higher importance attached to the given strategy.

4.2.3. School level factors

Disciplinary climate: The index of disciplinary climate (DISCLIMA) was derived from students' reports on how often in their lessons they did not listen to what the teacher said; *ii)* there was noise and disorder etc. Higher values on this index indicated a better disciplinary climate.

Teacher shortage: The index of teacher shortage (TCSHORT) was derived from four items measuring school principals' perceptions of potential factors hindering instruction at their school. These factors are a lack of: *i)* qualified science teachers, *ii)* a lack of qualified mathematics teachers, *iii)* qualified <test language> teachers, and *iv)* qualified teachers of other subjects. Higher values on this index indicated school principals' reports of higher teacher shortage at a school.

School type: Schools were classified either as public or private, according to whether a private entity or a public agency has the ultimate power to make decisions concerning its affairs.

Teacher behavior: This index was derived from school principals' reports on the extent to which the learning of students hindered by the following factors in their schools: i) teachers' low expectations of students; ii) poor student-teacher relations; iii) teachers not meeting individual students' needs; iv) teacher absenteeism; v) staff resisting change; vi) teachers being too strict with students; and vii) students not being encouraged to achieve their full potential.

Student behavior: The index of student-related factors affecting school climate was derived from school principals' reports on the extent to which the learning of students is hindered by the following factors in their schools (SC17): i) student absenteeism; ii) disruption of classes by students; iii) students skipping classes.

Quantity of teaching staff at school: The proportion of fully certified teachers (PROPCERT) was computed by dividing the number of fully certified teachers by the total number of teachers in a school. The proportion of teachers who have an ISCED 5A qualification (PROPQUAL) was calculated from this index.

Disciplinary climate (DISCLIMA): The index of disciplinary climate (DISCLIMA) was derived from students' reports on how often the followings happened in their lessons of the language of instruction was devoid of noise and disorder.

5. Analysis

The analysis was carried out in two parts: the first part involved all countries and students in the data set and the second part involved only six countries. This enables the testing of the dynamic model with respect to the entire sample and also with selected six countries as some sort of sensitivity analysis The PISA data fits the multi-stage sampling (i.e., students are nested within schools and schools within countries), and thus enabled the use of MLwiN (Goldstein et al., 1998) to examine the extent to which selected variables were associated with student achievement in reading literacy.. The analysis started with the null model after which student background factors (i.e. gender, socioeconomic status and aggregates of them at school and country level were added in Model 1. Thereafter, student attitudes, engagement and motivation in learning added in Model 2. In Model 3 and 4, school climate conditions and country level factors related to policy for teaching were considered. Model 4 was however abandoned as all selected variables at this level were not statistically significant.

6. Results

We first present the data for the entire sample. Following this, we briefly present our findings regarding the extent to which the results for the entire sample held across different individual countries

6.1. Entire-sample analysis

Model 0

Table 2 below presents results of three models used in the analysis. The null model presents the variance at student, school, and country levels without any explanatory variable (see Table 2: column 1). We can observe that 23.4% of the variance was at the country, 30.4% at the school level and 46.2% at the student level. We acknowledge that the percentage accounted for at the school level might have been over-estimated since the model did not include the intermediate classroom level.

Model 1

In Model 1 contextual variables measuring student background factors and their aggregate scores at school and country levels were entered. The following observations are made (see Table 2: column 2). First, sex and the parental educational level all have statistically significant effects on student learning in reading literacy. Particularly, female students coded as 1 appear to be doing better than their male counterparts in reading literacy. Also, the educational level of mothers of students also appears to contribute better to the reading literacy of students than the educational level of their fathers. Additionally, the aggregates of student background variables at the school level are statistically significant in their effect on student learning in reading literacy. However, when it comes to the aggregates at the country level, these characteristics had no significant effects, and therefore, were not included in the model. Model 1 explained 20.8 % of the unexplained variance component, of which 2.2 % was situated at the country level, 75.3 % at the school level and 22.5 % at the student level. Also, the likelihood statistic (X^2) shows a significant change between the empty model and model 1 ($p<.001$), suggesting that the latter model had a better fit to the data.

Model 2

In Model 2, variables relating to student attitude and motivation to learning were entered in the model. The following observations are made based on the third column of Table 2. First, selected variables in relation to opportunity to learn (i.e. lessons for improving skills, diversity in reading, joy/like reading) and subject motivation (i.e. use of control strategies, metacognision-summarising, metacognition-understanding and remembering) have a significant effect on reading literacy. This model explained an additional 39.2 % of the variance component, with 25.6 % at the country level, 27.5 % at the school level and 46.9 % at the student level. Also, the likelihood statistic (X^2) shows a significant change between Model 1 and Model 2 ($p<.001$).

Model 3

In Model 3, variables related to the school learning environment were entered in the model. The following observations are made based on the fourth column of Table 2. In this column, factors measuring the perceptions of students on the school learning environment (i.e. the proportion of qualified teachers and the disciplinary climate in schools have positive effects on student reading achievement. However, as to be expected, student absenteeism, low expectation from teachers and student disruptive behavior had negative effects. This model

explained an additional 41.6 % of the variance component, with 18.5 % at the country level, 72.6% at the school level and 8.9% at the student levels. Again, the likelihood statistic (X^2) shows a significant change between Model 2 and model 3 ($p<.001$), suggesting that the latter model had a better fit to the data.

Table 2. Parameter estimates and (Standard Errors) for the multilevel models used to investigate educational effectiveness in Reading literacy (students within schools, within countries) PISA Study 2009

	Model 0	Model 1	Model 2	Model 3
Factors				
Fixed Part (Intercept)	459.59 (6.39)*	752.05 (56.33)*	687.85 (8..54)*	693.38 (9.20)*
Student level				
Student Context				
Age		-7.75 (0.18)*	-6.25 (0.17)*	-6.41 (0.25)*
Sex (f=1, m=0)		-28.60 (0.23)*	-9.78 (0.23)*	-10.17 (0.25)*
Educational level of father (ISCED)		-5.85 (0.08)*	-3.59 (0.12)*	-3.60 (0.13)*
Educational level of mother (ISCED)		-5.32 (0.08)*	-4.42 (0.12)*	-4.51 (0.13)*
Opportunity to learn				
Diversity in reading			0.63 (0.12)*	0.63 (0.13)*
Joy/like reading			19.51 (0.13)*	19.37 (0.15)*
Strategies and motivation				
Use of control strategies			4.71 (0.12)*	4.67 (0.13)*
Meta cognition-Summarizing			16.99 (0.13)*	16.97 (0.15)*
Meta cognition-Understanding & Remembering			10.17 (0.12)*	9.99 (0.13)*
School level				
School context				
Aggregate age at school ID		-11.83 (3.57)*	-6.25 (0.17)*	-6.27 (1.14)*
Aggregate sex at school Id		-40.24 (1.72)*	-31.46 (1.41)*	-26.98 (1.49)*
Aggregate educational level of father (ISCED) at school ID		-25.26 (1.32)*	-19.63 (1.10)*	-16.77 (1.16)*
Aggregate educational level of mother (ISCED) at school ID		-25.67 (0.91)*	16.99 (0.12))*	18.36 (0.82)*
Climate factors				
Shortage test language teachers				-0.09 (0.44)
Proportion of qualified teachers				15.83 (1.36)*
Disciplinary climate				2.76 (0.50)*
Teachers too strict				3.27 (0.50)*
Low Teachers Expectation				-3.81 (0.42)*
Student absenteeism				-4.91 (0.42)*
skipping classes				-1.46 (0.50)*
Disruptive Behavior				-2.74 (0.50)*
Students Being Bullied				-2.76 (0.50)*
Academic Pressure				-5.31 (0.45)*
Country level				
Achievement public schools				NSS
Responsibility for course content, Regional/national				NSS
Regional/Nat influence on instructional content				NSS
Random Part				
Country level	23.4%	22.9 %	18.1 %	20.7 %
School level	30.4%	14.8 %	9.5 %	8.7 %
Student level	46.2%	41.5 %	32.6 %	33.3 %
Absolute	11238.3	8908.48	6766.5	7049.4
explained		20.8 %	39.8 %	37.3
Significance test				
X2	5465736	4600186.7	4149739.6	3781799
Reduction		865549.3	450447.1	692031
Degrees of freedom		8	5	7
p value		.001	.001	.001

Note *= statistically significant effect at .05 level

6.2. Second stage analysis: six countries

Table 3 below presents results of this aspect of the analysis. The objective at this stage of the analysis was to observe any similarities or differences that may exist between the results of the entire sample and that of the six individual countries. Doing so would help further empirically validate the dynamic model of educational effectiveness, given that it would help test whether the model holds not only for the entire sample but also for countries with different educational systems. Because school contextual factors were envisioned to be different across countries, in what follows we limited our attention to the student background factors, the opportunity to learn factors, and the school-climate factors. As can be observed in Table 3, there is a relative consistency across the six countries with regards to student background factors, and the opportunity to learn factors. When it comes

to the school-level factors, we again see that the model also holds across countries, at least to some extent. For example, academic pressure, student absenteeism appears to hold across almost all the selected countries. Some other factors appeared to hold in half of the countries (e.g., teachers's low expectation from their students), while there was just one factor that was observed as only significant in one country (e.g. Canada). This latter finding was not surprising, because the the other selected countries might have not suffered from such a shortage. Hence, it can be stated that, overall, the model appears to hold for individual countries as well.

Table 3. Checklist of how selected variables that were observed as statistically significant in select countries comparing with the total PISA 2009 sample.

Factors	Total Sample	UK	Canada	N Zealand	Korea	Japan	Tunisia	
Intercept	693.38 (9.20)*	448.19 (48.97)	447.33 (29.21)	186.95 (78.62)	569.48 (51.63)	587.75 (84.63)	429.42 (72.09)	
Student factors - Context								Out of four variables Yes = 18 Blank=6
Age	-6.41 (0.25)*	Yes	Yes	Yes				
Sex (F =1 M= 0)	-10.17 (0.25)*			Yes	Yes	Yes	Yes	
Educational level of father (ISCED)	-3.60 (0.13)*	Yes	Yes	Yes	Yes	Yes		
Educational level of mother (ISCED)	-4.51 (0.13)*	Yes	Yes	Yes	Yes	Yes	Yes	
Opportunity to learn								Out of five variables Yes= 25 Blank=5
Diversity in reading	0.63 (0.13)*		Yes	Yes				
Joy/like reading	19.37 (0.15)*	Yes	Yes	Yes	Yes	Yes		
Strategies and motivation								
Use of control strategies	4.67 (0.13)*	Yes	Yes	Yes	Yes	Yes	Yes	
Meta cognition-Summarising	16.97 (0.15)*	Yes	Yes	Yes	Yes	Yes	Yes	
Meta cognition-Understanding & Remembering	9.99 (0.13)*	Yes	Yes	Yes	Yes	Yes	Yes	
School Learning Environment								
Shortage test language teachers	-0.09 (0.44)		Yes					
Proportion of qualified teachers	15.83 (1.36)*				Yes		Yes	
Disciplinary climate	2.76 (0.50)*	Yes		Yes		Yes		
Teachers too strict	3.27 (0.50)*		Yes					
Low Teachers Expectation	-3.81 (0.42)*		Yes		Yes		Yes	
Student absenteeism	-4.91 (0.42)*	Yes		Yes	Yes	Yes		
skipping classes	-1.46 (0.50)*		Yes			Yes	Yes	
Disruptive Behavior	-2.74 (0.50)*	Yes				Yes		
Students Being Bullied	-2.76 (0.50)*					Yes		
Academic Pressure	-5.31 (0.45)*	Yes	Yes	Yes	Yes	Yes	Yes	

Note: YES: consistent results with the model representing all the countries, Blank spaces = inconsistent results (effects were not significant for the individual countries).

7. Discussion

This paper attempted to test the extent to which variables selected from PISA 2009 based on the Dynamic Model of Educational Effectiveness could be associated with student learning in reading literacy. As a further step of validating the model, we also tested the extent to which the model also held across six individual countries with different educational systems. Our first step of the analysis that concerned the entire sample empirically supported the Dynamic Model, by showing variables both at the student-level (i.e., student background variables, and opportunity to learn variables) and at the school level to have significant effects on student reading literacy performance. The second step of the analysis also provided some evidence suggesting that the model also holds across countries, especially with regards to the student background and opportunity to learn factors, and to some lesser degree with regards to the school-level factors. That some factors at the school level were more consistently observed across countries, whereas others were less consistently is also considered reasonable, especially when taking a closer look at the factors falling in each category. For example, as one would expect, student absenteeism and skipping of classes appear to have a "universal" effect, since skipping being absent reduces opportunity to learn. On the other hand, the extent to which a school is public or private did not have a consistent effect across countries, something that might have to do with the different functioning of these schools across different countries (e.g., in one country private schools might be more successful at establishing more productive learning environments, while in other countries they might not). Obviously, such results call for further research that will examine this and similar assumptions. Beyond validating the dynamic model of educational effectiveness, this study also corroborates pertinent prior research findings. For example, that the socio-economic status of students, their motivation and strategies for learning had significant effects on their learning gain was in accord with the argument of a significant link between student approaches to learning (i.e.

self-control, personal goals, motivation and self-belief) and their learning outcomes (Artelt, Baumert, Julius-McElvany, & Peschar, 2000) (see also Trowler, 2010; Willms, 2003). Furthermore, our analysis pointed to the importance of school climate conditions as important for student learning. This was consonant with prior study findings and theoretical arguments according to which, students perform better in schools with more disciplined classrooms, partly shaped by the resources, policies and practices of the systems and schools (OECD, 2010). A positive school climate was also suggested to be an important component of successful and effective schools (Koth, Bradshaw, & Leaf, 2008), and a necessary condition for learning (Creemers & Kyriakides, 2008)—an argument that was also corroborated by the study findings. In fact, Frank and Rosen (2010) suggested that a significant and positive relationship exists between school climate and reading achievement at both the student and classroom/school level; the present study suggests a similar trend. Also, as to be expected, student absenteeism, skipping classes and disruptive behavior accounted negatively to their learning gain in reading literacy Additionally, at the second stage of our analyzes, we observed a relatively high consistency across the six selected countries compared with the total sample at the student background level, opportunity to learn, student motivation and strategies for learning. This was similar for school climate factors, although not as pronounced as that of the level of student level.

8. Conclusion

One of the objectives of this paper was to explore how international comparative studies can help in developing the theoretical framework of EER. As a result, we tested the dynamic model through a secondary analysis of PISA 2009 data to determine if the model could be supported. Also, through this analysis our aim was also to provide implications for the improvement of international comparative studies and to also suggest policy options that might emerge regarding teaching and learning. We conclude this paper by considering these issues in turn.

The significance of PISA in providing a global perspective on educational policy and reform has been acknowledged (e.g. Baird et al., 2011; Lingard & Grek, 2008; Riddell 2008; Murphy, 2010). This could further be expanded to cover more developing countries and by extension countries in the African sub region, especially if one takes into consideration that such countries are largely absent from the PISA sample (with some notable exceptions, such some of the northern African countries like Tunisia that participated in the PISA 2009 study. Sampling more African countries has the potential to further enrich PISA data and ultimately, may contribute to the development of policy initiatives that address the specific needs of such countries. It has been argued that children in developing countries not only receive fewer years of education, but also attain comparatively lower quality education due to factors that include inadequate learning time, student and teacher absenteeism (Reddy, 2007). Although these factors appear to be global across countries and schools, they are more pronounced in developing countries (e.g., Glewwe & Kremer, 2005; Boissiere, 2004; Abadzi, 2007). Hence, by sampling more African countries, PISA will allow for better examining the extent to which these factors "travel" across countries, and especially across countries in which educational resources might be scarce and might be hindering the functioning of such factors.

A second area in which PISA studies could be improved relates to the inclusion of more teacher/classroom level variables (see a similar discussion in Mortimore, 2009). In our analysis, although we had originally thought of also including teacher-level variables (which are suggested by the Dynamic Model), we were not able to do so, because of limitations in the PISA dataset. By incorporating variables at this level, future PISA studies can help better empirically validate models of educational effectiveness, especially if one takes into consideration that teacher/classroom level variables are considered fundamental in such models. Furthermore, the model proposed here suggests that beyond examining students' progress in terms of learning outcomes, we need to collect longitudinal data for both teachers and students. Namely, we suggest that it is worth examining both the short term as well as the long-term effects of teaching on learning outcomes.

Turning now to the policy implications of this study, we note that a positive school climate has been consistently argued as an important component of successful and effective schools (i.e. Koth et al., 2008; Creemers & Kyriakides, 2008; Frank & Rosen, 2010; OECD, 2010). The results of the present study largely corroborate this finding, especially if one takes into consideration the results of the entire sample. In this respect, policy makers and schools could adopt a dynamic approach to school improvement (DASI) proposed by the dynamic model (Creemers & Kyriakides, 2012). DASI refers to factors at the school level which are related to key concepts of quantity of teaching, provision of learning opportunities, and quality of teaching. Specifically, the emphasis in DASI is on school policy which affects learning at both the level of teachers and students as follows: a) School policy for teaching and actions taken for improving teaching practice b) Evaluation of school policy for teaching and actions taken to improve teaching c) School policy for creating a school learning environment (SLE) and actions taken for improving the school learning environment, and d) evaluation of the school learning environment. By attending to such factors, schools can improve the learning environment afforded to students, which in turn, can boost student performance. The extent to which DASI can create an environment for such changes to occur across different countries is an open issue and an issue which we are currently working on

addressing.

References

Abadzi, H. (2007). Absenteeism and Beyond: Instructional Time Loss and Consequences", Washington, DC: *World Bank Policy Research Working Paper*, No. 4376 (October). Available from http://go.worldbank.org/YTOZYABWK0

Alexander, R. (2000). *Culture and pedagogy*. Oxford, UK: Basil Blackwell.

Artelt, C., Baumert, J., Julius-McElvany, N., & Peschar, J. (2000). *Student approaches to Learning: Results from PISA 2000*. Paris: OECD.

Baird, J., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T., & Daugherty, R. (2011). *Policy Effects of PISA*. Oxford University Centre for Educational Assessment

Boissiere, M. (2004). Determinants of Primary Education Outcomes in Developing Countries. *Background Paper for the Evaluation of the World, Bank's Support to Primary Education.* The World Bank Operations Evaluation Department (OED), Washington, D.C.

Campbell, R.J., & Kyriakides, L. (2000). The National Curriculum and Standards in Primary Schools: a comparative perspective. *Comparative Education, 36*(4), 383–395.

Creemers,B.P.M. (2006). The importance and perspectives of international studies in educational effectiveness. *Educational Research and Evaluation, 12*(6), 499–511.

Creemers, B.P.M., & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement, 17*(3), 347–366.

Creemers, B.P.M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools.* London: Routledge.

Creemers, B.P.M., & Kyriakides, L. (2010a). School factors explaining achievement on cognitive and affective outcomes: establishing a dynamic model of educational effectiveness. *Scandinavian Journal of Educational Research, 54*(3), 263–294.

Creemers, B.P.M., & Kyriakides, L. (2010b). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement, 21*(4), 409–427.

Creemers, B.P.M., & Kyriakides, L. (2012). *Improving Quality in Education: Dynamic Approaches to School Improvement.* London: Routledge.

Creemers, B.P.M., Kyriakides, L., & Sammons, P. (2010). *Methodological Advances in Educational Effectiveness Research*. London and New York: Taylor & Francis.

Dolin, J. (2007). PISA — an Example of the Use and Misuse of Large-Scale Comparative Tests. In Hopmann, Brinek, & Retzl (Eds.), *PISA zufolge PISA — PISA According to PISA* (pp. 93–125). Wien: Lit-Verlag.

Frank, E., & Rosen, M. (2010). The Importance of a Safe School and Classroom Climate for Student Achievement in Reading Literacy. *International Association for the Evaluation of Educational Achievement (IEA).* IRC Papers Available from http://www.iea.nl

Glewwe, P., & Kremer, M. (2005). *Schools, Teachers, and Education Outcomes in Developing Countries*. Handbook on the Economics of Education Harvard University, Brookings Institution, Center for Global Development, and NBER.

Goldstein, H. (2004). International comparisons of student attainment: some issues arising from the PISA study. *Assessment Education, 11*(3), 319–330.

Goldstein, H. (2008). *Comment peut-on utiliser les études comparatives internationales pour doter les politiques éducatives d'informations fiables?* [How may we use international comparative studies to inform education policy]. *Revue française de pédagogie, 164*, 69-76.

Hopmann, S.T., Brinek, G., & Retzl, M. (2007). PISA zufolge PISA — PISA According to PISA. Hält PISA, was es verspricht?[Does PISA Keep What It Promises?]. *Reihe Schulpädagogik und Pädagogische Psychologie, Bd.6.* Wien: Lit-Verlag.

Koth, C.W., Bradshaw, C.P., & Leaf, P.J. (2008). A Multilevel Study of Predictors of Student Perceptions of School Climate: The Effect of Classroom-Level Factors. *Journal of Educational Psychology*, *100,* 96-104.

Kyriakides, L. (2006). Using international comparative studies to develop the theoretical framework of educational effectiveness research: A secondary analysis of TIMSS 1999 data, *Educational Research and Evaluation*, *12*(6), 513–534.

Kyriakides, L., Campbell, R.J., & Gagatsis, A. (2000). The Significance of the Classroom Effect in Primary Schools: An Application of Creemers' Comprehensive Model of Educational Effectiveness. *School Effectiveness and School Improvement*, *11*(4), 501–529.

Kyriakides, L., & Charalambous, C. (2005). Using educational effectiveness research to design international comparative studies: turning limitations into new perspectives, *Research Papers in Education*, *20*(4), 391–412.

Kyriakides, L., & Creemers, B.P.M. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: a study testing the validity of the dynamic model. *School Effectiveness and*

School Improvement, *19*(2), 183–205.

Kyriakides, L., & Creemers, B.P.M. (2009). The effects of teacher factors on different outcomes: two studies testing the validity of the dynamic model. *Effective Education, 1*(1), 61–86.

Kyriakides, L., Creemers, B.P.M., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies for school factors: Implications for theory and research. *British Educational Research Journal, 36*(5), 807–830.

Lingard, B., & Grek, S. (2008). The OECD, indicators and PISA: An exploration of events and theoretical perspectives. ESRC/ESF Research Project on Fabricating Quality in Education. *Working Paper 2* http://www.ces.ed.ac.uk

Martin, M.O. (1996). Third International Mathematics and Science Study. In *TIMMS Technical Report*, *Vol. 1* (pp.1.1-19). Boston: MA Boston College (IEA).

McDonald, S.K., Keesler, V.A., Kauffman, N.J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher, 35*(3), 15–24.

Mortimore, P. (2009). Alternative models for analysing and representing countries' performance in PISA. *Paper commissioned by Education International Research Institute Brussels*.

Mullis, I.V.S., Martin, M.O., Gonzalez E.J., Gregory, K.D., Garden, R.A., O'Connor, K.M., Chrostowski, S.J., & Smith, T. A. (2000). *TIMSS 1999*- International Mathematics Report. The International Study Center at Boston College: IEA.

Murphy, S. (2010). The pull of PISA: Uncertainty, influence, and ignorance. Interamerican Journal of Education for Democracy, 3(1), 28-44.

OECD (2002). *PISA 2000 technical report.* Paris: OECD.

OECD (2009). *Key Factors in Developing Effective Learning Environments: Classroom Disciplinary Climate and Teachers' Self-Efficacy* Paris: OECD.

OECD (2010). *PISA 2009 Results: What Makes a School Successful? – Resources, Policies and Practices (Volume IV).* Available from http://dx.doi.org/10.1787/9789264091559-en

OECD (2011). PISA 2009 Technical Report (preliminary version), Online at
http://www.oecd.org/document/19/0,3746,en

Perry, L.B. (2008). Using PISA to Examine Educational Inequality. *ORBIS SCHOLAE, 2*(2), 77–86.

Postlethwaite, T.N. (2006). What is a good cross-national study? In K. Ross & I.J. Genevois (Eds.), *Cross-national studies of the quality of education* (pp. 105-120). Paris: UNESCO-IIEP.

Reddy, S. (2007). *School Quality, Perspectives from the Developed and Developing Countries*. Azim Premji Foundation. Available from http://www.azimpremjifoundation.org

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., & Schaffer, G. (Eds.) (2002). *World Class Schools: International Perspectives on School Effectiveness.* London: RoutledgeFalmer.

Reynolds, D., Sammons, P., De Fraine, B., Townsend, T., & Van Damme, J. (2011). Educational Effectiveness Research (EER): A State of the Art Review. Paper presented to the annual meeting of the International Congress for *School Effectiveness and Improvement,* January 4-7, *2011,* Limassol, *Cyprus.*

Riddell, A. (2008). *Factors Influencing Educational Quality and Effectiveness in Developing Countries: A Review of Research.* Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) and the German Federal Ministry for Economic Cooperation and Development (BMZ).

Scheerens, J., & Bosker, R.J. (1997). *The foundations of educational effectiveness.* Oxford: Pergamon.

Smithers, A. (2004). England's Education: what can be learned by comparing countries? Liverpool: *Centre for Education and Employment Research.*

Teddlie, C., & Reynolds, D. (2000). *The International Handbook of School Effectiveness Research*. London: Falmer Press.

Trowler, V. (2010). *Student engagement literature review.* Lancaster University: Department of Educational Research.

Willms, J.D. (2003). *Student engagement at school: A sense of belonging and participation results from PISA 2000.* Paris: OECD.

Yang, Y. (2003). Dimensions of Socio Economic Status and their Relations to Mathematics and Science