Knowledge and skills in Swedish comprehensive school during 40 years: A reassessment of the evidence based on the IEA studies

Jan-Eric Gustafsson
Department of Education, University of Gothenburg

Abstract

The IEA studies have been the major source of information about the level of achievement in the Swedish comprehensive school, which was implemented in the early 1960s. However, in the paper it is argued that the officially reported pattern of achievement is partially misleading, in that the level of achievement is underestimated in the studies conducted prior to 1995, while in the later studies achievement is somewhat overestimated. This is because schooling influences achievement more strongly than age, and in the early IEA studies age-based population definitions were used which caused the Swedish students with a comparatively high school start age to have fewer years of schooling than students in most other countries. A mathematical model is developed to adjust for the effects of this. The results from application of this model shows that the level of achievement of Swedish students was at a high international level in science and reading in the early 1990s, while after 1995 achievement has declined to a level close to the international mean.

Background and Study Purposes

In Sweden, as in many other countries, there has been a strong interest in obtaining information about how the level of achievement in different areas has developed over time. Sweden has participated in most of the IEA studies of student achievement that have been conducted since 1964, so these studies can potentially form a major source of information on the development over time. However, there are many methodological challenges involved in using results from these studies to analyze change in level of achievement. Thus, as has been shown by Medrich and Griffith (1992), the early IEA studies suffer from poorly documented sampling models and incomplete descriptions of the representativeness of the samples. While the technical quality of the IEA studies has been improved from the mid 1990s (Gustafsson, 2008), changes have been made in the basic design which makes it difficult to make comparisons over time. Furthermore, given that trend data has become available only during the last decade, use of the early studies must involve comparisons with other countries to provide any information about the change within Sweden.

However, given that no other source of information is available on the long-range development of knowledge and skills, it may be worthwhile to see if it is possible to extract meaningful information from the entire set of IEA studies on the development of the level of knowledge and skills in Sweden. This is the main purpose of the present paper. Another purpose is to approach some of the methodological challenges involved in making such comparisons over time.

One of the main methodological challenges of comparative studies of student achievement is to define the populations to be investigated in such a way that they are comparable across countries and over time. The major source of problems is that the school-start age varies across countries, which implies that if populations are defined on the basis of student age (e. g., 10 years old) the students will have a varying number of school-years behind them when the assessment is conducted. Similarly, if the populations are defined on the basis of grade (e. g., 4th grade), the samples of students from different countries will have different chronological age. If the sample is restricted to comprise students from a single grade or a year it thus is impossible to define populations which are completely comparable, which implies that comparisons will always be more or less biased. The amount and direction of bias depends both upon the population definition used and on the relative importance of one school year and one chronological year on student achievement. An age-based population definition is advantageous for countries with a low school-start age, and particularly so if one school year has a stronger influence on achievement than a chronological year, while a grade-based population definition is advantageous for countries with a high school start age, and particularly so if a chronological year has a stronger influence on achievement than has a school year.

In the IEA studies different population definitions have been used and these have changed over time. Three basic principles can be identified: (1) age based definition, which implies that the population includes students who belong to a certain birth cohort (e.g., 15-year-olds); (2) grade based definition, implying that the population comprises students who attend a particular grade; and (3) age-based grade definition, which implies that a grade based definition is used, but where the choice of grade is determined on the basis of which grade the majority of students in a certain birth cohort attend.

In the first IEA studies age based population definitions were used. However, when for example all 10-year-olds are to be included this generally implies that students will belong to two different grades, which is impractical in data collection. In several of the ensuing IEA studies the age-based grade definition was therefore adopted. However, the major IEA studies with a trend design which currently are running (i.e., PIRLS and TIMSS), rely on a grade based population definition. Thus, within IEA there has been a development from an age based population definition to a grade based population definition.

These changes make it important to find ways to assess the amount of bias associated with different population definitions. To determine the effects of different population definitions we need to know the relative importance of one year of schooling and one chronological year. Cahan and Cohen (1989) used a regression discontinuity design (RDD) to separate the effects of the two factors with respect to cognitive abilities, and concluded that schooling exerts twice a strong effect compared to chronological age. Recently, a series of studies have been conducted on IEA data which offer results for two adjacent cohorts, which is required for application of the RDD.

Luyten and Veldkamp (2008) applied an elaborated RDD design to data from grades 3 and 4 for 15 countries from TIMSS 1995. For science they found a grade effect of d = 0.33 and an age effect of d = 0.14. For mathematics the corresponding estimates were d = 0.44 and d = 0.16. Cliffordson (in press) used the Swedish data from TIMSS 1995 in grades 6, 7, and 8 in RDD designs, controlling for early and late school start. For science the grade effect was d = 0.28 and the age effect was d = 0.17. For mathematics the corresponding estimates were d = 0.24 and d = 0.17, respectively.

Gustafsson (2009a; see also Rosén & Gustafsson, 2006) used Swedish data from grades 3 and 4 in PIRLS 2001 to estimate the effect of age and grade on reading comprehension. The grade effect was d = 0.33 and the age effect d = 0.20. Luyten, Peschar and Coe (2008) used data from PISA 2000 to estimate effects of schooling on reading comprehension for 15 year olds. They found a weak grade effect (d = 0.12) and no age effect. However, estimates were imprecise, and Cliffordson and Gustafsson (2008) found a grade effect of around d = .20 on general cognitive abilities for 18 year olds. Given that it is not reasonable that general cognitive abilities are more affected by schooling than is reading comprehension, a more reasonable estimate of the grade effect on reading comprehension is d = .20.

In summary, the results show that the effect of a school year is stronger than the effect of a chronological year, and that the school year effect is stronger for younger than for older students. The school year effect also is stronger for knowledge and skills within mathematics and science than for general cognitive abilities and for reading. The results also show that up to age 15 there is an effect of chronological age which cannot be disregarded.

The strong effect of a school year implies that the early IEA studies underestimate the level of achievement in school systems with late school start. In Sweden school start age is 7, so compared to England, for example, where school starts at age 5, the Swedish students have two years less schooling. Thus, the Swedish results in the early studies are likely to have been underestimated. In contrast, the Swedish results in the recent IEA studies, which use a grade based population definition, are likely to be somewhat overestimated, because the Swedish students are older. In order to be able to describe the development of the level of achievement of the Swedish students over time, it therefore is necessary to make the estimates obtained from the early and the recent IEA studies more directly comparable.

Methodology and Data Sources

In this section the steps taken to make the estimates from studies using different population definitions more comparable are described, along with a short description of the data sources used.

Data

All the IEA studies investigating achievement in the fields of mathematics, science and reading that Sweden has participated in are relied upon in the current study. The different studies are presented below.

The first study within the field of mathematics (”First International Mathematics Study”, FIMS 1964) was conducted in 1964 (Husén, 1967a, b). The study included among others an age-based grade sample of 13 year-olds. In Sweden the participating students were in grade 7, while they were in grade 8 in most other countries. The second mathematics study (”Second International Mathematics Study”, SIMS 1980) was carried out during the years 1980-1982 (Robitaille & Garden, 1989). This study too comprised an age-based grade sample focusing 13-year-olds. The third mathematics study (”Third International Mathematics Study”, TIMSS 1995) was conducted in 1995 ((Beaton et al., 1996a, b) and employed new and more advanced techniques for measurement and design than previous studies. The study focused one population of young students (Population I, generally grades 3 and 4), and one population of older students (Population II, generally grades 7 and 8), and for each population students from two adjacent grades were sampled. Sweden did not participate with Population I, and for Population II, Sweden participated with grades 6 and 7. Of the 45 participating countries, only three were represented by grades 6 and 7, namely Denmark, Norway and Sweden. However, Sweden also participated with a smaller sample of students from grade 8, which was only reported in an appendix to the international report (Beaton et al., 1996a, b). The TIMSS study has since been repeated in 1999, 2003 and 2007 and the acronym now stands for “Trends in International Mathematics and Science Study.” Sweden participated with samples of grade 8 students in TIMSS 2003 (Mullis, Martin, Gonzalez, & Chrostowski, 2004a) and TIMSS 2007 (Mullis, Martin, Foy, Olson, Preuschoff, Erberber, Arora & Galia, 2008), and in TIMSS 2007 Sweden also participated with a grade 4 sample.

The first study of science (”First International Science Study”, FISS 1970) (Coomber & Keeves, 1973) was conducted within the framework of the large “Six Subject Survey” conducted in 1970-1971. The study comprised four different populations, and among these an age-based population of 10-year-olds (Population I) and an age-based population of 14-year-olds (Population II). The Swedish sample for Population I comprised older students from grade 3 and younger students from grade 4, while the sample for Population II comprised older students from grade 7 and younger students from grade 8. In both populations the Swedish students generally had one year less of schooling than students from other countries. IEAs second study of science (”Second International Science Study”, SISS 1984) was conducted in 1984 (Keeves, 1992; Postlethwaite & Wiley, 1991). In this study too Sweden participated both with Population I and Population II samples. A grade based population definition was used, but Sweden included two adjacent grades (grades 3 and 4 for Population I and grades 7 and 8 for Population II) to allow direct comparisons with FISS 1970. The TIMSS study described above not only comprised mathematics, but also science. As has already been mentioned, Sweden participated with samples from grade 8 in TIMSS 1995, TIMSS 2003 and TIMSS 2007, and also with a sample of students from grade 4 in TIMSS 2007. However, in TIMSS 1995 students from grades 6 and 7 were the official samples that were presented in the international and Swedish reports.

The first study of reading comprehension (”First International Reading Study”, FIRS 1970) was carried out as a part of the Six Subject Survey 1970 and was reported by Thorndike (1973). In this study there are results for Swedish students both in Population I (10-year-olds) and in Population II (14-year-olds). The second study of reading comprehension, here referred to as “Second International Reading Study” (SIRS 1991) was carried out in 1990/91 (Elley, 1994). In the study there were two age-based grade defined populations. Population I was an age-based of 9-year-olds and Population II of 14-year olds. Sweden participated with grade 3 and grade 8. So did some other countries, but most countries participated with grade 4 and grade 9. In 2001 the ”Progress in International Reading Literacy Study” (PIRLS), which focuses upon reading comprehension in grade 4 was conducted for the first time (Mullis, Martin, Gonzalez, & Kennedy, 2003). Sweden participated with samples of students both from grade 3 and grade 4. In 2006 the second round of PIRLS was conducted (Mullis, Martin, Kennedy, & Foy, 2007), and in this study Sweden only participated with students from grade 4.

Methods of analysis

The main method used is to adjust the observed results for Sweden to take into account the age and the number of years of schooling of the students in the samples, as compared with the mean age and mean number of years of schooling for all the countries. The procedures used are best described with a concrete example.

The FISS study in Population I described above will be used for the example. The Population I study comprised 16 countries, but the analyses reported here focus on the 12 countries that belong to what is now the EU/OECD groups of countries, or who have participated in later studies. The four excluded countries are Chile, India, Iran and Thailand. Table 1 provides descriptive data for the 12 countries.

Tabell 1. Results in FISS 1970 for Population I in 12 countries.

Country	Age	Grade	*d*
Japan	10.5	5.00	0.63
Sweden	10.5	3.53	0.20
Belgium (Fl)	11.0	5.10	0.15
USA	10.7	4.84	0.13
Finland	10.6	3.77	0.10
Hungary	10.7	4.29	0.00
Italy	10.7	4.96	-0.02
England	10.6	5.52	-0.12
The Netherlands	10.6	4.53	-0.17
Germany	10.5	4.61	-0.22
Scottland	10.6	5.64	-0.34
Belgium (Fr)	10.6	4.79	-0.35
Mean	***10.6***	***4.72***	***0.00***

The table presents information about the mean age and the mean grade of the students. It may be seen that the mean age for the students in most countries is close to 10.5, even though there is some variability. There is, however, more variability among the mean grades for the different countries. For Sweden the mean grade is 3.53, while the overall mean is 4.72. For England the mean grade is 5.52 and for Japan it is 5.00, just to mention two examples. Table 1 also presents the observed results, transformed into the d-scale. The transformation was done in such a way that the mean of the country means was first computed, along with the mean of the within country standard deviations. The deviation between each country mean and the overall mean was then computed and divided by the mean of the standard deviations. The countries are sorted according to level of performance, and it may be seen that Japan is the highest performing country (d = 0.60). Sweden is the second best performing country, but with a *d*-value of .20 Sweden is quite far behind Japan.

In order to adjust the Swedish results to be comparable over time and with other countries, the expected performance effect of the grade difference was first determined. Sweden has 1.19 school years less than the overall average, and according to the Luyten and Veldkamp (2008) estimate there is an effect of .33 on science for each grade for younger student. This implies that we need to add 1.19*0.33 = 0.39 to the observed value to compensate for the lower number of school years. The Swedish students also were somewhat younger than the overall mean (0.1) and with an age effect of 0.14, this implies that we need to add a further 0.1*0.14 = 0.01 to the observed *d*-value for Sweden. The adjusted d-value for Sweden thus amounts to 0.61, which is very close to the observed value for Japan.

Findings and Discussion

In the reanalysis of data the procedure described above has been adopted, and observed as well as adjusted values for the level of achievement have been computed. The results are presented below for the fields of mathematics, science and reading.

Mathematics

Table 2 presents observed and adjusted results for the studies in mathematics which Sweden has participated. The first column presents the observed d-values computed for the different studies. Positive values indicate achievement higher than the mean of the participating countries. It must be observed that these results in some cases are somewhat different that the officially reported ones, which is because the current analysis focuses on the subset of the countries that belongs to the EU/OECD group or that has participated in several studies. The following columns of Table 2 report the *d*-values for the difference between Sweden and the international mean with respect to age and grades, and the estimated effect of these differences based on the estimated effects of age and grade. In the last column the estimated effects have been added to the observed value to obtaine the adjusted value.

Tabell 2. Observerade och adjusted results in mathematics for Sweden 1964 – 2007.

	Observed value	Difference age	Difference grade	Effect age	Effect grade	Adjusted value
Population I						
TIMSS 2007 åk 4	-0.20	0.56	-0.12	-0.09	0.05	-0.23
Population II						
FIMS 1964	-0.56	-0.04	-1.00	0.01	0.24	-0.31
SIMS 1980	-1.04	-0.13	-0.94	0.02	0.31	-0.70

TIMSS 1995 åk 7	-0.12	-0.38	-1.08	0.08	0.26	0.22
TIMSS 1995 åk 8	0.27	0.63	-0.08	-0.11	0.02	0.18
TIMSS 2003	-0.15	0.63	0.05	-0.11	-0.01	-0.27
TIMSS 2007	-0.19	0.54	0.00	-0.09	0.00	-0.28

Note. For Population I it has been assumed that the effect of a chronological year is 0.16, and that the effect of a school year is 0.44. For Population II the corresponding estimates are 0.17 and 0.24.

For the single Population I study that Sweden has participated in (TIMSS 2007) both the observed and the adjusted value was somewhat below the mean for the comparison group of countries.

In the first two studies of Population II (FIMS 1964 and SIMS 1980) the Swedish results were very poor. In these studies the Swedish students had about the same age as the comparison group of countries, but they had been to school about one year less, because Sweden participated with students from grade 7, while most other countries participated with students from grade 8. Taking this into account with the adjusted value causes the Swedish results to get higher, but they are still considerably below the international mean.

In TIMSS 1995 the Swedish grade 7 students had improved their results, even though the observed value was still a bit below the international mean. Interestingly enough, Sweden also participated with a sample of grade 8 students, which was not included in the official reporting. The Swedish grade 8 students performed quite well, a bit below the international top level. This is the first study in which a Swedish sample had approximately the same number of school years as the students in the other countries. However, the Swedish grade 8 students were also older than the students in most other countries. Taking the effect of the age- and grade-differences into account, the Swedish adjusted achievement level is somewhat reduced, but is still above the international mean. The adjusted results for the grade 7 sample, is close to the adjusted value for the grade 8 sample.

Even though the Swedish results in TIMSS 1995 were quite good, the development since then has been quite negative. In TIMSS 2003 and 2007 Sweden participated with samples of grade 8 students, and the observed results for these were slightly worse than the results for the grade 7 students in TIMSS 1995. The adjusted values are below the international mean, and about .45 *d*-units lower than the TIMSS 1995 results.

In summary, the following pattern of results emerges for mathematics: In the early 1960s the Swedish mathematics results were poor in an international comparison, and this also was true in 1980. However, by mid 1990s, results had improved and the Swedish results were well above the international mean. Since then the results have deteriorated considerably.

Science

The observed and adjusted results for Sweden in the area of science are reported in Table 3.

Tabell 3. Observerade och adjusted results in science for Sweden 1970 – 2007.

	Observed value	Differenc e age	Differenc e grade	Effect age	Effec t grade	Adjuste d value
Population I						
FISS 1970	0.20	-0.10	-1.19	0.01	0.39	0.61
SISS 1984 åk 3	-0.14	-0.87	-1.56	0.12	0.52	0.49

SISS 1984 åk 4	0.30	0.13	-0.56	-0.02	0.19	0.47
TIMSS 2007 åk 4	-0.04	0.56	-0.12	-0.08	0.04	-0.08
Population II						
FISS 1970	0.20	-0.04	-1.11	0.01	0.31	0.52
SISS 1984 åk 7	-0.15	-1.00	-1.61	0.17	0.45	0.48
SISS 1984 åk 8	0.08	-0.00	-0.61	0.00	0.17	0.25
TIMSS 1995 åk 7	0.03	-0.38	-1.08	0.07	0.30	0.40
TIMSS 1995 åk 8	0.40	0.63	-0.08	-0.11	0.02	0.31
TIMSS 2003	0.08	0.63	0.05	-0.11	-0.01	-0.04
TIMSS 2007	-0.04	0.54	0.00	-0.09	0.00	-0.13

Note. For Population I it has been assumed that the effect of a chronological year is 0.13, and that the effect of a school year is 0.33. For Population II the corresponding estimates are 0.17 and 0.28.

Sweden participated with samples of students from Population I in the first two science studies. For grade 4 students the level of achievement was above the international average. In SISS 1984, Sweden also participated with a sample of students from grade 3, which had results slightly below the international mean. However, most other countries participated with samples from grades 4 or grade 5. Adjusting for the age- and grade-differences, the Swedish results are raised even further, and the adjusted results are close to those achieved by the Japan students, which was the country with the highest level of achievement in these studies.

After SISS 1984, it was not until TIMSS 2007 that Sweden participated in another study with Population I. Both the observed and adjusted results in this study show the Swedish results to be slightly below the international average. This indicates that there has been a decline in the level of achievement in science for students in the early grades of comprehensive school.

Sweden has participated in several science studies with samples of students from Population II. In FISS 1970 the observed result was well above the international mean. In SISS 1984 Sweden participated with samples both from grade 7 and grade 8. The latter group had a result slightly higher than the international mean, while the former was below the international mean. However, because of the lower number of school years for the Swedish students, the adjusted values were at the international top level, and close to those obtained by Japan.

In TIMSS 1995 Sweden participated, as has already been mentioned with samples from grade 7 (and grade 6), along with an unofficial sample from grade 8. The observed result for grade 7 was close to the international mean, while the observed result for grade 8 was close to the international top. Because the Swedish grade 8 students were older than the international mean, the adjusted value was marginally lower than the observed value. The adjusted value for the students in grade 7 was at the same high level as the observed value for grade 8.

Sweden participated in TIMSS 2003 and TIMSS 2007 with samples from grade 8. In both these studies substantial declines compared to 1995 were observed, and particularly so between 1995 and 2003. The adjusted values are even lower, because of the higher age of the Swedish students. As for mathematics, the decline implies that the Swedish grade 8 students in 2007 achieved worse in science than the grade 7 students did in 1995.

To summarize the findings for science, the results for the Swedish students in the lower grades were very good, except for the TIMSS 2007 study, and the results for the students in the higher grades were very good up too 1995, after which there has been a very considerable decline in the level of science performance.

Reading

Table 4 presents results from the studies of reading comprehension in which Sweden has participated.

Tabell 4. Observerade och adjusted results in reading comprehension for Sweden 1970 – 2007.

	Observed value	Difference age	Difference grade	Effect age	Effect grade	Adjusted value
Population I						
FIRS 1970	0.33	-0.16	-1.15	0.03	0.38	0.75
SIRS 1991	0.35	0.12	-0.48	-0.02	0.16	0.49
PIRLS 2001 åk 3	0.07	-0.56	-1.06	0.11	0.35	0.53
PIRLS 2001 åk 4	0.57	0.43	-0.06	-0.09	0.02	0.50
PIRLS 2006 åk 4	0.40	0.54	-0.07	-0.11	0.02	0.32
Population II						
FIRS 1970	-0.09	-0.08	-1.02	0	0.20	0.11
SIRS 1991	0.29	0.15	-0.38	0	0.08	0.37

Note. For Population I it has been assumed that the effect of a chronological year is 0.20, and that the effect of a school year is 0.33. For Population II the corresponding estimates are 0.0 and 0.20.

In the first two studies (FIRS 1970 and SIRS 1991) the Swedish students had very good results close to the international top for Population I. Adjusting the values to take into account the fact that the Swedish students had fewer years of schooling improves the results further.

In PIRLS 2001 Sweden participated with samples both from grade 3 and from grade 4. The latter sample had the best results among all participating countries, while the results for the grade 3 students were close to the international mean. The adjusted values for the two samples are at the same high level, which agrees with the results from the two earlier studies. The observed results in PIRLS 2006 indicate a decline, and Sweden has in this study lost the top position. The adjustment causes a further decline in the Swedish results.

In the first study of reading comprehension in Population II (FIRS 1970) the Swedish results were close to the mean, while the Swedish results in SIRS 1990 were well above the international average. The adjustments cause the Swedish results to be even better.

In short summary the studies show that the Swedish students in both Population I and II have had a very good reading comprehension compared to students in other countries from 1970 and onwards, even though there is a decline in the latest measurement for the Population I students.

Conclusions and Implications

According to the officially presented pattern of results for Sweden (e.g., National Agency for Education, 2004) the level of achievement in mathematics was poor in the first two studies, close to the international mean for the TIMSS 1995 study, and thereafter there has been a dramatic decline to a level just below the international mean. This somewhat strange description is due to the fact that the official Swedish results for TIMSS 1995 were based on the grade 7 sample, while the results for the grade 8 sample was only reported in an appendix to the international report (Beaton et al., 1996). In this way the quite good Swedish mathematics results in TIMSS 1995 were missed, and the decline for the TIMSS 2003 study appeared as difficult to reconcile with the fact that the level of achievement was still close to the international mean. However, this is because the change between 1995 and 2003 was determined through comparing the results for the 1995 grade 8 sample with the 2003 grade 8 sample (Mullis et al., 2004a).

For science, the observed results in the early studies were not particularly impressive, but according to the adjusted results, both Population I and II students performed at an international top level until the early 2000s.

For reading, the observed level of achievement was well above the international average up to the early 1990s, and after that time there has been a slight decline. However, the reassessment based on application of the adjustments for effects of age and grade differences show that the level of reading achievement was even higher than was shown by the observed results in the studies conducted before 2001.

Thus, the results show that the models used for defining the populations to be included in the studies systematically biased the Swedish estimates in a negative direction for the early IEA studies, that did not use a grade based population definition. The IEA studies conducted from 1995 and onwards have used a grade based population definition, and these studies yield a positive bias for Sweden. However, this bias is smaller because the somewhat higher age of the Swedish students does not have a large effect on the outcome.

The underestimation does of course not only apply to the Sweden, but to other countries with a late school start as well (e.g., Denmark, Norway, Finland, Latvia, and Lithuania?). However, in the current study no adjustments have been made for any other country than for Sweden. This, can be regarded as a methodological problem, given that the Swedish results have been determined in a relative fashion through comparison with the mean of the participating countries. However, since there would be adjustments both in an upward and a downward direction for the other countries, it is difficult to predict the amount of error introduced by applying the adjustment just for Sweden. This is a topic to be investigated in further research.

It should also be noted that the adjustments made here can only be meaningfully applied when there is a strict age-based decision rule for school start. If, for example, students do not start school because of low performance it makes no sense to apply the kind of adjustment used here. The same is true if students are retained in a lower grade because of poor performance.

The results presented here must thus be interpreted with caution, because they are counterfactual estimates, determined on the basis of a model which is based on a set of assumptions which may hold more or less true. Thus one assumption is that the estimates of the relative effect of a school year and a chronological year which have been established in previous research can be generalized to apply to all the studies reanalyzed here. This assumption is not likely to hold perfectly true, but the results should nevertheless be robust against a certain variation in the size of the estimated effects.

Another assumption is that it is only after formal school start that any schooling effects start to appear. However, it could be objected against the corrections that even though the children in England, for example, start school two years before the Swedish children, a large part of the Swedish children take part in other organized educational activities. Thus, most Swedish children attend pre-school from early age, which has as one aim to prepare children for school start. Only if pre-school does not develop the knowledge and skill in mathematics, science and reading is it reasonable to adjust for the smaller number of school years of the Swedish children. It is, however, reasonable to assume that the curriculum-bound knowledge and skills that to a large extent is focused upon in the international assessment primarily is developed through formal schooling. At the same time there is evidence to suggest that pre-school attendance has a positive effect on school achievement (e.g., Sylva, xxxx), which also should influence the results on the international assessments favorably. However, given that we currently know less about the effects of different types of preparations for schooling than about the effects of schooling it is difficult to assess the importance of violations of the assumption that the alternative activities do not have effect.

It may be noted, however, that the adjustments of the results when Sweden has participated with more than one grade, generally leads to the reasonable result that the level of achievement is the same, which provides support for the model applied to make the adjustments. It should also be emphasized that the fact that Sweden participated with an unofficial grade 8 sample in TIMSS 1995 makes it possible to conclude that the very good Swedish results in mathematics and science in the early to mid 1990s not only appear in the adjusted estimates, but in observed results as well.

The results of the current study also raise substantive and policy-related questions. One main question that needs to be answered is why there has been such a dramatic decline of the level of achievement of the Swedish students, from an international top level in the early 1990s to a level around the international mean 15 years later? This issue will not be discussed here, but it may be noted that Sweden in the early 1990s made radical changes to its educational system, transforming it from one of the most centralized and regulated systems in the world, to one of the most decentralized and deregulated educational systems in the world. There is reason to believe that one or more of the changes, singly or in combination, caused the decline but no definite explanation is yet available (National Agency for Education, 2009). It may also be noted that the radical change of the educational system was given impetus by the impression that the results of the Swedish were not as good as may be expected, given the high level of cost of the comprehensive school. This impression was largely based on the mediocre observed results achieved in the IEA studies focusing on mathematics and science in the 1980s.

References

Beaton, A. et al (1996a). *Mathematics Achievement in the Middle School Years. IEA's Third International Mathematics and Science Stud*y. Boston College, Chestnut Hill, MA, USA.

Beaton, A. et al (1996b). *Science Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study*. Boston College, Chestnut Hill, MA, USA.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development, 60,* 1239−1249.

Cliffordson, C. (in press). Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design applied to Swedish TIMSS 95 data. *Educational Research and Evaluation*.

Cliffordson, C., & Gustafsson, J.-E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence, 36,* 143-152.

Coomber, & Keeves, J. (1973). *Science education in nineteen countries*. Stockholm: Almqvist & Wiksell.

Elley, W. B. (Ed.). (1994). *The IEA study of reading literacy: Achievement and instruction in thirty-two school systems*. Oxford, England: Pergamon.

Gustafsson, J.-E. (2009a). Strukturell ekvationsmodellering. I G. Djurfeldt & M. Barmark (Red.) *Statistisk verktygslåda II.* Lund: Studentlitteratur.

Husén, T. (1967a). *International Study of Achievement in Mathematics*, Vol I. New York: John Wiley.

Husén, T. (1967b). *International Study of Achievement in Mathematics*, Vol II. New York: John Wiley.

Keeves, J. P. (Ed.) (1992). *The IEA study of science III: Changes in science education and achievement 1970 to 1984.* Oxford: Pergamon Press.

Keeves, J. P. & Schleicher, A. (1992). Changes in Science Achievement: 1970– 1984. I: J P Keeves (Ed.). *The IEA study of science III: Changes in science education and achievement 1970 to 1984.*

Oxford: Pergamon Press.
Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: Regression-discontinuity applied to TIMSS-95. *Oxford Review of Education*, *32*, 397–429.
Luyten, H., & Veldkamp, (2008). Assessing the effect of schooling with cross-sectional data. Between grades differences addressed as a selection-bias problem. Paper presented at the 3rd IEA International Research Conference (IRC 2008). Taipei City, Taiwan, September, 2008.
Luyten, H., Peschar, J., & Coe, R. (2008). Effects of Schooling on Reading Performance, Reading Engagement, and Reading Activities of 15-Year-Olds in England. *American Educational Research Journal*, *45(2), 319-342.*
Martin, M. O., Mullis, I. V. S., & Foy, P. (with Olson, J.F., Erberber, E., Preuschoff, C., & Galia, J.). (2008). *TIMSS 2007 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools*, Chestnut Hill, MA: Boston College.
Mullis, I. V. S., Martin, M. O., Gonzalez, E. O, Chrostowski, S. J. (2004a). *TIMSS 2003 International Mathematics Report.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
Mullis, I. V. S., Martin, M. O., Gonzalez, E. O, Chrostowski, S. J. (2004b). *TIMSS 2003 International Science Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P (2007). *IEA's Progress in International Reading Literacy Study in Primary School in 40 Countries.* Chestnut Hill, MA: Boston College.
Mullis, I. V. S., Martin, M. O., Foy, P., Olson, J. F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J. (2008). *TIMSS 2007 International Mathematics Report. Findings from IEA's Trends in Mathematics and Science Study at the Fourth and Eighth Grades.* Chestnut Hill, MA: Boston College 2008
Postlethwaite, N., & Wiley, D. E. (Eds.) (1991). *Science education in twenty-three countries*. Oxford: Pergamon Press.
Robitaille, D. & Garden, R.A. (1989). *The IEA Study of Mathematics II: Contexts and Outcomes of School mathematics*. Oxford: Pergamon Press.
Rosén, M., & Gustafsson, J.-E. (2006). Utvecklingen av läsförståelse mellan åk 3 och åk 4. I L. Bjar (Red.) *Det hänger på språket! Lärande och språkutveckling i grundskolan*. Lund: Studentlitteratur
Skolverket (2004). Internationella studier under 40 år. Svenska resultat och erfarenheter. Skolverkets aktuella analyser 2004. Stockholm: Skolverket.
Skolverket (2009). Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer. Stockholm: Skolverket.
Thorndike, R. L. (1973). *Reading Comprehension Education in Fifteen Countries*. Stockholm: Almqvist & Wiksell.