Utbildning & Demokrati 2013, vol 22, nr 3, 69–84
Tema: Storskaliga studier

# School effectiveness in science in Sweden and Norway viewed from a TIMSS perspective

*Marie Wiberg & Ewa Rolfsman*

Declining achievement in various core subjects has led to a debate on strategies to enhance student achievement. Identifying factors in the school environment that affect student performance in science, are therefore, of vital importance. The aim of this study is to identify school-level factors that are associated with eighth-grade students' achievement in science based on results from TIMSS (Trends in International Mathematics and Science Study) 2003 and 2007. Because the TIMSS data includes school-level factors at two different time points, we expected to find factors that influence science performance by Swedish and Norwegian students. Multilevel analysis was used, and this framework allowed us to account for the influence of the students' home backgrounds. After controlling for student background, our results show that there are only a few school-level factors that are associated with student achievement in science, and the influence of these factors differ between Sweden and Norway.

Keywords: Nordic context, school questionnaires, principals, multilevel analysis.

## Introduction

Declining achievement levels among students in Sweden in several core subjects has brought about a debate concerning strategies for improving student achievement. A significant amount of research

---

*Marie Wiberg* is associate professor and lecturer in statistics at the Department of Statistics at Umeå School of Business and Economics, Umeå University, SE-901 87 Umeå, Sweden. Email: marie.wiberg@stat.umu.se
*Ewa Rolfsman* has a PhD in pedagogy and is lecturer in educational measurement at the Department of Applied Educational Science at Umeå university, SE-901 87 Umeå, Sweden. Email: ewa.rolfsman@edusci.umu.se

has been conducted in Sweden regarding student achievement, and these studies have reviewed factors that influence student success at different levels of the educational system and from different perspectives. The link between students' socioeconomic backgrounds and academic achievement is well established (see e.g., Skolverket 2009). Against this background, a pressing issue is to identify other factors that can be modified and that might enhance academic achievement. This is particularly important because recent data has shown that the variability between schools in Sweden has increased over the years (Skolverket 2012).

The focus of this study was to identify school-level factors associated with student achievement in science based on the Trends in International Mathematics and Science Study (TIMSS) data collected in 2003 and 2007. The results of our study are relevant to teachers, policy makers, researchers, and anyone interested in school effectiveness and the prerequisites for student learning and academic achievement.

Identifying factors in the school environment that affect academic achievement from different perspectives is an important part of showing to what degree schools are successful in their task of providing good educational opportunities to their students. Such knowledge will allow for improvements in educational development and achievement. In this context, the principal is a key actor because he or she is ultimately responsible for the school's success and the students' academic achievements, which are critical measures of the "effectiveness" of the school.

Efforts to improve school effectiveness have attracted a sizeable amount of international research in a variety of different fields, including research on education policy, curriculum development, and school organisation (Cremers, Kyriakides & Sammons 2010). The basic premise behind school effectiveness research is that student achievement varies between schools and is related to the extent that a school "…'adds value' by realizing the potential of the student body through efficient organization and effective instruction." (Martin, Mullis Gregory, Hoyle & Shen 2000, p. 9).

Modern research into school and education effectiveness is based on a reaction to research conducted in the 1960s and 70s showing that schooling accounted for a rather small proportion of student achievement. One of the first school effectiveness studies was the classic study *Fifteen Thousand Hours* (Rutter, Maughan, Mortimore, Outson & Smith 1979) that was conducted in a school in London. This study showed that the climate in which the learning took place was directly associated with student outcomes. In general, a school's climate was described as the deep-rooted values and norms that were manifested on a daily basis by the teachers and other professionals within the school. The authors called this the school's 'ethos'. In Sweden, Lennart Grosin (1993, 2002) has stressed the importance of the school's leadership in relation to the school's pedagogical and social climate and how this affects student outcomes. The climate of the school is affected by the expectations and attitudes among the educational professionals of the school, and the principal is the primary authority in this regard.

Very little research has been carried out on the relationship between school effectiveness and student outcomes in terms of academic achievement, especially in regard to the role of the school principal. This is the conclusion from a review of research on school leaders between 2000 and 2010 (see Johansson & Bredeson 2011, Møller 2011, Ärlestig & Johansson 2011). Furthermore, those studies that have been performed have been based on interview data and there is a lack of studies based on large-scale data (Johansson & Bredeson 2011). The current study will, therefore, make a contribution to this field of research because it is based on large-scale data in regard to the influence of school principals on student achievement. In addition, the methods used in the TIMSS data collection makes it possible to separate the effects of school variables from the effects of the students' home environments, and this is in line with the results from previous research on school effectiveness that used the TIMSS data (Martin, Mullis, Gregory, Hoyle & Shen 2000, Neuschmidt, Hencke, Rutkowski & Rutkowski 2008). This is important because it is well known that there is a general correlation between students' socioeconomic background and their level of academic achievement (see e.g. Mullis, Martin & Foy 2008, Skolverket 2009).

School effectiveness can be defined in different ways, including student achievement in different subjects (see e.g. Neuschmitdt, Hencke, Rutkowski & Rutkowski 2008), social development, and other non-cognitive outcomes (see e.g. Creemers & Kyriakides 2010). School-level variables related to school effectiveness can be categorized into context and climate variables (Ma, Ma & Bradley 2008). Examples of context variables are school location, school size, resources, and equipment. Examples of climate variables are administrative polices, values, and expectations of students and parents.

The term "school effectiveness" can be defined in different ways. Bert Creemers, Leonidaas Kyriakides, and Pam Sammons (2010) elucidate that the terms "school effectiveness", "teacher effectiveness", and "educational effectiveness" are used inconsistently in the literature although these are three separate but related concepts. In this study we use the term "school effectiveness" because our data is composed primarily of school-level data and student outcomes that are solely based on cognitive achievements, i.e., test results from the TIMSS. Previous studies have used TIMSS data to measure student achievement in mathematics with respect to school effectiveness (Wiberg & Andersson 2010, Wiberg, Rolfsman & Laukaityte 2013). The results of those studies suggested that only a few school-level factors appear to be related to school effectiveness in Sweden. To further investigate the impact of the school-level factors measured by TIMSS, the focus of this paper is on school effectiveness in relation to student achievement in science. The main purpose of this study is to identify factors, from the perspective of the principals, which might help explain why some schools are more effective than others in terms of their students' achievement in science. The data came from TIMSS 2003 and 2007 and we compared student results from Sweden and Norway.

## Methodology

### Participants

This study used data from TIMSS 2003 and 2007 that was related to achievements in science for eighth-grade students from Norway and Sweden (IEA 2003, IEA 2007). Student achievements in science, student questionnaires, and school questionnaires were used. The school questionnaires were used as a measure of the context and the climate of the school. The school questionnaires covered several broad areas and included questions about student behaviour, resources and technology at the school, the teachers, and the context and the climate in which the learning took place. The survey also included questions about instruction and learning in mathematics and science, the principal's own role at the school, and questions about parental involvement. The student questionnaires were used to rule out the influence of the students' home backgrounds and to control for these factors.

### Statistical theory

The theory of multilevel analysis, especially hierarchical linear modelling (e.g. Gelman & Hill 2007, Raudenbush & Bryk 2007) in combination with the ideas of Xin Ma, Lingling Ma and Kelly Bradley (2008), was used to examine school-level factors when controlling for the student backgrounds. The framework of multilevel analysis was chosen due to the sampling procedure used in TIMSS (Kyriakides & Charalambous 2005) and to control for factors that are not connected to school effectiveness. We are aware that using hierarchical linear modelling is somewhat limiting because it only allows us to identify linear relationships between the variables. However, this level of analysis was believed to give valuable information in this case. The IEA IDB analyzer (IEA 2012) was used to prepare the data files for analysis with SPSS 20.

The formal analysis was performed in four steps. In the first step the variables from the student questionnaires in regard to the students' home backgrounds was examined. Based on results from previous studies (Wiberg & Andersson 2010, Wiberg, Rolfsman & Laukaityte 2013), we used two variables; student's sex and if the student's father was born in the country or not. In addition to these variables, two factors were constructed. The first was science self-concept, and this assessed the student's attitude towards science. The second factor was the student's socioeconomic status as measured by variables that previous TIMSS studies have shown to be linked to socioeconomic status, i.e. it includes home possessions such as having a calculator, computer, study desk, dictionary and books at home. The construction of the two factors was done in a similar way as in Marie Wiberg, Ewa Rolfsman and Inga Laukaityte (2013) study and inspired by the Ebrahim Mohammadpour and Mohamed Ghafar (2012) study with reference to how they had constructed the students' socioeconomic. In addition, their choice of variables for measuring mathematics self-concept inspired our choice of variables for measuring science self-concept. The student variables, including the constructed factors, will be referred to as student factors from this point (Table 1). Because the amount of missing data was low in the home background variables, ranging from 0% (for sex) to 5.4% (for science self-concept in Norway in 2003), we used listwise deletion to exclude missing data (Tabachnick & Fidell 2007). We are aware that this may be limiting and it might have been better to impute missing data, but this choice was made because it was only a small number of cases that were removed.

*Table 1. Descriptions of the student-level factors together with their mean and standard deviation within parenthesis for the first two factors and the proportion (%) for the latter two.*

| Factor | Description of factor | Norway | | Sweden | |
|---|---|---|---|---|---|
| | | 2003 | 2007 | 2003 | 2007 |
| SSC | **Science self-concept** Contains the item: How much do you agree with these statements about learning science? 1) I usually do well in science. 2) Science is more difficult for me than for many of my classmates. 3) Science is not one of my strengths. 4) I learn things quickly in science.<br>The original options were 1: Agree a lot, 2: Agree a little, 3: Disagree a little, and 4: Disagree a lot. The coding was reversed for the positively formulated questions and the responses were averaged and classified into three categories. 1: Low, average is less than or equal to 2; 2: Medium, average is greater than 2 and less than 3; 3: High, average is greater than or equal to 3. Note, in Sweden in 2003 the students answered the 'science' item for the subjects Chemistry, Earth Science, Geography, and Physics and these responses were weighted together in this study. | 2.35 (.65) | 2.19 (.73) | 2.15 (.67) | 2.20 (.68) |
| SES | **Socioeconomic status:** Contains items:<br>1) About how many books are there in your home? (Do not count magazines, newspapers, or your school books.)<br>2) Do you have any of these at your home- a) Calculator, b) Computer (do not include PlayStation, Game Cube, Xbox, or other TV/video game computers), c) Study desk/table for your use, d) Dictionary.<br>1) Book ownership: 1: 0–10 books, 2:11–25 books, 3: 26–100 books, 4: 101–200 books, and 5: more than 200 books. Recoded as: Low (1,2), Medium (3,4), and High (5).<br>2) Home resources (calculator, study desk, computer, dictionary). Yes/No. Number of positive responses was counted: 1 = Low, one or zero resources, 2 = Medium, two resources, and 3 = High, three or more resources.<br>Both indicators were averaged and categorized into: 1 = Low, average less than or equal to 1.5, 2 = Medium, average is higher than 1.5 and lower than 2.5 and 3 = High, average is higher than 2.5. | 2.55 (.35) | 2.50 (.37) | 2.55 (.37) | 2.52 (.37) |
| SEX | The gender of the student was recoded as 0 for female and 1 for male. (% male) | 51 | 50 | 51 | 52 |
| FB | Coded as 1 if the father was born in the country, otherwise 0. (% native fathers) | 93 | 88 | 90 | 83 |

In the second step of the analysis, we examined which schools were expected to be effective by using constructed regression models. We assumed that a school is more effective if its mean science achievement was higher than predicted from examining multiple regressions of the identified home background factors. The mean differences between the five science plausible values and the expected scores from the regressions were calculated for both countries at both time points. Schools were regarded as more effective if they were in the top third in their country in science achievement, mid-effective if they were in the middle third in their country and less effective if they were in the bottom third in their country.

In the third step of the analysis, we tried to identify school variables that might have an impact on the school's effectiveness. We reduced the number of school variables by constructing reasonable school-level factors through inspiration from previous studies (Mohammadpour & Ghafar 2012, Wiberg, Rolfsman & Laukityte 2013), reviews of the literature, and the variables available in TIMSS as described below and with mean and standard errors (and in one case percentage) given in Table 2.

**Teacher Competence:** Contains the item: How would you characterize each of the following within your school? a) Teachers' job satisfaction, b) Teachers' understanding of the school's curriculum goals), c) Teachers' degree of success in implementing the school's curriculum, and d) Teachers' expectations for student achievement. The responses were reverse coded, averaged, and categorized into: 3: High, average is less than or equal to 2, 2: Medium, average is greater than 2 but less than 3, and 1: Low, average is greater than 3.

**School Attendance:** Contains the item: How often does each of the following behaviours occur among eighth-grade students in your school, and; If the problem occurs, how severe a problem does it represent? A. Frequency in your school; B. Severity of problem in your school. a) Arriving late at school, b) Absenteeism, c) Skipping class/hours/periods. Recoded as 1: Low if response is never to all statements in A or not a problem in all statements in B, 2: Medium, most of the other response combinations, 3 = High, responded "serious problem" for two or three statements in A or "serious" problem for two statements in B.

**Negative student behaviour:** Contains the item: How often does each of the following behaviours occur among eighth-grade students in your school, and if the problem occurs, how severe a problem does it represent? A. Frequency in your school; B. Severity of problem in your school a) Classroom disturbance, b) Intimidation or verbal abuse of other students, c) Intimidation or verbal abuse of teachers or staff. The answers were categorized as 1: Low if response is "never" to all three statements in A or "not a problem" to all three statements in B. 2: Medium, most of the other response combinations, and 3: High, responded "serious problem" for two or three statements in A or "serious" problem for two statements in B.

**School location:** Contains the item: How many people live in the city, town, or area where your school is located? Recoded as 1 = Urban (original code 4-6), 0 = rural (original code 1-3).

Lack of school resources for science instruction: Contains the item: Is your school's capacity to provide instruction affected by a shortage or inadequacy of any of the following? – a) Science laboratory equipment and materials, b) Computers for science instruction, c) Computer software for science instruction, d) Calculators for science instruction, e) Library materials relevant to science instruction, f) Audio-visual resources for science instruction. Responses were averaged and categorized as 1: Low, average of items (a-f) is less than 2; 2: Medium, most other combinations; and 3: High, average of items (a-f) is greater than or equal to 3.

**Parental Involvement:** Contains the item: Does your school ask parents to do the following: a) Attend special events, (e.g., science fair, concerts, sporting events), b) Raise funds for the school, c) Volunteer for school projects, programs, and trips, d) Ensure that their child completes his/her homework, e) Serve on school committees, (e.g. Selecting school personnel, reviewing school finances). Reponses were recoded as 1 if parents were involved and zero otherwise. The number of responses were summarized and categorized as 1: Low if 0 or 1; 2: Medium if 2 or 3; and 3: High, if 4 or 5.

**Teacher Professional Competence:** Contains the item: In the past two years, what percentage of your eighth-grade teachers have been involved in professional development opportunities for mathematics and science targeted at the following: a) Supporting the implementation of the national or regional curriculum, b) Designing or supporting the school's own improvement goals, c) Improving content knowledge, d) Improving teaching skills, and e) Using information and communication technology for educational purposes. Recoded into five categories and averaged and categorized into 1: Low if the average was less than or equal to 1.5; 2: Medium if the average was above 1.5 and below 3; and 3: High if the average was above 3.

*Table 2. School-level factors mean and standard deviation in parentheses.*

| Factor | Description of factor | Norway | | Sweden | |
|---|---|---|---|---|---|
| | | 2003 | 2007 | 2003 | 2007 |
| TC | Teacher competence | 2.36 (.48) | 2.32 (.49) | 2.44 (.57) | 2.38 (.53) |
| SA | School Attendance | 1.86 (.44) | 2.02 (.41) | 2.16 (.53) | 2.29 (.56) |
| NSB | Negative Student Behaviour | 1.81 (.54) | 2.26 (.57) | 1.86 (.58) | 1.95 (.60) |
| SL | School location (%) | 73 | 70 | 58 | 58 |
| SSR | Lack of school resources for science instruction | 2.26 (.62) | 1.62 (.70) | 1.90 (.59) | 1.39 (.67) |
| PI | Parental Involvement | 2.60 (.52) | 2.67 (.46) | 2.38 (.61) | 2.46 (.59) |
| TPC | Teacher Professional Competence | 1.91 (.48) | 2.34 (.69) | 1.78 (.51) | 2.37 (.58) |

A few principals did not answer any of the items in the school questionnaire so these schools were omitted from the analysis. For schools where only some items were missing, we used multiple imputations to replace missing values (see e.g. Schaefer & Graham 2002). This choice was made because exclusion of these schools would have excluded all of the students within that school. To use this method we must assume that the data is missing at random, and even if such an assumption were not valid this would only have a minor impact on the estimates and standard errors according to Linda Collins, Joseph Schaefer & Chi-Ming Kam (2001). Our linear imputation model included all of the variables that we wanted to include as predictors and response variables in the latter models that we subsequently analysed. In other words, the analysed model and imputed model were the same.

In the fourth step of the analysis, hierarchical linear modelling was carried out with HLM 6 (Raudenbush, Bryk, Cheong, Congdon & du Toit, 2004), which is built on the theory described in Stephen Raudenbush & Anthony Bryk (2002) and the suggestions on how to estimate effective schools from Xin Ma, Lingling Ma and Kelly Bradley (2008). For both years we examined the null, context and full models. The full model contains all possible school-level and student-level factors as well as aggregated student-level factors, the context model contains student-level factor and aggregated student-level factors, and the null model contained none of the factors. We used the five science achievement plausible values as dependent variables and grand mean centering of the variables. In the full model, the association between school-level factors and student science achievement was quantified while controlling for the home background of the students. At the first level, students' home background factors were entered. At the second level, aggregated student-level factors and the previously defined school-level factors were entered and weighted with school weights.

Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_1(H_1) + \beta_2(H_2) + \beta_3(H_3) + \beta_4(H_4) + r_{ij}.$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(aH_1) + \gamma_{02}(aH_2) + \ldots + \gamma_{04}(aH_4) + \gamma_{05}(S_1) + \ldots + \gamma_{011}(S_7) + u_{0j}$$

$$\beta_{0j} = \gamma_{00} + \mu_{0j},\ \beta_{1j} = \gamma_{10},\ \beta_{2j} = \gamma_{20} \cdots\ \beta_{11j} = \gamma_{11j},$$

where $Y_{ij}$ is the science achievement for each student $i = 1,2 \ldots n_j$ in school $j = 1,2,\ldots J$ , $\beta_{0j}$ is the mean science achievement of school $j$ and $r_{ij}$ is the random error of student $i$ in school $j$. Further, $\gamma_{00}$ is the grand science mean for all schools and $\mu_{0j}$ is the random school effect (the deviation of school $j$:s mean from the grand mean). $H_1$, $H_2$, $H_3$, and $H_4$ are the student-level factors in Table 1, $aH_1,\ldots,aH_4$

are aggregated student-level factors at school level and $S_1,\ldots,S_7$ are the school-level factors previously described with characteristics given in Table 2.

As suggested by Xin Ma, Lingling Ma and Kelly Bradley (2008), we used two measures to examine the models: intraclass correlation and the proportion of between-school variances at each level as explained by the hierarchical linear model. The intraclass correlation is constructed by dividing the between-school variance by the total variance and represents the percentage of variance in science achievement at the school. A low intraclass correlation value indicates that there is only a small added value for using multilevel analysis compared to the use of ordinary linear regression. The proportion of between-school variances explained by the multilevel analysis is constructed by dividing the difference in between-school variances for the full model and the null model by the between-school variances for the full model. If the between-school variance is high, it could be an indicator that school-level factors may have an effect on science achievement.

## Results

Our results show that there is a difference between Norway and Sweden regarding student achievement in science. The students' achievement in Sweden is consistently higher compared to the Norwegian students, although both Sweden and Norway had a general decline in science performance between 2003 and 2007 (Table 3). The results also reveal that the more effective schools in Sweden showed the largest decline while in Norway the decline was more evenly spread among the three different school types. The multilevel analysis using our defined student-level and school-level factors indicates that there are factors that influence school effectiveness in Sweden and Norway and that these factors are different at different time points (see Table 4). What is most evident is the rather large influence of the students' home background, which tends to explain most of the variance seen in the data as shown in Table 3. We also show that the intraclass correlation is attributable to schools and the total variance is best explained by between-school variance. The between-school variance was reasonably high indicating that school-level factors have an effect on science achievement.

*Table 3. Average science achievements with standard errors in parentheses. The average science achievements for the less, mid, and more effective schools in Norway and Sweden in 2003 and 2007 are shown. The intraclass correlation (ICC) attributable to schools and the proportion of the total variance explained by between-school variance are obtained from the multilevel analyses.*

| Country | Norway | | Sweden | |
|---|---|---|---|---|
| | 2003 | 2007 | 2003 | 2007 |
| Average | 493.9 (2.2) | 486.8 (2.2) | 524.3 (2.7) | 510.7 (2.6) |
| Less effective schools | 470.0 (2.9) | 464.4 (2.6) | 484.1 (3.5) | 482.0 (3.4) |
| Mid effective schools | 494.8 (2.0) | 483.9 (2.1) | 525.8 (3.0) | 512.4 (2.0) |
| More effective schools | 516.7 (2.4) | 514.1 (2.3) | 559.4 (2.5) | 539.5 (3.0) |
| ICC attributable to schools | 0.06 | 0.06 | 0.12 | 0.09 |
| Between-school variance | 0.30 | 0.45 | 0.60 | 0.51 |

The results from the multilevel analyses of the two countries at the two time points are shown in Table 4. We have excluded school-level factors from the table that were non-significant when using a significance level of .05. Three out of the four student-level factors were significant in both countries and at both time points, but the forth factor (sex) was only significant in Sweden in 2003. Only a few school-level factors were significant in the two countries at the two time points. In Norway 2003, none of the school-level factors were significant, but in 2007 negative student behaviour was significant in Norway. In Sweden, aggregated socioeconomic status was significant in both 2003 and in 2007 but among the school-level factors, only lack of school resources in science was significant in 2003.

*Table 4. Results from the multilevel analysis showing the value of the coefficients of the significant factors.*

| | Level 1 | | | | | Level 2 | | |
|---|---|---|---|---|---|---|---|---|
| | Intercept | SSC | SES | FB | SEX | aSES | NSB | SSR |
| *Norway* | | | | | | | | |
| 2003 | 500.04 | 26.77 | 43.13 | 40.02 | | | | |
| 2007 | 489.56 | 39.41 | 51.44 | 30.42 | | | 11.33 | |
| *Sweden* | | | | | | | | |
| 2003 | 521.82 | 27.06 | 27.06 | 25.35 | -8.04 | 87.45 | | -8.76 |
| 2007 | 515.26 | 38.81 | 71.29 | 25.76 | | 71.29 | | |

SSC = Science-Self-Concept, SES = Socioeconomic status, FB = Father born in country, SEX = Sex, aSES =aggregated socioeconomic status, NSB = Negative student behaviour, and SSR = Lack of school resources in science.

## Concluding remarks

In this study a comparison was made between Norway and Sweden. In line with previous results from analysis of TIMSS data in regard to student achievement in mathematics (Wiberg, Rolfsman & Laukaityte 2013) the Swedish students had higher levels of achievement in science than the Norwegian students. However, in contrast to results for mathematics achievement, both Sweden and Norway had a general decline in regard to science achievement between 2003 and 2007.

Our analysis included a comparison between three types of schools; the more effective schools, the less effective schools, and the mid effective schools. The results show that there is quite a large difference in achievement between the more effective schools and the less effective schools, in particular in Sweden. A comparison between the two countries reveals a decline in achievement in both countries regardless of school type. In contrast to Norway, this seems to be related to the level of achievement in Sweden. The result showing that the more effective schools in Sweden exhibited the largest decline in achievement over the four years between the TIMSS data collections. This is something that should be investigated further. One can only speculate whether this could be related to the increase in school segregation in Sweden (cf. Skolverket 2012).

Regarding *student*-level factors and their ability to explain school effectiveness as manifested in student achievement, our results show that three out of four factors (see table 4) were significant in both countries in 2003 as well as in 2007. Regarding *school*-level factors, only a few of these were significant and these were specific in terms of country and year. This result is in line with previous studies based on TIMSS data in mathematics and therefore, is not surprising (cf. Mohammadpour & Ghafar 2012, Wiberg & Andersson 2010, Wiberg, Rolfsman & Laukaityte 2013). Nevertheless, these results require further consideration in relation to issues of validity if they are to be useful tools for school development. An overarching question is why only these school-level factors are significant and why they are only significant during specific years. Furthermore, it is noteworthy that parental involvement, teacher (professional) competence, school location and school attendance do not seem to be of importance for the effectiveness of a school in teaching science because neither of these school-level factors turned out to be significant. Noteworthy is that aggregated socioeconomic status, i.e. the socioeconomic context in the schools, had large coefficients at both time points in Sweden but was non-significant in Norway at both time points. A limitation of this study is that we constructed student-level factors and school-level factors to be able to compare the two countries at different time points. It was, however, sometimes problematic to construct meaningful factors that can be used over time because some items and some scales differed between the years in which the data was collected.

In this study, the selection of variables regarding student-level factors and school-level factors was based primarily on variables that proved to be of significant value in previous studies in the field (Mohammadpour & Ghafar 2012, Wiberg & Andersson 2010, Wiberg, Rolfsman & Laukaityte 2013). This may be a key to our results showing that three out of the four student-level factors were significant in 2003 as well as in 2007 in both countries, which is in contrast to our previous results from studying student achievement in mathematics (Wiberg & Andersson 2010, Wiberg, Rolfsman & Laukaityte 2013). The forth student-level factor sex, was only significant in Sweden in 2003. An important challenge for further studies using the TIMSS data is to continue to develop meaningful school-level factors that can be used in different countries and at different time points. Further studies should also include more than two time points to be able to better investigate changes among school-level factors and their relevance to student outcomes.

**Acknowledgement**
This study has been conducted within the project *Analysis and modelling of Swedish students' performance in TIMSS and PISA in an international perspective* (2008-4027) which is funded by the Swedish Research Council.

## References

Collins, Linda; Schaefer, Joseph & Kam, Chi-Ming (2001): A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6 (4), 317-329.
Creemers, Bert; Kyriakides, Leonidas & Sammons, Pam (2010): *Methodological Advances in Educational Effectiveness Research*: Milton park, Abington: Routledge.
Gelman, Andrew & Hill, Jennifer (2007): *Data Analysis Using Regression and Multilevel/hierarchical Models*. New York: Cambridge University press.
Grosin, Lennart (1993): School effectiveness research as a point of departure for school evaluation. *Scandinavian Journal of Educational Research*, 37 (4) 317-330.
Grosin, Lennart (2002): Rektorer i framgångsrika skolor. *Nordisk pedagogik 22* (3), 158-175.
IEA (2003): *TIMSS 2003 international database*. Obtained from the official TIMSS 2003 IEA web site http://www.iea.nl.
IEA (2007): *TIMSS 2007 international database*. Obtained from the official *TIMSS 2007 International Database* CD.
IEA (2012): IEA IDB analyzer version 2.0 [Software] http://www.iea.nl/data.html.
Johansson, Olof & Bredeson Paul (2011): Framtida forskningsperspektiv på rektor – vilken forskning saknas. In O. Johansson (ed.), *Rektor en forskningsöversikt 2000-2010* kapitel 4, 61-74 (Vetenskapsrådets rapportserie, 4). Stockholm: Vetenskapsrådet.
Kyriakides, Leonidas & Charalambous, Charalambos (2005): Using educational effectiveness research to design international comparative studies: turning limitations into new perspectives. *Research paper in Education*, *20 (4)*, 391-412.
Ma, Xin; Ma, Lingling & Bradley, Kelly (2008): Using multilevel modelling to investigate school effects. In Ann O´Connell & Betsy McCoach (red) *Multilevel Modelling of Educational Data*, pp. 59-110. Information Age Publishing: Charlotte, NC.
Martin, Michael; Mullis, Ina; Gregory, Kathleen; Hoyle, Craig & Shen, Ce (2000): *Effective schools in science and mathematics*. IEA's third international mathematics and science study, IEA: Chestnut Hill, MA.
Mohammadpour, Ebrahim & Ghafar, Mohamed (2012): Mathematics achievement as a function of within- and between school differences. *Scandinavian Journal of Educational Research,* DOI:10.1080/00313831.2012.725097.

Mullis, Ina; Martin, Michael & Foy, Pierre (in collaboration with Olson, John; Preuschoff, Corinna; Erberber, Ebru; Arora, Alka & Galia, Joseph) (2008): *TIMSS 2007 International Mathematics Report. Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education.

Møller, Jorunn (2011): Research about principals in Norway. In Olof Johansson (Ed.), *Rektor en forskningsöversikt 2000-2010,* , 112-133 (Vetenskapsrådets rapportserie, 4). Stockholm: Vetenskapsrådet.

Neuschmidt, Oliver; Hencke, Juliane; Rutkowski, Leslie & Rutkowski, David (2008): Effective schools in Arab educational systems: an analysis of TIMSS 2003, *3rd IEA International Research Conference*. Taipei.

Raudenbush, Stephen & Bryk, Anthony (2002): *Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd edition,* Thousand Oaks, CA: Sage Publications.

Raudenbush, Stephen; Bryk, Anthony; Cheong, Yuk Fai; Congdon, Richard & du Toit, Mathilda (2004): Hierarchical linear and nonlinear modelling. SSI scientific software.

Rutter, Michael; Maughan, Barbara; Mortimore, Peter; Ouston, Janet & Smith, Alan (1979): Fifteen Thousand Hours: Secondary Schools and Their Effects on Children. London: Open book.

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

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

Skolverket (2012): *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid* [Equity in Swedish comprehensive school? A quantitative analysis of equity over time.] (report 374). Stockholm: Swedish National Agency for Education.

Tabachnick, Barbara & Fidell, Linda (2007): *Using Multivariate Statistics*, (Fifth ed.). Pearson Education, Inc.

Wiberg, Marie & Andersson, Ewa (2010): *School effectiveness in Sweden compared with countries in Europe and Asia-Pacific.* Paper presented at the 4th IEA International research conference in Gothenburg, July 1-3, 2010. Retrieved December 1th, 2012 from www.iea-irc.org/fileadmin/IRC_2010_papers/TIMSS/Wiberg_Andersson.pdf

Wiberg, Marie; Rolfsman, Ewa & Laukaityte, Inga (2013): *School effectiveness in mathematics in Sweden and Norway 2003, 2007 and 2011*. Paper presented at the 5th IEA International Research Conference in Singapore, July 28-30, 2013. Retrieved October 2nd 2013 from http://www.iea.nl/fileadmin/user_upload/IRC/IRC_2013/Papers/IRC-2013_Wiberg_etal.pdf

Ärlestig, Helene & Johansson, Olof (2011): Publicerad forskning om rektorer Sverige. I O. Johansson (red.), *Rektor en forskningsöversikt 2000-2010*, kapitel 3, 43-60 (Vetenskapsrådets rapportserie, 4). Stockholm: Vetenskapsrådet.