School-Effectiveness in Mathematics in Sweden compared with countries in Europe and Asia-Pacific

Marie Wiberg, Department of Statistics, Umeå University, Sweden, marie.wiberg@stat.umu.se
Ewa Andersson, Department of Applied Educational Sciences, Umeå University, Sweden, ewa.andersson@edmeas.umu.se

Abstract

There is an increased focus on educational quality and learning outcomes in Sweden, especially since international comparative studies like TIMSS has put the limelight on declining study results in core subject like mathematics. School-effectiveness is therefore an urgent issue. The aim of this study is to identify factors that contribute to the explanation why some schools is effective but others are less effective in terms of the students' academic achievement in mathematics on TIMSS. This is conducted by examining Sweden in contrast with school-effectiveness in some high achieving countries in Europe and Asia-Pacific. The framework of multilevel analysis was used, since it is important to separate the effect of school-level variables from the effect of home environment and to take care of the sampling design used in TIMSS 2007. The results show that different educational system identifies different school level factors when controlling for home background. Noteworthy is that very few school level factors were significant in the full models. The single most important factor for school-effectiveness appears to be student behavior in school. Offering enrichment or remedial in mathematic seem to have different effects in different countries. Professional development opportunities for the teachers and the use of different incentives do not seem to have any effect in any of the countries, except in Chinese-Taipei. Disappointingly, we were unable to detect any significant school level factors in Sweden. A possible reason for this is that we excluded variables that have been altered in any of the countries, since we do believe that such changes might be a threat to the validity. Hence, it is possible that important questions in the Swedish context were excluded.
Keywords: *multilevel analysis, home background, principals questionnaire, School-Effectiveness, country comparisons*

Introduction

International comparative studies (e.g. TIMSS) have put the limelight on declining study results among students in Sweden in core subject like mathematics. In Sweden, this has brought about a debate about learning outcomes and quality in education, although the concept of school-effectiveness, much like the concept of accountability, is not often used in the public debate (cf. Eklöf, Andersson, & Wikström, 2009). However, school-effectiveness is an urgent issue since it is important that the schools and their students achieve their core objectives. Furthermore, changes in the Swedish society in terms of a shift towards a more multi-cultural society along with other long and short term trends makes it urgent to investigate and compare school-effectiveness in countries quite different from the Swedish school context (cf. Skolverket, 2009).

A vast amount of studies have been conducted regarding school-effectiveness from different perspectives (see e.g. Kyriakides & Charalambous, 2005; Luyten, Visscher, & Witziers, 2005, Martin, Mullis, Gregory, Hoyle, & Shen, 2000). Many studies use students' achievement in core subjects (e.g. mathematics, language and science) as criteria for school-effectiveness (Neuschmidt, Hencke, Rutkowski, & Rutkowski, 2008). This is also the point of departure for this study. Furthermore, it is a well known fact that in general there is a correlation between students' socio-economical background and school achievement (see e.g. Mullis, Martin, & Foy, 2008; Skolverket, 2009). It is therefore important to separate the effect of school variables from the effect of students' home environment, in line with what previous research on effective schools using TIMSS data has suggested (Martin, Mullis, Gregory, Hoyle, & Shen, 2000, Neuschmidt, Hencke, Rutkowski, & Rutkowski, 2008). Consequently a school could be viewed as "…effective to the extent that it "adds value" by realizing the potential of the student body through efficient organization and effective instruction" (Martin, Mullis, Gregory, Hoyle, Shen, 2000, p. 9).

Variables related to school effectiveness can be categorized into context variables and climate variables (Ma, Ma & Bradley, 2008). Context variables describe the “hardware” of the school, e.g. school location and resources, characteristics of the student and the teacher body. Climate variables, often referred to as evaluative variables, describe the “software” of the school, with characteristics descriptive of the learning environment, e.g. administrative polices, values, and expectations of students, parents, and teachers. Nevertheless, it is important to consider that such variables may be perceived in different ways in different countries and that there are differences between the school systems and the way mathematics is emphasized (see Mullis, Martin, Olson et al., 2008). Hence, a pressing issue is, not only to find explanations to declining results, but rather to identify key factors related to school-effectiveness that can be influenced. This study deals with school effectiveness in terms of factors related to the climate and the context of high achieving schools in TIMSS 2007 in mathematics. The main purpose of this study is to identify factors that contribute to the explanation why some schools are effective in terms of their students’ achievement in mathematics, by contrasting Sweden to high achieving countries in Europe and Asia-Pacific.

Methodology

Data/participants: Data from TIMSS 2007 8th grade in Mathematics (IEA, 2007) was used from Sweden, three countries from Asia-Pacific (Chinese-Taipei, Republic of Korea and Japan) and four from Europe (The Czech Republic, Hungary, Russian federation and Slovenia). The choice of the countries was based on two criteria. First of all, the average mathematics achievement scale score for each country were above the average score in TIMSS 2007. Secondly, the country had followed the sample design stated in TIMSS (Mullis, Martin, & Foy, 2008). Students’ mathematical achievement, students’ questionnaire and the school questionnaire were used. Furthermore, data from the qualitative descriptions of each country was used for comparison between the countries regarding aspects of the different school systems and the way they emphasize mathematics (see Mullis, Martin, Olson et al., 2008). The result of the review has been related to the results from the statistical analysis.

Statistical analysis: In order to identify school indicators associated with school effectiveness a multilevel analysis approach was chosen (see e.g. Gelman & Hill, 2007; Snijders & Bosker, 1999). This approach was chosen since we wanted to take account of the sampling procedure used in TIMSS (Kyriakides & Charalambous, 2005) and hence be able to control for factors which are not connected to school effectiveness, e.g. factors connected to home background of the students. Multilevel analysis has previously been used by a few researchers to analyze TIMSS (see e.g. Martin, Mullis, Gregory, Hoyle, & Shen, 2000; Webster & Fisher, 2000). The statistical analysis was carried out in six steps.

In all analysis the IEA IDB analyzer was used to prepare the files and analyze them together with PASW 18.0. The first step was to reduce the number of variables from the students’ questionnaires, i.e. the students’ home backgrounds. These variables were examined with multiple regressions with mathematics achievement scores as dependent variables. We used all five plausible values in this study. Examples of home background variables supported by the literature and available in TIMSS studies are e.g. home possessions (i.e. calculator, computer, study desk, books, and dictionary), parents’ education, and students’ sex. All home background variables which were measured according to TIMSS standard was examined at this step.

One variable used were parents’ highest education, a variable which has a large amount of missing values (about 19% over all examined education systems). Also the variable amount of homework were just below 10% of missing values while in the other variables there were less than 2% missing values. In order to be able to use as many students as possible missing values were replaced using multiple imputations available in PASW, a method recommended by Schaefer & Graham (2002). This procedure assumes that the data is missing at random (MAR) but according to Collins, Schaefer & Kam (2001) to incorrectly assume MAR does only have a minor impact on estimates and standard errors. Further, since we are using large samples and imputed data has better coverage and less bias than to only analyzing complete cases we consider our result fairly robust. As imputation model we used all variables that we wanted to include as predictors and response variables in our latter models so that analyzed model and imputed model was the same. Then, we ran the multiple regression and significant variables were kept in the multiple regression models. The home background variables were recoded to make more sense in the analysis, e.g. home possession was recoded to zero if no possession and 1 if possession.

The second step was to examine which schools were expected to be effective. This was done by assuming that a school is more effective if its mean mathematics achievement is higher than predicted from the multiple regressions of our identified home background variables. Hence, the mean mathematics difference between the five mathematics plausible values and the expected scores were calculated in each country. Schools were regarded as effective if they were in the top third in their country in mathematics achievement and less effective if they were in the bottom third in their country in mathematics achievement.

In the third step we reduced the number of home background variables by using principal component analysis (PCA) with varimax rotation. The material was weighted with TOTWGT since we had eight countries in the analysis. After examining different combinations we decided to retain three home background factors which explained 68.42% of the variance. We also kept the variables time spent on math homework (BS4MSHWM), labeled HW and if father was born in country (BS4GFBRN), labeled FB, for further analysis. HW and FB were not included in the factors since they behaved differently than the other variables in the factors. The three factors were constituted as follows; H1 contained home possession of computer (BS4HT02) and internet connection (BS4HT05), H2: contained books at home (BS4GBOOK) and parents' highest education (BSDGEDUP). Finally, H3 contained home possession of dictionary (BS4HT04) and study desk (BS4HT03).

The fourth step was to identify school variables which could have an impact on the school effectiveness. In order to weight the countries equally a manual calculated senate weight was introduced which gave all schools the same weight within a country. The school variables were examined with correlation to the mean mathematics difference obtained in the second step. As for the home background variables relevant literature suggestions and available variables in TIMSS determined examined variables. Note, we excluded all variables where any of these eight countries had made country specific adjustments, e.g. alteration of the options.

In the fifth step we reduced the number of school variables by again using PCA with varimax rotation. The material was weighted with the manual calculated senate weights. Since some principals had not answered a number of the items of interest we decided to omit those schools. It would probably be better to impute missing data here but since the amount of missing values was relatively low in most countries we decided to not impute values. After the final analysis we decided to retain six factors which explained 60.7 % of the variance. The factors scales reliability was examined with Cronbach's alpha and factors with many variables had in general an alpha above .7 but the factors with only two variables were around .4. The six school level factors, which include context as well as climate variables, were:

[S1] *Student behaviour* (BC4GFP5-6, 8-12):This category include variables from question 18A, which focusing on different kind of negative student behaviour, e.g. arriving late at school, vandalism, theft etc.

[S2] *Teachers' professional development opportunities* (BC4GPDIK, BC4GPDTS, BC4GPDIC, BC4GPDUT, BC4GPDSG). This category include question 13 which contain the teachers professional development opportunities, focusing on the participation rate.

[S3] *Town/school size* (BC4GTENR, BC4GEENR, BC4GCOMU): This category include question 1-2.

[S4] *Teachers´ incentives* (BC4MEPTR, BC4MEPOS, BC4MRRTM): This category include variables from question 14 and 17 which include strategies for evaluation of teacher practice and strategies to recruit or retain teachers.

[S5] *Students opportunities to learn* (BC4MSORM, BC4MSOEM): This category include question 10 focusing on strategies for promoting students' learning, i.e. offering enrichment or remedial in mathematics.

[S6] *Student body (*BC4GSBEA, BC4GSBED). This category include question 3, i.e. percentage of students from economically disadvantaged or economically affluent homes.

In the final sixth step, multilevel analysis was carried out with HLM 6 (Raudenbush, Bryk, Cheong, Congdon & du Toit, 2004) which is built on the theory described in (Raudenbush & Bryk, 2002) and suggestions on how to estimate effective schools from Ma, Ma & Bradley (2008). By using multilevel analysis we can control for home background and concentrate our study on school level variables of interest for school effectiveness. For each of the eight countries three kinds of models were fit: 1) null model 2) home/context model 3) full model. We used the five mathematics achievement plausible values as dependent variables and grand mean centering of the variables.

Null model: In this model we quantified the between-school variance of the average mathematics score with no predictors. Mathematics achievement for each student was estimated as a function of the school average plus a random error.

Level 1 (within schools): $Y_{ij} = \beta_{0j} + r_{ij}$,

Level 2 (between schools): $\beta_{0j} = \gamma_{00} + u_{0j}$,

where Y_{ij} represents mathematics achievement for each student $i = 1,2...n_j$ in school $j = 1,2,...J$, β_{0j} represents the mean mathematics achievement of school j and r_{ij} represents the random error of student i in school j. Further, γ_{00} represents the grand mathematics mean for all schools and μ_{0j} represents the random school effect, the deviation of school j :s mean from the grand mean.

Home/context model: In this model we quantified the between-school variance that is due to home background measures and the differences in the average home background. Mathematics achievement for each student was estimated as a function of the school average plus random error. The aggregated mean of the home background measures was entered at the second level.
Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_1(H1) + \beta_2(H2) + \beta_3(H3) + \beta_4(HW) + \beta_5(FB) + r_{ij}$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(H1m) + \gamma_{02}(H2m) + \gamma_{03}(H3m) + \gamma_{04}(HWm) + \gamma_{05}(FBm) + u_{0j}$$

$$\beta_{1j} = \gamma_{10},\ \beta_{2j} = \gamma_{20} \cdots\ \beta_{5j} = \gamma_{50},$$

where $H1$, $H2$, $H3$, FB and HW are connected to students' home background as defined earlier. Further, $H1m$, $H2m$, $H3m$, HWm, and FBm are the contextual effect of the home background factors and variables.

Full model: In this model we quantified the association of school factors with student mathematics achievement, while controlling for students' home backgrounds. The mathematics school mean achievement varies randomly around a grand mean for all schools. The first level was weighted with house weights and the second level and the second level was weighted with school weights. On the first level we examined all possible home background measures for each educational system. On the second level we examined all possible school level factors as well as the mean of the home background measures. Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_1(H1) + \beta_2(H2) + \beta_3(H3) + \beta_4(HW) + \beta_5(FB) + r_{ij}.$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(H1)m + ... + \gamma_{05}(FBm) + \gamma_{06}(S1) + ... + \gamma_{011}(S6) + u_{0j}\ \beta_{0j} = \gamma_{00} + \mu_{0j},\quad \beta_{1j} = \gamma_{10},$$

$$\beta_{2j} = \gamma_{20} \cdots\ \beta_{11j} = \gamma_{11j},$$

where $S1,...,S6$ are the school level factors.

Finding and Discussion

Effective schools were evaluated according to average mathematics achievement, see Table 1. The average mathematics achievement within Asia-Pacific varied in the interval of 570-598 and schools within Europe, Sweden excluded, varied in the interval of 501-517. Noteworthy is that less effective schools in Asia-Pacific still has an higher average achievement than more effective schools in some of the countries in Europe.

[Take in Table 1 about here]

In Table 2 the intraclass correlation (ICC) attributable to schools and the proportion of variances explained by the HLM models are summarised together with the model which showed the best fit. In general, the ICC was very low for all countries and varied between 0.03 (rep. of Korea) and 0.28 (Russian federation). In general, a low ICC is an indication that there is not that much added value of using HLM compared to regular linear regression. The proportion of between-school variance ranged from 0.39 to 0.77. In general, high proportion of between-school variance is an indicator of the existence of school effects (Ma, Ma & Bradley, 2008). Note, in neither Korea, nor Hungary nor Sweden was a full model ever significant.

[Take in Table 2 about here]

In Table 3 the result from the full or in some cases the home/context models are summarised. When using the full HLM models all three home background factors were significant in all eight countries and the HW was significant in all countries except in Chinese-Taipei. The FB variable was only significant for Slovenia. Noteworthy, in contrast to some previous studies in the field (se e.g. Neuschmidt, Barth, Rutkowski, & Rutkowski, 2009), our analysis does not imply any gender differences although that was a variable we initially examined. The H2m factor was significant in all countries except Slovenia.

[Take in Table 3 about her]

The school level factors showed a more blended pattern and no clear regional pattern was found. Overall, very few school level factors were significant in the full models, a result in line with other studies (see e.g. Neuschmidt, Barth, Rutkowski, & Rutkowski, 2009). Most east European countries had the first school factor significant, except Hungary, while in Asia-Pacific this school factor was only significant for Japan. Some aspects related to effective schools seem to be more prevalent, especially in some of the countries. Slovenia is the country were three out of six school factors seem to be related to school effectiveness. Sweden, on the contrary, is the only country where none of the factors seem to be of importance for promoting school effectiveness. In general, when we had controlled for home background, school factors accounted only for a minor percentage of additional variance in most countries. The pattern of the school factors are summarised below;

Town/school size [S3] seems to be of importance for promoting success, but only in two of the countries, i.e. Chinese-Taipei and Slovenia. This may be related to their almost identical emphasis on approaches and processes in the indented curriculum in mathematics (see Mullis, Martin, Olson et al., 2008).

Student body [S6] seems to be a factor of importance for school efficiency, but only in Slovenia.

Student behaviour [S1]: In line with previous studies, this is the single most important factor for school effectiveness, or rather lack of success. This is a factor that mainly seems to be present in the east European countries. The reason for this may be related to socio-cultural differences and more overarching changes in the society and its structure.

Students' opportunities to learn [S5] seem to have different effects in different countries. In the Czech Republic this is a strategy that seem to be promote success among the students while that strategy in Japan seem to have the opposite effect. The reason for this might be related to the fact that the emphasis on mathematics and the approaches and processes in the intended mathematics curriculum differ in some aspect between the two countries (see Mullis et al., 2008). Compared to Japan, which already seem to have a stronger focus on advanced mathematics, the students in the Czech republic may benefit more from programs aiming at enhance students progress.

Teachers' professional development opportunities [S2] do not seem to be relevant for school effectiveness in any country.

Teachers' incentives [S4] does not seem to have any effect in any of the countries, except in Chinese-Taipei where incentives have a negative effect on school-effectiveness.

Conclusion and Implications

We aimed to identify factors in different educational systems which showed if a school is effective. Moreover we wanted to examine if it was possible to learn from high achieving countries what could be influenced in the Swedish educational system. Our overall conclusion is that it appears that when we control for students' home background, the school level factors are of less importance, especially in Sweden where we were unable to detect any significant school level factors. A reason for this might be that we excluded variables due to alterations in the questionnaires since we do believe that such changes might be a threat to the validity, i.e. alterations might mean that the principals do not answer identical questions. Nevertheless, some of these excluded altered questions might be important for school effectiveness in the Swedish context. However, the fact that there are differences in the way the subject is emphasized in different countries may also be a factor that is influential on school effectiveness.

In this study we have only used the school questionnaire as measures of the context and climate of the school, i.e. the principals' view on the school and how it works. Furthermore, we have excluded variables that might have been relevant not just for Sweden but also for the other countries. In order to get a more comprehensive view of the schools and their efficiency, views from teachers, students and, if possible, their parents should be included and weighted together since they constitute important interested parties. The fact that there are still quite large amount of unexplained variance on school level, imply that one should construct country specific school factors and fit more multilevel models to each educational system. Finally, if we want to find out what we should influence in Sweden maybe we should study Sweden over time or the closest educational systems, e.g. the Nordic countries.

Our choice to exclude all variables where any of the countries have made alternations may have been too limiting. In further studies a less strict strategy for exclusion may be applied. However, the fact that national adaptations are made to the international version of the questionnaire and what it means for interpretation of results in the TIMSS context, is a validity issue that deserves special attention in further research.

References

Collins, L. M., Schaefer, J. L., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6 (4)*, 317-329.

Eklöf, H., Andersson, E. & Wikström, C. (2009). The concept of accountability in Education: Does the Swedish school system apply? *Cadmo 2*, 55-66

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University press.

IEA. (2007). TIMSS 2007 international database. Obtained from the official TIMSS 2007 International Database CD.

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

Luyten, H., Visscher, A., & Witziers, B. (2005). School effectiveness research: From a review of the criticism to recommendations for further development. *School Effectiveness and School Improvement*, *16* (3), 249-279.

Ma, X., Ma, L., & Bradley, K. (2008). Using multilevel modelling to investigate school effects. In A. A. O´Connell & D. B. McCoach (Eds.) *Multilevel modelling of educational data*, Chapter 3, pp. 59-110. Information Age Publishing: Charlotte, NC.

Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., & Shen, C. (2000). Effective schools in science and mathematics. IEA's third international mathematics and science study, IEA: Chestnut Hill, MA.

Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 International Mathematical Report. Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eight Grades.* Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education.

Mullis, I. V. S., Martin, M. O., Olson, J. F., Berger, D. R., Milne, D., Stanco, G. M. (Ed.). (2008).*TIMSS 2007 Encyclopedia. A Guide to Mathematics and Science Education Around the world* (Vol.1-2*).*Boston College: TIMSS & PIRLS International Study Center, Lynch

School of Education.

Neuschmidt, O., Barth, J., Rutkowski, L., & Rutkowski, D. (2009). Effective schools in Arab educational systems: an analysis of TIMSS 2007. IERI Spring Academy Hamburg, Germany, May 2009.

Neuschmidt, O., Hencke, J., Rutkowski, L., & Rutkowski, D. (2008). Effective schools in Arab educational systems: an analysis of TIMSS 2003*, 3rd IEA International Research Conference*. Taipei.

Raudenbush, S. W., & Bryk, A. S., (2002). *Hierarchical linear models: Applications and data analysis methods. 2nd edition,* Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W. Bryk, A., Cheong, Y. F., Congdon, R., & du Toit, M. (2004). Hierarchical linear and nonlinear modelling. SSI scientific software.

Schaefer, J., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, *7 (2)*, 147-177.

Skolverket (2009). *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer. Stockholm*: Skolverket

Snijders, T. Bosker, R. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modelling*. ORT: Sage Publishers.

Webster, B. J. & Fisher, D. L. (2000). Accounting for variation in science and mathematics achievement: a multilevel analysis of Australian data third international mathematics and science study (TIMSS). *School Effectiveness and School Improvement, 11, 3,* 339-360.

Table 1: Average mathematics achievement within countries and in less and more effective schools together with the standard errors in parenthesis.

Countries	Average	Less effective schools	More effective schools
Chinese Taipei	598 (4.5)	546 (4.0)	645 (5.7)
Rep. of Korea	597 (2.7)	568 (3.3)	629 (3.7)
Japan	570 (2.4)	537 (3.0)	605 (3.7)
The Czech Rep.	504 (2.4)	471 (2.1)	552 (4.3)
Hungary	517 (3.5)	476 (4.6)	562 (5.8)
Russian Fed.	512 (4.1)	460 (2.7)	566 (4.5)
Slovenia	501 (2.1)	475 (2.7)	525 (2.1)
Sweden	491 (2.3)	465 (2.6)	517 (2.4)

Table 2: The ICC attributable to schools and the proportion of between school variance explained by the HLM models and the HLM models that showed the best fit.

Country	ICC	Proportion of between school variance	Model
Chinese Taipei	0.07	0.75	Full model
Rep. of Korea	0.33	0.71	Home/context model
Japan	0.11	0.58	Full model
The Czech Rep.	0.11	0.77	Full model
Hungary	0.14	0.68	Home/context model
Russian Fed.	0.29	0.39	Full model
Slovenia	0.06	0.46	Full model
Sweden	0.07	0.50	Home/context model

Table 3: Size of the significant factors in the HLM models when using the full models or the home/context models.

	Level 1					Level 2						
Countries	H1	H2	H3	FB	HW	H2m	S1	S2	S3	S4	S5	S6
Chinese Taipei	21.4	27.6	16.1			29.4			8.4	-7.7		
Rep. of Korea	46.5	33.9	20.5		-10.5	25.4						
Japan	12.8	22.2	12.8		-13.5	37.4	-6.9				-9.0	
The Czech Rep.	12.0	18.2	9.7		-16.6	71.4	-4.9				4.9	
Hungary	14.7	27.1	10.2		-7.8	35.5						
Russian Fed.	6.7	15.7	9.3		-11.5	45.9	-17.6					
Slovenia	18.4	24.1	4.2	23.0	-6.5		-3.8		16.7			6.0
Sweden	15.5	19.9	6.4	19.8	-4.9	27.7						

H1= home possessions (computer and internet connection), H2 = books at home and parents' highest education, H3 = home possession (dictionary and study desk), FB = father born in country, and HW = time spent on math homework. S1 = student behaviour, S2 = teachers' professional development opportunities, S3 = town/school size, S4 = teachers' incentives, S5 = students' opportunities to learn, and S6 = student body.