An Investigation of International Science Achievement Using the OECD's PISA 2006 Data Set

Dr. Todd M. Milford, University of Victoria, Canada tmilford@uvic.ca
Dr. John O. Anderson, University of Victoria, Canada anderson@uvic.ca

Abstract

This study uses hierarchical linear modeling (HLM) to analyze data from PISA 2006 for nations experiencing high rates of immigration (i.e., Germany, Spain, Canada, the United States, Australia and New Zealand). The outcome measures used were achievement scores in science (i.e., scientific literacy). The variables examined at the student level were science self-efficacy, science self-concept, immigrant status and socioeconomic status. The variables examined at the school level were student level aggregates of school proportion of immigrants and school socioeconomic status. In the HLM null models, the intraclass correlations for the all countries except for Germany ranged from .16 to .29 (Germany's was between .57 and .68). In the final models, at level-1 country, immigrant status tended to negatively influence achievement (i.e., non-native students are predicted to have lower performance), while science self-efficacy and science self-concept positively influenced achievement. The student level ESCS variable also impacted achievement positively. At the school level, level-2, school mean ESCS or school proportion of immigrants was found to significantly influence the level-1 predictors; however, a good deal of variability across nations was observed. The findings from this study demonstrate distinct national differences in the relationships between science self-beliefs, immigrant status and academic achievement.

Keywords: Science Achievement, PISA 2006, HLM, Immigrant Students

Subject Problem

The theoretical perspective which proposes that student background characteristics are more important to academic success than school characteristics is generally linked to the work of Coleman, Campbell, Hobson, McPartland, Mood, Winfield, and York (1966) and Jencks, Smith, Acland, Bane, Cohen, Gintis, Heyns, and Michelson (1972). These studies came to similar conclusions regarding the amount of variance that can be explained by educational factors. After taking into account student background characteristics (e.g., ability and family background), little variance in student achievement remained (Creemers, 2006). These theories propose heterogeneity of student responses to school characteristics.

As a direct response to this theoretical perspective, studies in school effectiveness research (SER) emerged through the works of George Weber and Ron Edmonds (Raptis & Fleming, 2003). School effectiveness research proposes that specific school attributes are associated with academic achievement and argues that substantive differences in student performance among schools can be ascribed to the quality of the schooling itself (Goldstein & Woodhouse, 2000). A major proposition of this framework is that schools are the foremost factor influencing academic success, and the way to improve student outcomes is to identify and reproduce characteristics of good schools. An implicit assumption in this framework is that students are relatively homogeneous and respond equally to school characteristics.

The field of SER has called for – among other things – the expansion of studies from the traditional ethnocentric approach to a more cross-national one (Creemers, 2006). Factors that appear to work in one country cannot be assumed to work in another. Kyriakides (2006) argues that there is importance in conducting comparative studies in order to identify direct and indirect effects upon student achievement at the national level.

This study informs and builds on cross-cultural comparative theory by examining if differences in academic achievement are associated more with individual student or school membership. Thus, the examination of whether or not students are homogenous within schools and, if not, which differences in background are significant will be useful in understanding variation in student outcomes. It focuses on the identification of select factors that account for the student and school variance in scientific literacy in six nations of high immigration (i.e., Australia, Canada, Germany, New Zealand, Spain, and the United States) in the Programme for International School Achievement (PISA) 2006 data set.

Design Procedure

PISA provides a publicly accessible, large, and significant data base as an "ongoing, periodic international comparative study of the proficiency in Mathematics, Science and Reading of 15 year old students" (Turner & Adams, 2007, p. 238). PISA has been administered in 2000, 2003, 2006, and 2009, with current planning projected to 2015. After each assessment cycle, the OECD makes available the results as well as comprehensive reports detailing literacy outcomes and background variables (i.e., student, family, and school and system factors). These PISA publications may be accessed at the OECD website: http://www.pisa.oecd.org.

Hierarchical linear modeling (HLM) (Byrk & Raudenbush, 2002) is a data analysis technique used to address hierarchical data structures that is a more advanced form of simple and multiple linear regression (Snijders & Bosker, 1999). Compared to classical regression, multilevel (hierarchical) modeling is useful for data reduction and essential for prediction. One of the assumptions underlying traditional regression is that observations of one individual are not systematically related to those of another (Willms, 1999). This assumption is violated with data such as PISA when student level and school level data is combined into an analysis. Multilevel analysis allows variance in outcome variables to be analyzed at multiple hierarchical levels, between-group and within-group; thus, is appropriate for use with nested data such as PISA.

The goal of multilevel analysis is to predict values of some dependent variable based on a function of predictor variables at more than one level (Luke, 2004). In the PISA situation, HLM uses level-1 student and level-2 school variables to help explain variation in scientific literacy scores while accounting for the variance at each level. HLM allows for estimating both the effects of group-level variables on school mean score and on the slopes of the individual characteristics to predict outcome variables (Willms & Smith, 2005).

Here, the relationship between socio-economic status, immigrant status, science self-beliefs, and science literacy for the countries of Australia, Canada, Germany, New Zealand, Spain, and the United States are modeled using hierarchical linear modelling (HLM). The rationale for choosing these six countries for comparison is simply that the United Nations, in its most recent report on world population, has identified Western Europe, North America, and Oceania as major destinations of net migration (UN, 2006). These counties are representative of those regions that will receive the majority of these immigrants.

To investigate school effects on the SES, immigrant and science self-beliefs (science self-concept and science self-efficacy) gradients two models were developed. Model one (the baseline model) consisted of student SES (ESCS in the PISA dataset) immigrant and science self-beliefs in level-1. The level-2 model had no school level predictors but both the intercept and the slopes included random school-level variation which allowed for an estimation of the association of student-level SES, immigrant status and science self-belief on science literacy. The level-2 models added school-level predictors of school SES and school proportion of immigrants to both the intercept term and the slope equations:

Level-1

$$\text{Science}_{ij} = \beta_{0j} + \beta_{1j}\text{SES}_i + \beta_{2j}\text{Non-Native}_i + \beta_{3j}\text{Science-Belief}_i + \text{error}_{1j}$$

Level-2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{School trait } 1_j + \text{error}_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}\text{School trait } 1_j + \text{error}_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}\text{School trait } 1_j + \text{error}_{2j}$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31}\text{School trait } 1_j + \text{error}_{3j}$$

A significant coefficient on the predictor in the intercept equation (γ_{01}) indicated an association to school mean scientific literacy scores. A significant coefficient on the school-level predictor in the SES, and/or immigrant gradient equation (γ_{11}) indicated an association of the school trait with the relationship of student-level predictor to scientific literacy. The analyses were conducted using the computer analysis program SAS (Statistical Analysis Software, nd) and the subroutines Proc Mixed (available from the PISA website) with normalized student final weighting (OECD, 2007, Annex A8).

Findings and Analyses

The null models for the PISA 2006 science results (i.e., the intra-class correlation [ICC] or proportion of student achievement variance attributable to schools) were similar across all domains for Spain, Canada, Australia, New Zealand, and the United States (p = .16–.29). See Table 1. However, the ICC was considerably higher in Germany at .57 for science. This is not surprising as Germany formally tracks students (Marks, 2005). Because of this, students in Spain, Canada, Australia, New Zealand, and the United States will account for more variance at the student level. Predictably, with the addition of student-level variables, the within-school variance is reduced by approximately 20–30% for Spain, Canada, Australia, New Zealand, and the United States (with Germany lower because it had less variance to model at this level).

Table 1
Intraclass Correlations (ICC) Derived from the Null Models for All Countries and Science

Country	ICC for science literacy
Australia	.19
Canada	.20
Germany	.57
New Zealand	.17
Spain	.18
The United States	.23

The conditioned means for the six countries analyzed for scientific literacy were split into two groups with Australia, Canada, Germany and New Zealand (range of 507.17–531.90) consistently above the OECD average (i.e., mean of 500, standard deviation of 100) compared to Spain, and the United States (range of 473.85–495.64) consistently below.

In the final model for science literacy in Australia, immigrant status, science self-efficacy, science self-concept, and student ESCS significantly predicted achievement at the student level. On the intercept, school-level ESCS was significant at level 2, and there was significant variability observed across schools. Additionally, at the school level, attending a school with a one-unit greater mean ESCS will increase the predicted school mean score (originally 519.99) by almost 49 units (half a standard deviation). The student immigrant slope is also made steeper by attending a school with all immigrant students by 25.67 units. The self-concept slope is increased by 5.92 units as the school mean ESCS increases by one unit. The immigrant, science self-beliefs, and the ESCS slope varied significantly across schools.

Each of the five other nations included in this section of the analysis (i.e., Canada, Germany, New Zealand, Spain, and the United States) showed the same significant predictors at the student level as well as relative size and direction of influence on the conditioned mean (e.g., immigrant status negatively influenced the conditioned mean, while science self-efficacy, science self-concept, and ESCS positively influenced the conditioned mean). Additionally, school-level ESCS was also significant at level 2 on the intercept with significant variability across schools for all six nations. However, variability was observed at the school level. The overall final models are summarized in Table 2.

Table 2
Final Models for Scientific Literacy for Countries included in this Study

	Parameter Estimate by Nation*					
Variable	Australia	Canada	Germany	New Zealand	Spain	U.S.A.
Fixed effects						
Intercept (γ_{00})	519.99	513.17	493.20	535.05	505.89	477.39
School ESCS (γ_{01})	48.97	42.39	92.94	54.61	21.68	44.39
Immigrant Slope						
Intercept (γ_{10})	-22.76	-21.24	-36.87	-22.61	-42.63	-17.07
Proportion immigrants (γ_{11})	25.67					
Self Efficacy Slope						
Intercept (γ_{20})	26.42	25.25	18.77	37.75	22.63	22.92
Self Concept Slope						
Intercept (γ_{30})	21.34	19.68	12.08	13.61	15.08	11.52
School ESCS (γ_{31})	5.92			14.22	3.67	13.03
ESCS Slope						
Intercept (γ_{40})	14.99	12.99	6.61	19.46	14.75	21.59
School ESCS (γ_{41})				-11.65		
Proportion immigrants (γ_{42})				26.39		
Random Effects						
Intercept (μ_{0j})	370.37	1267.43	1177.75	405.34	506.15	1034.94
Immigration slope (μ_{1j})	305.64	1091.73			1633.39	327.01
Science self efficacy slope (μ_{2j})	55.43	77.62		54.60	18.69	58.15
Science self concept slope (μ_{3j})	48.15	107.79		96.51	47.18	
ESCS slope (μ_{4j})	66.65	48.69			39.22	
Student level effect (*r*)	5708.69	5406.42	3657.99	6191.57	4889.75	6399.07

*Significant at $p<.05$

Contributions to the Teaching and Learning of Science and to NARST members

There are a number of reasons why this study is important in the development of cross-cultural theory and practice in school effectiveness research. Firstly, the migration of people from one nation to another between 1990 and 2000 reached an all-time high of 2.6 million annually and is projected to average 2.2 million for the next 50 years (United Nations Department of Economic and Social Affairs/Population Division [UN], 2006). This increase has given rise to a greater research emphasis on understanding the heterogeneity of immigrant academic achievement (e.g., Ammermueler, 2007; Huang 2000; Marks 2005). A nation's school system plays a critical role in educating immigrant children and facilitating their participation in the larger society (Huang, 2000). However, numerous studies have documented lower achievement for immigrant students compared to non-immigrant students on international academic assessments (Marks, 2005). Additional studies have uncovered reasons for these lower results in immigrant academic achievement, including differing levels of socio-economic status (SES), home background, and motivation (Blair & Qian, 1998). A key objective is to understand and explore student science achievement in the framework of these background variables across nations of high immigration.

Additionally, large datasets such as PISA developed by the Organization for Economic Cooperation and Development (OECD) can be utilized to help uncover what works and what does not work for students. Studies such as PISA have demonstrated sizable variations in educational outcomes (i.e., reading, mathematics, and science) between countries. Consistently, countries such as Canada, Finland, Japan, Singapore, and South Korea rate much better on academic outcome measures than the other countries, and the same countries consistently come out ahead in the rankings regardless of the domain measured ("How to be on top", 2007). However, the achievement scores from programs such as PISA typically provide limited information, due either to the way they are interpreted or the way the results are disseminated.

According to Anderson, Rogers, Klinger, Ungerleider, Glickman, and Anderson (2006, p. 707), "The publication of these rankings implies that variation in student performance is solely due to school effects". Simple rankings of nations compared to the OECD mean score (see http://www.pisa.oecd.org)—known as the *league tables*—fail to account for the contextual relationships of scores to background traits of the students, schools, or communities.

Ancillary to the usefulness of large international databases such as PISA is the recognition that science literacy is one of the three domains measured (along with mathematical literacy and reading literacy). In our current technologically based society, an understanding of fundamental scientific concepts and theories and how this can be applied to the challenges we face collectively is more important than ever (OECD, 2006). Mathematics and science preparation is linked to future needs (Wang, 1998), and no nation can afford poor academic achievement or high dropout rates among its young people without jeopardizing its economic future (Raptis & Fleming, 2003). Hanushek, Jameson, Jameson, and Woessmann (2008) used a country's performance on international tests of mathematics and science to demonstrate that those countries with higher test scores experienced far higher economic growth rates on the order of 10% of gross domestic product (GDP) over the last half century.

In this study, it is argued that one possible approach to understanding more about why some variables explain effectiveness across countries while others do not is through the use of multilevel educational effectiveness models by conducting secondary analyses of data from international comparative studies. If patterns of significant student-level and school-level variables are relatively stable across countries, then we expect few differences in how school systems support students. However, if differences exist in how student achievement is supported across national school systems, then we expect these differences to emerge in the final models.

References

Ammermueler, A. (2007). Poor background or low returns? Why immigrant students in Germany perform so poorly in the Programme for International Student Assessment. *Education Economics*, *15*(2), 215–230.

Blair, S. L., & Qian, Z. (1998). Family and Asian students' educational performance. *Journal of Family Issues*, *19*(4), 355–374.

Byrk, A., & Raudenbush, S. W. (2002). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Winfield, F., & York, R. (1966). *Equality of educational opportunity*. Washington: US Government Printing Office.

Creemers, B. P. M. (2006). The importance and perspectives of international studies in educational effectiveness. *Educational Research and Evaluation*, *12*(6), 49–511.

Goldstein, H., & Woodhouse, G. (2000). School effectiveness research and educational policy. *Oxford Review of Education*, *26*(3/4), 353–363.

Hanushek, E. A., Jameson, D. T., Jameson, E. A., & Woessmann, L. (2008). Education and economic growth: It's not just going to school, but learning something while there that matters. *Education Next.* Retrieved from http://findarticles.com/p/articles/mi_m0MJG/is_2_8/ai_n25128922.

How to be on top. (2007, October 18). *The Economist,* 81–82.

Huang, G. C. (2000). Mathematics achievement by immigrant children: A comparison across five English-speaking countries. *Education Policy Analysis Archives*, *8*(25). Retrieved from http://epaa.asu.edu/epaa/v8n25/.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: A reassessment of the effects of family and schooling in America*. New York: Basic Books.

Kyriakides, L. (2006). Introduction international studies in educational effectiveness. *Educational Research and Evaluation*, *12*(6), 489–497.

Luke, D. A. (2004). *Multilevel modeling.* Thousand Oaks, CA: Sage.

Marks, G. N. (2005). Accounting for immigrant non-immigrant differences in reading and mathematics in twenty countries. *Ethnic and Racial Studies*, *28*(5), 925–946.

UN (United Nations Department of Economic and Social Affairs/Population Division). (2006). *World population prospectus*: The 2006 Revision, Volume III: Analytical Report.

OECD (Organisation for Economic Co-operation and Development). (2007). *OECD's PISA survey shows some countries making significant gains in learning outcomes.* Retrieved from http://www.oecd.org/dataoecd/47/61/35070367.pdf.

Raptis, H., & Fleming, T. (2003). Reframing education: How to create effective schools. *C.D. Howe Institute Commentary*, *188*, 1–24.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: U.K.: Sage.

Turner, R., & Adams, R. J. (2007). The Programme for International Student Assessment: An overview. *Journal of Applied Measurement*, *8*(3), 237–248.

UN (United Nations Department of Economic and Social Affairs/Population Division). (2006). *World population prospectus*: The 2006 Revision, Volume III: Analytical Report.

Wang, J. (1998). International achievement comparison: Interesting debates on inconclusive findings. *School Science and Mathematics*, *98*(7), 376–382

Willms, J. D. (1999). Basic concepts for hierarchical linear modeling with applications for policy analysis. In G. J. Cizek (Ed.), *Handbook of educational policy* (pp. 473–491). San Diego, CA: Academic Press.

Willms, J. D., & Smith, T. (2005). *A manual for conducting analyses with data from TIMSS and PISA. Report prepared for the UNESCO Institute for Statistics*. Retrieved from http://www.unb.ca/crisp/pdf/Manual_TIMSS_ PISA2005_0503.pdf.