School effectiveness in the Nordic countries in relation to PISA and TIMSS

Ewa Rolfsman[1] (ewa.rolfsman@edusci.umu.se), Marie Wiberg[2] (marie.wiberg@stat.umu.se), & Inga Laukaityte[2] (inga.laukaityte@stat.umu.se)
[1]Department of Applied Educational Science, Umeå University, Sweden.
[2]Department of Statistics, USBE, Umeå University, Sweden.

Abstract

In a Nordic perspective, the finish students overall achieve the highest score on PISA (Programme for International Student Assessment), while the Swedish students exhibit declining results. The results of the Swedish students have drawn attention to the quality of education and the role of the educational professionals and the efficiency of the school. It is therefore of vital importance to investigate whether these results can be related to school level factors in a Nordic perspective. However, TIMSS (Trends in International Mathematics and Science Study) and PISA exhibit similarities as well as differences as they target different subjects. In addition, the results on TIMSS and PISA differ between countries. The aim of this study is to investigate whether school level factors can contribute to the explanation of the results for the Nordic countries participating in PISA 2009 and, if so, identify factors that can be influenced in order to enhance students' achievement. We focus on school effectiveness in relation to PISA, since all Nordic countries participate in PISA. However, the results are contrasted to results from TIMSS for Sweden and Norway. In order to separate the effect of school level variables from the effect of student's home environment and to take care of the sampling design used in TIMSS and PISA, multilevel analysis was used. The results show that only a few school level factors were significant, and only in Sweden and Finland. Furthermore, school level factors in Sweden and Norway on PISA differ from school level factors based on TIMSS data.

Keywords: PISA, TIMSS, School effectiveness, School questionnaire, Principal

Introduction

A common view of Swedish schools nowadays is that the knowledge level is dropping. A large group of students do not reach the objectives of the Swedish comprehensive school, which in turn means that they are not competent to proceed to secondary education. This group has also increased over time (Swedish National Agency for Education, 2012a).

To find explanations for low or declining performance of students at different levels of education is a field of research that has attracted great interest and it is nowadays well established that there is a clear relationship between social background and success in studies (see, e.g. Swedish National Agency for Education, 2009). OECD (2013) has established, there is a link between socioeconomic disadvantage and poor performance in school, but emphasize that education can make a difference based on observations that there are countries that exhibit progress in reversing this trend "in narrowing the gap between advantaged and disadvantaged students while simultaneously improving overall performance" (OECD, 2013, p. 2). However, Sweden is not among these countries. In fact, Sweden is one of the countries that exhibit a decline over the years, both in performance and in equity level. An urgent issue

is therefore to promote student success and thereby also promote equity in education.

The impression of the Swedish school and the outcomes of the students, are not univocal. High achieving students as well as low achieving students exhibit a decline over the years in achievement, but this range between subjects. Results from PISA (Programme for International Student Assessment) (Swedish National Agency for Education, 2012b) show that almost one out of six students was high-achieving in one of the subjects, but not in all subjects. Nevertheless, overall results from the last PISA study in 2009 reveal that Sweden exhibits the largest decline in reading over the years among the Nordic countries. Norway as well as Iceland, exhibit falling results until 2006, but increase their results in the latest PISA study. Denmark has been fairly stable over the years. However, Sweden and Finland was the only Nordic countries that exhibit a decline in 2009 compared to in 2006. Still, the Finish students achieve the highest results among the Nordic countries (Thornval, 2012). In addition, the school segregation, previously considered as keen to counter in Sweden, rather has become increasingly prominent in recent years (Swedish National Agency for Education, 2012a; OECD, 2013). This is however, also a factor that is highlighted as a factor associated to the Swedish decline in PISA (Egelund, 2012). An urgent issue is therefore to investigate the effect of schooling, regardless of student characteristics.

A field of research, focusing on identifying factors that enhance success in their students, is school effectiveness research. A basic assumption within this field of research is that a school is successful – or effective – if it "adds value" to the students in terms of different forms of outcomes, regardless of the characteristics of the students within a school (Martin, Mullis, Gregory, Hoyle, Shen, 2000). Extensive research has been conducted regarding school effectiveness from different perspectives over the years (se e.g. Cremers, Kyriakides & Sammons 2010). One of the first school effectiveness studies ever conducted was the nowadays classical study by Rutter et al. (1979) showing that the climate in which the learning take place is essential for student outcomes.

Generally, a school climate could be described as deep founded values that come into daily practice by the educational professionals. In this context the principal is a key actor since he or she is ultimately responsible for the school, including student achievement. The significance of the leadership, and thereby factors at the school level, has been the issue in several studies. However, research on school effectiveness using large-scale data, focusing on the principal, has been sparse in a Nordic context (Johansson & Bredeson, 2011). This study will make a contribution to this field of research, since we make use of large-scale data from PISA and TIMSS (Trends in International Mathematics and Science Study) focusing on the school level from the perspective of the principals.

Our previous results from studies based on student achievement in mathematics, shows that only a small number of school factors appear to be positively related to the student performance (Wiberg & Andersson, 2010; Rolfsman & Wiberg, 2012; Wiberg & Rolfsman, 2013). However, these studies are limited to data from TIMSS and student performance in science and mathematics. In order to achieve more knowledge about school level factors of importance for student achievement, research should include data from other large-scale studies and include achievement data from other subjects, in particular since there is evidence that student performance differ between TIMSS and PISA (Wu, 2010).

School effectiveness can be defined in different ways, but many studies use students'

achievement in core subjects as criteria for school effectiveness (Neuschmitdt, Henke, Rutkowski, & Rutkowski, 2003). This is also the case in this study, while controlling for the effect of the student's home environment. In addition, in line with Martin, Mullis, Gregory, Hoyle, Shen, (2000) we view a school as effective if 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). This is also an incentive for us to choose to focus on the school level. Given that it is well documented that there is a relationship between social background and school performance (see e.g. Swedish National Agency for Education, 2009), it is vital to separate the effect of the schools' influence and that resulting from the students social background. This is something that requires the use of multi-level analysis, in line with what previous research on school effectiveness has demonstrated.

The aim of this study is to investigate whether school level factors can contribute to the explanation of the results for the Nordic countries participating in PISA 2009 and, if so, identify factors that can be influenced in order to enhance students' achievement. Furthermore, the results will be contrasted to results from TIMSS, regarding school efficiency for Sweden and Norway, since they are the only Nordic countries participating over the years in both PISA and TIMSS with students who are 15-years old and grade 8 respectively.

Methodology

Data/participants: Data from PISA 2009 (OECD, 2009) was used from the Nordic countries, i.e. Finland, Denmark, Norway, Sweden and Iceland. PISA is given every third year to a randomly selected group of 15 year olds within the subjects mathematics, reading, and science. At each assessment one subject is given special focus. The examined achievement was reading comprehension because it was the subject which was emphasized in PISA 2009. The results are contrasted to mathematics achievement and science achievement among students from Sweden and Norway from 8th grade on TIMSS 2003, 2007 and 2011.

Statistical analysis: Each of the five countries was examined for reading achievement with PISA 2009 data. Multilevel analysis (see e.g. Gelman & Hill, 2007) was used based on the fact that we wanted to take into account of the sampling procedure used in PISA (Kyriakides & Charalambous, 2005) and hence be able to control for factors which are not connected to school effectiveness, e.g. factors connected to the students' home background (Ma, Ma, & Bradley, 2008). We constructed student level home background factors and school level factors from the available variables in PISA which were as similar as possible to previous studies (Rolfsman & Wiberg, 2012; Wiberg, Rolfsman & Laukaityte, 2013) in order to facilitate comparisons and based on what previous research indicate that might be of interest. Regarding school level variables, our point of departure was the existing question areas in the questionnaire. Regarding student level variables; we used sex [SEX] and whether the father was born within the examined country or not [FB] together with two constructed factors [READ] and [SES] defined as follows:

[READ] = Reading activities includes questions focusing on students' attitudes towards reading and reading habits..

[SES] = Socioeconomic status is measured by number of books in the students' homes

and home possessions (study desk, student's own rom, computer, and a dictionary).

In Table 1, mean and standard deviations are given for the student level variables and student level factors for the five countries. Note that average reading activities, socioeconomic status and sex was similar in all five countries, but the proportion of native fathers was much lower in Denmark as compared with the other countries.

Table 1. Student level factors READ and SES with their, mean and standard deviation and the proportion of the student level variables FB (% of fathers who are natives in the Nordic countries) and SEX (% of men).

Countries	READ	SES	FB	SEX
Denmark	2.21 (0.02)	2.20 (0.01)	70.8	51.3
Finland	2.31 (0.01)	2.25 (0.01)	93.8	50.8
Iceland	2.23 (0.01)	2.31 (0.01)	92.4	50.9
Norway	2.13 (0.02)	2.29 (0.01)	88.0	49.0
Sweden	2.22 (0.01)	2.32 (0.01)	81.6	49.4

READ = Reading activities, SES = socioeconomic status, FB = Father born within country, SEX = Student's sex.

In order to investigate school effectiveness, we examined if the mean reading achievement is higher than predicted from multiple regression of the student level variables and student level factors within each country. More specifically, we examined the mean difference between the five reading plausible values and the expected scores from the multiple regressions of the student level factors and variables in each country. In general, we regarded school as *more effective* if the schools were in the top third in its country in achievement in reading achievement, *mid effective* if they were in the middle third, and *less effective* if they were in the bottom third in reading achievement.

Next we aimed to examine the five countries with multilevel analysis, where we can control for the student level factors and pay special attention to the school level factors. The construction of school level factors was performed as follows. Based on the questions included in the 2009 PISA school questionnaires, answered by the school principal, our point of departure were the six question areas. Items which were problematic to recode into a meaningful factor were excluded. We are aware that this might have a small impact on the result but overall, quite few variables were excluded. The constructed school level factors followed the question areas in the school questionnaire, except for the two question areas regarding school curriculum and assessment and the school's policy and practice respectively. These two question areas were framed as two factors each, named D1, D2 and F1, F2. The constructed school level factors for this study were the following:

[A] **School location** Based on question 4 and 5, focusing on the school location and the prevalence of competing schools in the surroundings of the school.

[B] **Student and teacher body:** Based on question 8, focusing on the percentage of students in the school having a first language that is not the test language.

[C] **Lack of school's resources:** Includes items where the principal is asked about into what extent the school's capacity to provide instruction is hindered by a shortage or a lack of resources (question 11), such as qualified teachers, computers for instruction, library materials.

[D1] **School curriculum:** Includes items from questions regarding what the school can offer their students e.g. art club, art activities, debating club or debating activities, collaboration with local newspapers (question 13) but also what the school can offer their students whose first language is not the test language, e.g. additional instruction, class size reduced (question 14).

[D2] **Student assessment**: Includes items from questions regarding evaluation and assessment practice and test use, in terms of frequency of use of different test formats, e.g. standardized tests, teacher-developed tests (question 15) and use of tests for different purposes, e.g. to group students for instructional purposes or to inform parents about their child's progress (question 16).

[E] **School climate:** Includes items focusing on factors in the learning environment that constitute obstacles for student learning e.g. poor student-teacher relations, disruption of classes by students, staff resisting change (question 17) but also items focusing on parents expectation towards the school (question 18).

[F1] **Accountability and evaluation:** Includes items focusing on the schools policy and practice (strategies) for using achievement data for accountability purposes, e.g. achievement data are posted publicly or used in decisions about instructional resource allocation to the school (question 22), and for evaluation of the teachers, e.g. principal or senior staff observations of lessons (question 23).

[F2] **Administration and management:** Includes items focusing on the schools policy and practice (strategies) regarding responsibility and management issues e.g. the principal responsibility for selecting teachers for hire, formulating the school budget, establishing student assessment policies, inform teachers about possibilities for updating their knowledge and skills (question 24, 26).

[G] **Principal's sex:** Recoded as 0 for male and 1 for female.

Descriptive statistics for the school level factors for the five Nordic countries are given in Table 2. There are noticeable differences in the mean between these countries. One particular interesting difference is that Finland differs from the other countries in having a very low mean of school level factor F1, i.e. meaning that assessments are not used regularly to the same extent as in the other countries for accountability and evaluation purposes. Finland also exhibits the lowest mean with reference to students having a first language that is not the test language (factor B).

Table 2. School level factors' descriptions with mean and standard deviation within parenthesis for the first eight factors and proportion of males for principals' sex.

	Denmark	Finland	Iceland	Norway	Sweden
A	2.20 (0.05)	2.10 (0.05)	2.03 (0.00)	1.76 (0.05)	2.17 (0.05)
B	0.94 (0.04)	0.72 (0.03)	0.86 (0.00)	1.02 (0.03)	1.15 (0.04)
C	1.80 (0.05)	1.98 (0.05)	1.54 (0.00)	2.39 (0.05)	1.86 (0.05)
D1	1.45 (0.04)	2.08 (0.05)	2.06 (0.00)	1.79 (0.04)	1.93 (0.04)
D2	2.12 (0.04)	2.22 (0.04)	2.46 (0.00)	2.27 (0.03)	2.43 (0.04)
E	2.32 (0.05)	2.77 (0.04)	2.43 (0.00)	2.72 (0.04)	2.57 (0.04)
F1	1.72 (0.06)	0.58 (0.05)	1.12 (0.00)	1.67 (0.05)	1.68 (0.06)
F2	2.25 (0.04)	2.03 (0.05)	2.39 (0.00)	1.99 (0.04)	2.47 (0.04)
G	33.87	43.94	44.36	37.19	49.45

There were missing data both at the student level and at the school level. The missing data range was in general low in the student home background variables and factors ranging from 1.05% (for FB in Norway) to 8.61% (for READ in Denmark). Variable SEX had no missing values for any of the studied countries. We used listwise deletion to exclude missing data (Tabachnick & Fidell, 2007). We are aware that imputing missing data would probably have been better but this choice was made due to the small amount of cases which were removed. For handling the missing data at school level, the full-information maximum likelihood procedure was used (Schafer & Graham, 2002).

We used the software MPLUS 7 (Muthén & Muthén, 2011) to conduct multilevel analysis. For each country we examined the null, context and full model. All five plausible values for reading were used as dependent variables and all variables were grand mean centered. First we ran the null model and calculated the intraclass correlation (ICC) to test how homogeneous the data are within group level clusters. In the null model reading achievement for each student was estimated as a function of the school average with a random error.

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

Level 2 (between schools): $\beta_{0j} = \gamma_{00} + u_{0j},\ j = 1,\ldots,J,$

where Y_{ij} denotes reading achievement for student *i* within school *j*, β_{0j} is the average reading achievement for school *j*, and r_{ij} is the error term representing a unique effect associated with student *i* in school *j*. Level 2 term γ_{00} denotes the grand mean of reading achievement and u_{0j} is the error term representing a unique effect associated with school *j*.

Second, we examined the home context models for reading achievement. In the home context models the reading achievement for each student was estimated as a function of the school mean achievement and the student effect. The students' home background factors were at the first level of the model, $(H_1,\ldots,H_f)$ and aggregated student level factors $(aH_1,\ldots,aH_f)$ were placed at the second level. Student level factors were weighted by total student weights (W_FSTUWT).

Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(H_1) + \beta_{2j}(H_2) + \ldots + \beta_{fj}(H_f) + r_{ij},\ i = 1,\ldots,N,$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(aH_1) + \gamma_{02}(aH_2) + ... + \gamma_{0f}(aH_f) + u_{0j}, \quad j = 1, \ldots, J$$
$$\beta_{1j} = \gamma_{10}, \beta_{2j} = \gamma_{20}, \ldots, \beta_{fj} = \gamma_{f0}$$

,

Third, the full models were examined in which the association of school factors with student reading achievement was quantified while controlling for the student home background. The full models thus contained student level factors, aggregated student level factors and school level factors ($S_1, ..., S_7$) weighted with school weights (W_FSCHWT).

Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(H_1) + \beta_{2j}(H_2) + ... + \beta_{fj}(H_f) + r_{ij}, \quad i = 1, \ldots, N,$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(aH_1) + \gamma_{02}(aH_2) + ... + \gamma_{0f}(aH_f) + \gamma_{0(f+1)}(S_1) + ... + \gamma_{0l}(S_l) + u_{0j}$$
$$\beta_{0j} = \gamma_{00} + \mu_{0j}, \ \beta_{1j} = \gamma_{10}, \ \beta_{2j} = \gamma_{20} \ldots \ \beta_{lj} = \gamma_{lj}.$$

The Bayesian information criterion (BIC) was used to decide which model fitted the data best. One advantage of BIC (Schwartz, 1978) is that as long as the sample remains constant we have the possibility to compare competing models (whether or not they are nested) (McCoach & Black, 2008).

Results

The average PISA reading achievements among the Nordic countries were the lowest in Sweden and Denmark, and the highest in Finland (Table 3). Noteworthy, are the differences in average achievement in the more effective schools, were the Danish students exhibited the highest performance, followed by Finland, while Iceland, Norway and Sweden performed at a similar level. Furthermore, the less effective schools in Finland performed higher than the mid effective schools in all the other Nordic countries, except Denmark.

Table 3. Average reading achievement with standard errors in parenthesis; on average, and in less-, mid-, and more effective schools in the Nordic countries. The ICC attributable to schools and the between school variances with the proportion of between school variances within parenthesis from the multilevel analyses.

	Denmark	Finland	Iceland	Norway	Sweden
Average	495 (2.1)	536 (2.3)	500 (1.4)	503 (2.6)	497 (2.9)
Less effective	452 (2.9)	505 (2.2)	467 (3.1)	469 (2.8)	457 (4.0)
Mid effective	508 (2.1)	527 (1.9)	500 (2.2)	504 (2.2)	499 (2.4)
More effective	601 (2.1)	562 (2.4)	532 (2.7)	536 (3.1)	538 (4.2)
ICC attributable to schools	0.159	0.086	0.135	0.103	0.194
Between school variance	1137.61 (0.54)	655.14 (0.49)	1304.13 (0.57)	902.67 (0.56)	1989.29 (0.76)

The intraclass correlation (ICC), which shows the amount of the total variance in reading achievement that is attributable to schools, and the between school variances from the multilevel models are given in Table 3. The ICC ranged from 0.09 (Finland) to 0.19 (Sweden) suggesting that it could be relevant to examine multilevel models for these five countries. If

the ICC value is very low, not much information is added by using multilevel models as compared to linear regression. An interesting result in Table 3 is the low between school variance in Finland compared to the other countries, and in particular in comparison to Sweden which is the country that exhibits the largest difference between schools.

The results from the multilevel analysis are shown in Table 4. All student home background variables and factors were significant and had high coefficient values for all five Nordic countries. The fact that sex was significant is in line with e.g. Neuschmidt, Barth, Rutkowski, & Rutkowski (2009) but in opposite to Wiberg and Andersson (2010) who examined mathematics achievement in TIMSS 2007. In general, only a few school level factors were significant in the full models, a result which is similar to previous research on TIMSS 2003, 2007 and 2011 (e.g. Neuschmidt, Barth, Rutkowski, & Rutkowski, 2009; Wiberg & Andersson, 2010; Wiberg, Rolfsman & Laukaityte, 2013).

On level 2, the aggregated socioeconomic factor (SESm) was significant and very high in all five Nordic countries. In Denmark the aggregated mean of natives fathers were significant (compare Table 1) while in Iceland and Norway the aggregated reading activities were significant. Only Finland and Sweden had any significant school level factors. In Finland, the principal's sex (factor G) was significant in the sense that if the school has a female principal it detracts value. In Sweden, the location of the school (factor A) adds value while lack of school resources (factor C) detracts value.

Table 4. Parameter estimates for the home context and full multilevel models.

	Level 1				Level 2					
	READ	SES	FB	SEX	READm	SESm	FBm	A	C	G
Denmark	36.14***	26.43***	38.99***	13.77***		79.33***	42.86***			
	(1.56)	(3.01)	(3.43)	(2.36)		(14.47)	(8.74)			
Finland	39.17***	23.91***	37.81***	30.91***		94.62***				-8.57*
	(1.65)	(2.84)	(7.88)	(2.54)		(17.91)				(4.20)
Iceland	41.30***	26.19***	26.19***	23.83***	42.98***	58.82***				
	(1.75)	(3.27)	(7.24)	(3.17)	(13.05)	(20.46)				
Norway	40.18***	29.36***	31.82***	25.84***	28.70**	71.18***				
	(1.56)	(2.70)	(4.30)	(2.34)	(9.02)	(14.08)				
Sweden	40.62***	33.95***	32.82***	19.65***		90.34***		8.80**	-11.38*	
	(1.97)	(2.99)	(4.45)	(2.70)		(21.15)		(3.31)	(4.72)	

*** - p-value < 0.001, ** - p-value < 0.01, * - p-value < 0.05.
READ = Reading activities, SES = Socioeconomic status, FB = Father born, SEX=Sex, READm, SESm, and FBm are aggregated student level factors. A = Structure and organization of school, C = School resources, G = Principal's sex.

Discussion and implications

Our results show that only Finland and Sweden exhibit any significant school level factors in PISA. The results for Sweden, has provided us with two major factors that should be feasible to influence; aspects related to the location of the school (factor A) and lack of schools resources (factor C). The fact that there are quite large differences between schools in Sweden, and Sweden also has the highest amount of the total variance in achievement attributable to the schools among the Nordic countries (cf Table 3), is something that could be expected to be reflected in the presence of several school level factors. The results for Sweden could rather be summarized as; to the extent that the effectiveness of the Swedish schools is

related to school factors, the only factors that seem to be of importance are related to structural factors. Consequently, "softer" factors like the climate of the school and the administration and management seem to lack importance. The results for Finland will, according to our view, rather give raise to new questions than provide us with a foundation for school factors that could be influenced, e.g. *Why does a female principal detract the value of the school? Why is this a significant factor only in Finland?*

As noted in previous studies (Wiberg, Rolfsman & Laukaityte, 2013; Wiberg & Rolfsman, 2013), after controlling for student' home background, school level factors accounted for only a small percentage of the additional variance in the countries. A general conclusion is thus that when controlling for students' home background the school factors appear to be of less importance. It is however interesting to note that socioeconomic status appear to add more value than reading activity in contrast to a study of mathematics achievement in TIMSS 2003, 2007 and 2011 in Norway and in Sweden where mathematics-self-concept had a higher coefficient value than socioeconomic status (Wiberg, Rolfsman & Laukaityte, 2013). A similarity with that study, however, is that the aggregated socioeconomic status coefficient values were high and significant in all Nordic countries.

The results from PISA, when examining school effectiveness, are completely different in comparison to our previous results on TIMSS. When examining school effectiveness in relation to mathematics achievement in TIMSS in Sweden and Norway in 2003, 2007 and in 2011 (Wiberg, Rolfsman & Laukaityte, 2013), no significant school level factor were found in Sweden, but in Norway where the factors differ over the years. In 2003, it was School location; in 2007 Lack of school resources and in 2011 it was Poor student attendance at school and School climate teachers. In a comparison between Sweden and Norway, where we use science achievement as a criteria for school effectiveness, some school level factors were found to be significant in Sweden (in 2003, School attendance and in 2007, Teacher competence). Some school level factors were also found in Norway, but only in 2007 (Negative School behavior and School location).

An explanation to the differences found in TIMSS may be related to the difficulties of comparing results at three different time points. In some cases, items in the questionnaires have been altered to such large extent that some items had to be excluded which may have had an impact of the feasibility of identifying school level factors of importance for success in the two countries. However, the differences between Sweden and Norway may also be related to other aspects. PISA and TIMSS school questionnaire may be measuring different aspect and in relation to different criteria. This raises the question whether there is a difference between school level factors in relation to different kind of success among the students. This is however, a validity issue that requires further analyses and is something that should be pursued in further research.

References

Creemers B., Kyriakides, L., & Sammons P. (2010): *Methodological Advances in Educational effectiveness Research*: Milton park, Abington: Routledge.

Egelund, N. (2012). Conclusion. In N. Egelund (Ed.), *Northern Lights on PISA 2009 – focus on reading.* Copenhagen: Nordic Council of Ministers. TemaNord 201:501

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

Johansson, Olof & Bredeson Paul (2011): Framtida forskningsperspektiv på rektor – vilken forskning saknas. [Future research perspectives on principals – which research is lacking? in Swedish]. In O. Johansson (red.), *Rektor en forskningsöversikt 2000-2010* [Principal – a review of research 2000-2010], kapitel 4, 61-74 (Vetenskapsrådets rapportserie, 4). Stockholm: Vetenskapsrådet.

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.

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.

McCoach, B. D. & Black, A.C. (2008). Evaluation of model fit and adequacy. In A.A. O' Connel & D.B. McCoach (Eds.), Multilevel modeling of educational dats (pp. 245-272). Charlotte, nC: Information age publishing inc.

Muthén, L. K. & Muthén, B. O. (2011). MPLUS (Version 7) [computer software]. Los Angeles, CA.

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. (2003). Effective schools in Arab educational systems: an analysis of TIMSS 2003, *3rd IEA International Research Conference*. Taipei.

OECD (2009). The PISA 2009 International data base. Data retrieved April 2013 from http://pisa2009.acer.edu.au/

OECD (2013). *Are countries mowing towards more equitable education systems?* PISA in Focus, 2013/02 (February).

Rolfsman E., & Wiberg, M. (2012): *School effectiveness and Mathematical Achievement results from TIMSS in Sweden and Norway* (Manuscript submitted for publication).

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979): Fifteen thousand hours: secondary schools and their effects on children. London: Open book.

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

Schwartz, G. (1978). Estimating the dimensions of a model. Annals of Statistics, 6, 461-464

Swedish National Agency for Education (2009). *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer.* [What influences the results in Swedish comprehensive school? An overview of the importance of different factors, in Swedish] *Stockholm*: Skolverket

Swedish National Agency for Education (2012a): *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid* [Equity in Swedish comprehensive school?

A quantitative analysis of equity over time, in Swedish] (rapport 374): Stockholm: Skolverket.

Swedish National Agency for Education (2012b): Högpresenterande elever, höga prestationer och undervisningen [High performing students, high performance and teaching, in Swedish] (rapport 379): Stockholm: Skolverket

Tabachnik, B.G. and Fidell, L.S (2007). Using Multivariate Statistics, (Fifth ed.). Pearson Education, Inc.

Thornval, M. (2012). Sammenfatning In N. Egelund (Ed.), *Northern Lights on PISA 2009 – focus on reading.* Copenhagen: Nordic Council of Ministers. TemaNord 201:501

Wiberg, M., & Andersson, E. (2010): School effectiveness in Sweden compared with countries in Europe and Asia-Pacific. Paper presented at IEA International research conference, Gothenburg, July 1-3, 2010. Retrieved December 1th, 2012 from www.iea-irc.org/fileadmin/IRC_2010_papers/TIMSS/Wiberg_Andersson.pdf

Wiberg, M. & Rolfsman, E. (2013). School effectiveness in Science in Sweden and Norway viewed from a TIMSS perspective. Accepted for publication.

Wiberg, M., Rolfsman, E., & Laukaityte, I. (2013). School effectiveness in mathematics in Sweden and Norway 2003, 2007, and 2011. Paper presented at the 5th IEA research conference, 26-28 June, 2013, Singapore.

Wu, M. (2010). Comparing the similarities and differences of PISA 2003 and TIMSS, OECD Education Working papers, No 32, OECD Publishing http://dx.doi.org/10.1787/5km4psnm13nx-en