Running head: Hierarchical Linear Models in School Boards

The Use of Hierarchical Linear Models in School Boards: Reflections from Applied Researchers

Greg Rousell
Halton Catholic District School Board

Virginia Sheriff
Limestone District School Board

Rosanne Brown
Peel District School Board

Kathy Russell-Kwan
Dufferin-Peel Catholic District School Board

Chris Conley
Durham District School Board

Paper Presented to:
The Canadian Society for the Study of Education, CERA
May 23, 2009

Contact:
Greg Rousell, B.A.
Halton Catholic District School Board
802 Drury Lane, Burlington, ON L7R 2Y2
RousellG@hcdsb.org
905-632-6314 ext. 326

Abstract

Applied researchers in schools boards from across Southern and Eastern Ontario explored the use of Hierarchical Linear Models. Community level variables were aggregated at the school level to explore the effect on student achievement on a large scale assessment. Consistent with previous research, socio-economic and demographic factors predicted achievement on the assessments (Willms 2003a); however the variance explained by these factors differed between school boards. This paper describes general trends in results across the participating school boards, a discussion of the findings, and a description of the professional learning community in Ontario that allowed for the exploration of this type of analysis.

Introduction

The creation of the Education Quality and Accountability Office (EQAO) by the Ontario provincial government and the use of its data have raised the profile of accountability for student achievement across the province's publicly funded school districts. Subsequent initiatives from Ontario's Ministry of Education (e.g., Managing Information for Student Achievement, Literacy and Numeracy Secretariat, Student Success, and Learning for All) have continued to focus on the use of data in the education system. School boards have largely used the descriptive statistics provided by EQAO or Student Success for their decision making; however, a large body of research has explored the relationship between student achievement and community level factors (Raudenbush & Willms, 1995; Willms, 2003a). As part of Ontario's Hierarchical Linear Models – Learning and Research Community (HLM-LRC) a group of researchers from both English public and Catholic school boards located in Southern and Eastern Ontario collaborated to explore the use of HLM with their achievement, demographic, and community level data.

This paper presents the methodology, experience, and reflection of applied researchers as they explore the extent to which student factors and community contexts impact EQAO achievement on the EQAO Junior Division (Grades 4-6) Assessments of Reading, Writing and Mathematics.

Hierarchical Linear Models – Learning and Research Community

Sponsored by the Canadian Centre for Knowledge Mobilization (CCKM), the Ontario Hierarchical Linear Models - Learning and Research Community (HLM-LRC) was established in June 2007 to build capacity in the practice of HLM. Membership includes researchers in school boards, educational associations and organisations, Ministry of Education research

groups, universities, learning organisations, and service delivery organisations (i.e., health care providers). Participant familiarity with the software and process to conduct an HLM analysis was varied. Several members had prior HLM training through university courses and training workshops such as KSI International's HLM statistical modelling training workshop, led by Dr. J. Douglas Willms. Other members had no prior experience or knowledge of HLM. While the structure of school board data makes HLM an ideal method of analysis, it is not used regularly by applied researchers in education. The goal of the HLM – LRC is not only to promote and expand the use of HLM but also to create a network and develop and share resources. As member familiarity and utility of HLM increases, the network will continue to serve to facilitate and promote deeper discussions regarding the application of HLM to a variety of data sets and contexts.

In the interest of providing hands-on opportunities to develop resources and expand the experience and knowledge of members, the HLM-LRC initiated three projects to explore education data in a variety of contexts and levels. The three projects explored education data at national, provincial and local school board levels and included data from:

- The Organisation for Economic Co-operation and Development Programme for International Student Assessment (PISA),
- Education Quality and Accountability Office (EQAO) provincial data,
- Board level EQAO and student data sets.

These projects provided an opportunity not only to increase the capacity for HLM analysis but also to develop a variety of resources. The resources developed from these projects (i.e., Workbooks, HLM organizers etc.) have been posted on the HLM-LRC wiki (http://HLM-LRC.wikispaces.com) and have been designed to make HLM conceptually and procedurally

accessible regardless of previous experience or background. These resources are available upon request. Regularly scheduled meetings provided discrete opportunities to work in a focused manner which, given the regular demands of the applied researchers' roles, could not have been afforded otherwise.

The Use of HLM in School Boards

Doug Willms of the Canadian Research Institute for Social Policy (CRISP), University of New Brunswick, has been a leader in research exploring relationships between community measures and academic achievement (Raudenbush & Willms, 1995; Willms 2003a; Willms, 2003b; Willms, 2009). Following this line of research, project participants were interested in exploring the relationships between student (e.g., gender, English Language Learner needs, and special education information) and community factors and student achievement on EQAO's Junior Assessments of Reading, Writing and Mathematics. Project participants worked collaboratively to refine the proposed data sets by identifying data common to all participants. Student demographics used by participants were all common and provided in the EQAO source data. While determining a common community data set was initially anticipated to be a challenge, it was quickly determined that all participants were using the Social Risk Index (SRI) (HRDC, 2003) for a variety of local projects and initiatives (details of the SRI can be found in subsequent sections of this paper). The resources developed to support this group and opportunities to troubleshoot and problem solved were considered to be invaluable opportunities to foster a greater understanding of the HLM.

School Board Data

Using hierarchical linear models for school board data has the potential for increased accountability given the demand for data-based (data-informed) decision-making (Klinger,

2007). The dominant outcome data come from EQAO as all of Ontario's publicly funded schools are required to administer the assessments in Grades 3 and 6. Large scale assessments have potential to enhance the understanding of student achievement and school performance on important educational outcomes. The information these assessments provide can (with appropriate analyses and interpretation) lead to a better understanding of and support for the performance and quality of schools and education (Klinger, Rogers, Anderson, Poth & Calman, 2006).

The relationships of schooling to student achievement are complex. Combining relevant contextual and process variables can increase the understanding of education and support quality practices in education. The benefit of using HLM in educational research is that these analyses acknowledge the nested structure of educational systems and simultaneously includes the influence of factors at different levels of the system (Klinger, 2007). HLM can separate fixed effects and random effects as well as identify the variation at the different levels of the data. However, this type of analysis does have some limitations. A relatively large and complete sample is needed and it is cautioned that the same results may be obtained through regression procedures (Klinger, 2007). Furthermore, while a great deal of information is currently collected at the both the student and school level, very little is collected at the classroom level.

Education Quality and Accountability Office (EQAO).

EQAO is an independent, arm's-length agency of the Ontario provincial government that provides parents, guardians, teachers, and the public with reliable and valid information about student achievement. EQAO reports provide information for educators, parents, guardians, policy makers and others in the education community to use to improve student learning and instructional practice.

EQAO conducts a range of province-wide assessments. The Grade 3 (or Primary) assessment was instituted in 1996-1997. This assessment involves all students in publicly-funded schools, occurs annually, and provides information on what students have learned in reading, writing, and mathematics. In 1998-1999, EQAO introduced an annual Grade 6 (or Junior) assessment that measures student achievement in reading, writing and mathematics up until the end of Grade 6.

Social Risk Index (SRI).

The conceptual model for the Social Risk Index (SRI) was developed by Human Resources Development Canada (HRDC) as a tool for providing a general picture of potential risks in communities. The index contains nine variables that profile the socio-economic context of communities (HRDC, 2003). The variables, drawn from Statistics Canada's Census, include:

- Average household income;
- Unemployment rate;
- Proportion of residents 20 years and older who lack a high school diploma;
- Proportion of owner-occupied dwellings;
- Mobility over one year;
- Knowledge of Canada's official languages;
- Proportion of recent immigrants;
- Proportion of lone parent families;
- Percentage of household income that is comprised of government payments.

These variables were aggregated to provide a single index score.

Methodology

The data sets identified for use by participants included the following:

- Outcome: Junior Division (Grades 4-6) EQAO Assessments of Reading, Writing and Mathematics.
- Level 1 Student Data: Area (each school board identified regions according to their particular geographical context); Gender; English Language Learners (ELL); students with Individual Education Plans (IEPs) (excluding gifted); and Gifted students.
- Level-2 School Data: School Size; Area; Social Risk Index (SRI); and Income (for some school boards only).

Construction of the Social Risk Index (SRI)

Statistics Canada uses geographic areas to report census data, starting nationally and then disaggregating the information at each smaller level (provincial, regional, municipal, etc.). The smallest level of geography that Statistics Canada reports on, and that school boards are interested in, is the Dissemination Area (DA). Each DA is bounded by a variety of features including roads, rivers, railways, and transmission lines which create neighbourhoods comprised of 400 to 700 people (Statistics Canada, 2007). It is important to note that since a DA is defined by the number of people within its boundaries, DAs can have very different sizes. One very small DA in an urban area may include an apartment complex whereas a very large DA in a rural area may include several small communities and farms.

To create the Index, school boards compare the DA information for the nine variables to a cut-point for risk. The method to calculate the cut-points varied slightly between boards. In some boards, the regional average for the variable was used as the cut-point whereas other boards used a median value or calculated the 95% Confidence Interval for each variable. When

using the Confidence Interval, DAs that were outside of the 95% Confidence Intervals were considered to be at risk. When a DA falls outside of the cut-point it is coded as a 1 to represent risk. This recoding standardizes the direction of the risk, for example as the unemployment rate increases, the associated risk increases. On the other hand, as average household income decreases, risk increases. After recoding all of the SRI variables for each DA, a new variable is created that sums the recoded values across all variables. This new variable presents the level of social risk for each DA according to a 9 point scale. School level SRI is then calculated by averaging the scores across *all* students in a school.

For each of the boards, the SRI has been an opportunity to discuss the value and utility of community measures and also provides a common ground for discussions between boards. With the evolution of local discussions, boards have begun to modify the SRI to address local concerns and realities. In addition to accommodating board contexts, the SRI is also being adapted to address changes made by Statistics Canada in their 2006 Census for the collection and reporting of variables. Recent work by Brown (2009) has prompted the review of these data within school boards.

Performing Hierarchical Linear Modelling using SPSS 17.0 Advanced Statistics Compared to the HLM Software Package

One board opted to perform hierarchical linear modelling using the Statistical Package for the Social Sciences (SPSS). The software version used was the most recent release (17.0) of the Advanced Statistics module. SPSS employed different procedures and output interpretations than those associated with the HLM software package. Four notable differences between the two statistical packages were observed: (1) the source data file requirements; (2) different

terminology used in the statistical tools; (3) the procedures used to build the models; and (4) the output generated and its interpretation.

With respect to the differences between the source data file requirements, where HLM requires the creation of separate data files for each variable level, SPSS requires the creation of a single file with all variables at all levels. The SPSS software can also handle missing data and does not require any sorting of the level variables. The HLM software package, on the other hand, requires both the Level-1 and Level-2 files to be sorted by the common identifier and cleaned of any missing data. Furthermore, while initially a challenge, the differences in terminology used by each statistical package was quickly overcome. The challenge this posed was related more to discussing results across the two packages than in completing the analyses.

Looking at the differences between SPSS and HLM in constructing models, SPSS uses scripts or it's menu system of dialogue boxes to build them whereas HLM builds a model using a visual representation of the formula $y|x = \beta_{0+} \beta_1 x_1 + \ldots + \beta_i x_i + r$ ($y = mx+b$), where $\beta_0=\gamma_{00}+\mu_0$. Another difference between the two software packages is that SPSS reports on the variance contributed by the multi-level model without distinction between the proportions of variance contributed by each level in the model. As this inquiry represented an initial attempt at using the HLM facility in SPSS, it is not clear whether or not this difference in data output would be problematic for practical analyses.

Once the terminology, procedural, and interpretive differences are learned, using SPSS 17.0 Advanced Statistics to perform hierarchical linear modelling yields useful information. There were clear differences between the software packages in the amount and type of output. Unfortunately, time did not permit comparisons between analyses of the same data run using the HLM package and SPSS 17.0 Advanced Statistics.

Results

The results reported in this paper are general trends rather than specific school board results. This limited reporting of results is intended to maintain privacy and confidentiality of the individual school boards. The outcomes data examined are limited to a single year of EQAO data for the junior division assessments.

EQAO Junior Division (Grades 4-6) Assessment of Reading

The variance attributed to the two levels differed between boards. In the Null Models, Level-2 (school level) variances ranged from 5.1% to 15.4% of the total variance.

Level-1 variables.

All boards found the gender and IEP (excluding gifted) variables to be significant predictors of student achievement (p <.000) in reading. Females outperformed males on the assessment and students with IEPs (excluding gifted) underperformed compared to students without IEPs (excluding gifted). Students identified as gifted had higher achievement than non-gifted students at most of the participating boards (p <.000). English language learners were significant for only one board (p <.000), with these students underperforming compared to other students. Level-1 models accounted for between 11.4% and 23.5% of the total variance.

Level-2 variables.

All boards except one found the Social Risk Index (SRI) to be a significant predictor of student achievement (p values ranged from .001 and .051). Students at schools with lower levels of risk had higher scores on the assessment. Where SRI was not significant, income was a significant predictor (p=.018); students at schools with higher median incomes had higher scores on the assessment.

Most of the boards had areas that were significantly different than others (*p* values ranged between .000 and .006). Level-2 models accounted for between 2.9% and 5.4% of the total variance. The variance accounted for by the complete models ranged between 16.8% and 28.2% of the total variance.

EQAO Junior Division (Grades 4-6)Assessment of Writing

The variance attributed to the two levels differed between boards. In the Null Models, Level-2 (school level) variances ranged from 5.4% to 14.9% of the total variance.

Level-1 variables.

All boards found the gender and IEP (excluding gifted) variables to be significant predictors of student achievement in writing ($p < .000$). Females outperformed males on the assessment and students with IEPs (excluding gifted) underperformed compared to students without IEPs (excluding gifted). Gifted students had higher achievement than non-gifted students ($p < .000$) for most of the participating boards. English language learners were significant for only one board ($p < .000$), with these students underperforming compared to the other students. Level-1 models for writing accounted for between 7.6% and 21.6% of the total variance.

Level-2 variables.

All boards except one found the Social Risk Index to be a significant predictor of student achievement (*p* values ranged from .001 and .025). Students at schools with lower levels of risk had higher scores on the assessment. Where SRI was not significant, income was a significant predictor ($p=.015$) where students at schools with higher median incomes had higher scores on the assessment.

Most of the boards had areas that were significantly different than others (*p* values ranged between .000 and .003). Level-2 models accounted for between 0.9% and 6.9% of the total variance. The variance accounted for by the complete models ranged between 8.5% and 25.7%.

EQAO Junior Division (Grades 4-6) Assessment of Mathematics

The variance attributed to the two levels differed between boards. In the Null Models, Level-2 (school level) variances ranged from 5.0% to 24.8% of the total variance.

Level-1 variables.

All boards found IEP (excluding gifted) to be a significant predictor of student achievement in mathematics ($p < .000$) with these students having lower achievement on the assessment. Students identified as gifted were found to be significant at most boards ($p < .000$) with these students performing better on the assessment. English language learners were found to be significant for one board ($p < .001$) with these students having lower achievement on the assessment. Level-1 models accounted for between 2.4% and 12.4% of the variance.

Level-2 variables.

One board reported Social Risk Index as significant ($p < .000$) and another board reported that it approached significance ($p = .092$). One board reported an area that was significantly different than others ($p=.016$). The variance accounted for Level-2 ranged between 1.7% and 4.3%. The variance accounted for by the complete models ranged between 2.4% and 15.4%.

Discussion

Given the use of similar data sets and methodology, of particular note in the results are the differences in the variances attributable to Level-2. In these analyses, the Level-2 variances in the Null Models indicate how much variation in the student's achievement can be attributed to

the schools in which they are nested. In boards where this variance is low, the school that the student attends appears to have little impact on his or her achievement.

Not surprisingly, two variables were consistent predictors among all participating boards for the literacy assessments (reading and writing). Females and gifted students consistently scored higher on the assessments than males and non-gifted students (respectively), and students with IEPs (excluding gifted) had lower achievement. Only one board reported that students identified as English language learners had lower predicted performance on the assessment; however, this finding may be due to lower overall numbers of these students in other boards.

Level-2 variables were fairly consistent among school boards. The SRI or income (as a proxy measure of risk) predicted achievement and most boards had schools in geographic areas that they were able to identify as over or underperforming on the assessment. While these variables are significant, they account for a very low proportion of the overall variance in student achievement. At the most, one board was able to attribute 8.6% of the *total* variance to all of the Level-2 factors.

Students' achievement in mathematics followed a similar pattern for Level-1 variables with the notable exception that the effect of gender was not significant. IEP (excluding gifted), Gifted and English language learners were all predictors. However, the Level-2 variables were very different. Only two boards had the SRI as predictive of student achievement in mathematics (one of these boards reported that the variable only approached significance) and only one board was able to identify a geographic area that predicted achievement. The majority of the boards had a much lower variance explained by these models than the models for the literacy assessments.

Challenges

Some of the challenges that the group experienced were the following:

1. Time –Meeting on a regular basis was somewhat problematic due to the limited time of the researchers.
2. Geography – The researchers were affiliated with school boards from across Southern and Eastern Ontario.
3. Experience Level – Participants had a range of knowledge and experience that needed to be integrated so that all could participate and contribute.
4. Different Software Used – One of the challenges of using SPSS 17.0 Advanced Statistics to perform hierarchical linear modelling instead of HLM software was the additional learning curve required. While most subgroup members used the HLM software package, the resources that were developed did not provide support for the SPSS module.

Next Steps

This project has been invaluable for increasing the capacity of participating school boards in using this HLM analysis. Moving forward, group participants would like to continue with the collaborative professional learning community model to look at other areas of mutual interest. Participants are interested in both transitioning from a focus on process to product and to expanding knowledge of other techniques. This project has been extremely successful in building a network of people with similar professional interests who are committed to ongoing learning and collaboration.

The primary objective of this project was to learn collaboratively about HLM as a tool to explore the nested nature of student and school data within school boards. As such, it is the

collective learning process rather than the product of individual analyses that was the focus of the project and this paper. It is important to note that this focus on the process does not diminish the value of the individual results. In fact, these findings are now being used for evidence informed decision making internally within participating school boards.

The range and availability of data in school boards have increased significantly in recent years. There are currently a number of education initiatives in place in Ontario that are related to data and evidence, the result being that school boards have access to more data than ever before. From the group's collective experience, the majority of analyses within school boards have traditionally been descriptive statistics focused primarily on student achievement data. This project has also been an opportunity to look at the impacts of other factors on student achievement using inferential statistics. This project has also allowed the participants to build local capacity for further analysis in this area. Through the exploration of HLM the group has found that it is a valuable tool that fits well with the nested nature of much of school board data. It should be noted that HLM remains one tool that researchers can use. However, in situations where the Level-2 variance in the Null Model variance is very low there may be little benefit to this type of analysis.

The professional learning community model is a powerful vehicle for exploring new ideas and building on the knowledge and strengths of all members of the community. The members of this subgroup brought a range of knowledge and experience to the project. The participants all found this process to be an invaluable networking and learning opportunity. The number of research personnel within school boards is generally quite limited, so participation in a collaborative learning project such as this is a welcome chance to further professional knowledge in an applied setting.

References

Brown, R. (2009). *The learning opportunities index: An overview.* Adobe Connect PowerPoint presentation presented at the Association of Educational Researchers of Ontario (AERO) Special Interest Group (SIG) The Use of Community Measures. Burlington, ON: April 22.

Human Resources Development Canada (HRDC). (2003). *Understanding the early years: Early childhood development in the Dixie Bloor community of Mississauga, Ontario.* Ottawa: Human Resources Development Canada.

Klinger, D. A. (2007). *Examining the potential use of HLM to inform educational practice and policy.* Adobe Connect PowerPoint presentation presented at the Association of Educational Researchers of Ontario (AERO) Special Interest Group (SIG) Hierarchical Linear Modeling (HLM), Kingston, ON: April 16.

Klinger, D. A., Rogers, W. T., Anderson, J. O., Poth, C. & Calman, R. (2006). Contextual and school factors associated with achievement on a high-stakes examination. *Canadian Journal of Education, 29*(3), 771-797.

Raudenbush, S. W. & Bryk, A. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Second Edition. Thousand Oaks, CA: Sage publications, Inc.

Raudenbush, S. W. & Willms, J. D. (1995). *The estimation of school effects.* Journal of Educational and Behavioral Statistics, 20(4), 307-335.

Statistics Canada (2007). *Dissemination area: detailed definition.* Retrieved, May 1, 2009 from, Statistics Canada Web site: http://geodepot.statcan.ca/Diss2006/Reference/COGG/LongDescription_e.jsp?GEO_LEVEL=35&REFCODE=10&TYPE=L

Williams, F. (1983). *Executive communication power: Basic skills for management success.* Englewood Cliffs, N.J.: Prentice-Hall.

Willms, J.D. (2003a). *Ten Hypotheses about socioeconomic Gradients and Community differences in Children's Developmental Outcomes*. Report prepared for Human Resources Development Canada.

Willms, J.D. (2003b). *PISA 2000: Socioeconomic Status and Reading Performance of French- and Italian-speaking Swiss Students.* Report commissioned by the Service de la Recherche en Éducation. Genève, Switzerland.

Willms, J.D. (2009). *Raising and leveling the learning bar in Ontario schools.* Keynote presentation at the annual Ontario Education Research Symposium, Toronto, ON.