Northern Lights on PISA

UNITY AND DIVERSITY IN THE NORDIC COUNTRIES IN PISA 2000

Svein Lie • Pirjo Linnakylä • Astrid Roe

Editors

Programme for
International Student Assessment

DEPARTMENT OF TEACHER EDUCATION AND SCHOOL DEVELOPMENT
UNIVERSITY OF OSLO, NORWAY

FOREWORD

In late August 2001, about three months before the first international PISA results were published, members from the national PISA groups in Denmark, Finland, Norway and Sweden met in Sigtuna outside Stockholm to discuss the possibility of publishing a Nordic PISA report. The purpose of such a report would be to shed light on the Nordic results in PISA 2000 to see what similarities and differences there were within the Nordic countries and between the Nordic and other countries. Three editors were appointed, and author groups based on special fields of interests were formed. The following chapters are written by researchers who are all connected to the PISA project groups in Denmark, Finland, Norway or Sweden. We regret that we did not manage to include a contribution from Iceland.

We aimed to have the report published within a year. However, things always take more time than expected. And more importantly, both the scope of the report and the ambitions of the contributors increased as time went on. Now at last the report is finished and as editors we want to thank the authors for their contributions and for their patience. We hope that this book will be a source of interest and inspiration for teachers, teacher educators and policy makers, as well as for education researchers.

Oslo and Jyväskylä, May 2003

Svein Lie *Pirjo Linnakylä* *Astrid Roe*

List of authors

Peter Allerup, Department of Educational Psychology, The Danish University of Education, Copenhagen, Denmark (nimmo@dpu.dk)

Marit Kjærnsli, Department for Teacher Education and School Development, University of Oslo, Norway (marit.kjarnsli@ils.uio.no)

Kaisa Leino, Institute for Educational Research, University of Jyväskylä, Finland (kaisa.leino@ktl.jyu.fi)

Svein Lie, Department for Teacher Education and School Development, University of Oslo, Norway (svein.lie@ils.uio.no)

Pirjo Linnakylä, Institute for Educational Research, University of Jyväskylä, Finland (linnakyl@jyu.fi)

Antero Malin, Institute for Educational Research, University of Jyväskylä, Finland (antero.malin@ktl.jyu.fi)

Jan Mejding, Department of Educational Psychology, The Danish University of Education. Copenhagen, Denmark (jm@dpu.dk)

Bengt-Olov Molander, Department of Curriculum Studies and Communication, Stockholm Institute of Education, Stockholm, Sweden (bengt-olov.molander@lhs.se)

Torben Pilegaard Jensen, AKF, Institute of Local Government Studies, Copenhagen, Denmark (TPJ@akf.dk)

Erik Knain, Department for Teacher Education and School Development, University of Oslo, Norway (erik.knain@ils.uio.no)

Rolf V. Olsen, Department for Teacher Education and School Development, University of Oslo, Norway (r.v.olsen@ils.uio.no)

Astrid Pettersson, Stockholm Institute of Education, Stockholm, Sweden (Astrid.pettersson@lhs.se)

Astrid Roe, Department for Teacher Education and School Development, University of Oslo, Norway (astrid.roe@ils.uio.no)

Karin Taube, Department of Education, The Mid Sweden University. Härnösand, Sweden (karin.taube@mh.se)

Are Turmo, Department for Teacher Education and School Development, University of Oslo, Norway (are.turmo@ils.uio.no)

Jouni Välijärvi, Institute for Educational Research, University of Jyväskylä, Finland (jouni.valijarvi@ktl.jyu.fi)

TABLE OF CONTENTS

1 NORTHERN LIGHTS ON PISA

Svein Lie, Pirjo Linnakylä, and Astrid Roe

1.1 What is PISA 2000?

The OECD Programme for International Student Assessment (PISA) represents a new commitment by the governments of OECD countries to monitor the outcomes of education systems in terms of student achievement *in reading literacy, mathematical literacy and scientific literacy*. The three domains represent knowledge and skills that are regarded as relevant for adult life. Denmark, Finland, Iceland, Norway and Sweden participated in the first round of the survey, which took place in 32 countries during 2000, as a collaboration between the governments of the participating countries and the OECD. Between 4000 and 10 000 students were surveyed in each country. In 2000 the primary focus was on reading literacy. The survey will be repeated every three years, with the primary focus shifting to mathematics in 2003, science in 2006 and back to reading in 2009.

The framework and design of the PISA study is reported in two international publications (OECD 1999, 2000). The PISA concept of literacy is much wider than the concept commonly associated with the term *literacy*; the ability to read and write. It is regarded as a range of competencies, and the three PISA domains of literacy emphasise the ability to undertake a number of fundamental processes in a wide range of situations. To do this one needs to understand some basic ideas and principles and be able to use them adequately in different situations in everyday life. Thus, a broad understanding of key concepts is thought to be equally as important as the possession of specific knowledge. The assessment of cross-curriculum competencies is an integral part of PISA.

PISA 2000 was a paper and pencil test of 7 hours, 2 hours for each student. The test items were a mixture of multiple-choice items and tasks requiring students to construct their own responses. The items were organised in units based on a text passage setting out a real-life situation. The students also answered a 30-minute background questionnaire, with questions about their home background, reading interests and attitudes towards school. School principals were given a 30-minute questionnaire asking about their schools.

The first international report with the main initial results was published in 2001 (OECD 2001), and a thematic report on reading was published in 2002 (OECD 2002).

1.2 Focus on the principle of equity

This report presents some comparative findings of the PISA 2000 assessment in the Nordic countries. The main focus of the report is on the results related to the principle of equity and how it is realised in student achievement in the light of PISA findings.

The principle of equity has a long tradition in the Nordic education system (Husén 1974). Providing all students with equal access to education and removing obstacles to learning, especially among students from a disadvantaged background, have been leading objectives in Nordic education policy. However, attaining high overall performance while, at the same time, evening out disparities in learning outcomes are key aims not only in the Nordic countries but in the other OECD countries as well (Husén 1989; OECD 2001).

The Nordic strategy for building up both high quality and equality in education has been based on constructing a publicly funded comprehensive school system without selecting, tracking or streaming students during their basic education until the age 16. Part of the strategy is to spread the school network so that pupils have a school near their homes whenever possible or if this is not feasible, e.g. in rural areas, to provide free transportation to more widely dispersed schools. Inclusion of special education and instructional efforts to minimise low achievement are also typical to Nordic educational systems.

An investigation of equal opportunities for all can be approached by comparing regions, or urban and rural areas, or students of various target groups, such as socio-economic, ethnic, linguistic or gender groups. In the history of provision of equal educational opportunities in the Nordic countries, geographic barriers presented the first challenge. Next, equality was demanded for different socio-economic groups, then for gender groups and lastly for immigrant students (Husén 1974; OECD 2001, 2002; Fredriksson 2002).

Husén (1974) defines provision of equal educational opportunities as a conservative interpretation of the principle of equity. According to Husén, a more liberal view places emphasis on the active removal of instructional and pedagogical obstacles for the most disadvantaged students and provision of special support for learners from weaker socio-economic backgrounds or with lower capabilities. A more radical interpretation, in turn, adds to the previous ideas the aim of reducing inequality of learning achievement (Välijärvi 1994). Although this radical interpretation has not gained much ground in the Nordic countries, they all have typically sought the same goal by providing special support for the weakest students, e.g. by means of diverse special education arrangements.

In the Nordic countries there was a fairly large consensus in the 1970s in favour of adopting a liberal interpretation of educational equity in the development of the school system. More conservative circles were afraid, though, that this would lead to forced uniformity and, consequently, to loss of

individuality, creativity and talent. On the other hand, among progressives there were concerns about the increased freedom and autonomy of schools and teachers, which might lead to differentiation among schools, with weaker student receiving little attention and the most talented students being favoured (e.g. Hirvi 1994).

In recent decades the aspiration for social justice through education, which is typical of the Nordic countries, has been accompanied by the quest for economic success and competitiveness (Rinne et al. 2000). In the past decade, too, neo-liberalistic traits have found their way into Nordic education policies, . National competitiveness has also been associated with competition between schools in terms of curricular ideas and profiles, pedagogical innovations, and learning achievements, even in the Nordic countries, where some countries have already published ranking lists for schools. The trend towards increased autonomy of schools and school-based profiles, establishment of private schools, curricular differentiation, and increased external funding, has also raised concerns about the deterioration of Nordic educational equity and social cohesion (Söderberg 2001).

In addition, increased possibilities for parents to influence selection of their children's school, as well as curricula and the operation of the schools, have been based on the idea of stronger 'customer' orientation (Brown 1990). The desire of parents to influence schools is getting more and more evident in all Nordic countries. In this respect, the choice of school, where possible, is the obvious means available to parents for exerting influence. In choosing a school, parents select both the quality of education and the social climate, as well as a specific profile, such as Montessori pedagogy, emphasis on music, sports, mathematics or languages or religious orientation. At the same time, this has meant abandoning the system of school districts and uniform curricula typical of the policy of equity, as well as decreasing the influence and expertise of the educational administration and teachers. Those arguing for stronger parental choice accentuate the importance of individuality and talent, but forget that not all parents have the same possibilities, or wealth, for such individual choices.

Generally speaking, parents with the highest level of education who live in big cities know best what choices there are, and also have the greatest interest in actually choosing a school for their children. These parents are also best prepared to make efforts to change things in school (Söderberg 2001). In addition, a family's social network may serve as a channel conveying information about best schools, course options, and educational opportunities. Parents with little education tend to have a more distant relationship with school issues, and their friends neither know much about these matters nor find them very important. The educational background, occupation, and related economic status of parents have a bearing on the resources which the family can invest to support their children's learning, such as literature, computers, hobbies, magazines, language courses, or private tuition. Hence, the economic, cultural and social capital of the family influences their children's learning in various ways, either as promoting or hindering factors (Bourdieu 1986).

Social and cultural reproduction has not, so far, been as manifest in the Nordic countries as it has been in some countries in Central Europe (OECD 2001). In the Nordic countries, the students' integrated education has been further supported by a well-established general library system which is accessible to those coming from disadvantaged homes. Various reading materials, music recordings, and Internet connections available in libraries have promoted, for example, reading and web literacy also among those young people who do not have an extensive home library or computer facilities, or whose parents do not subscribe to many papers and magazines.

In what ways has the Nordic education system managed to even out economic, cultural and social discrepancies? Have decentralisation, school autonomy, choice of school and a competitive orientation in education policy diminished equal opportunities to learn? Is there a trade-off between quality and equity? Answers to these questions are sought here primarily from the perspective of learning achievements, contrasting them with students' gender, socio-economic and cultural home background, as well as with differences between schools.

1.3 Some general results

1.3.1 Achievement in Nordic and other countries

In the following pages we will present some of the main results that will be focused on and further analysed in this report. The overall achievement results in each of the three domains are presented and displayed in figures 1.1-1.3. The cognitive scores are Rasch scale scores with an OECD mean of 500 and standard deviation of 100 score points. In the figures, the countries are ranked according to the mean scores. The spread of the distribution of cognitive scores is given for each country as a standard deviation (S.D.) in the third column. In addition, the 5^{th}, 25^{th}, 75^{th} and 95^{th} percentiles are marked in the bar graphs to the right. The dark area in the middle of the bar is the 95% confidence interval for the mean. These three figures will be the basis for discussion and reference throughout the report.

It should be kept in mind that what for simplicity is referred to in the tables as reading, mathematics and science, represents the three domains *reading literacy, mathematical literacy* and *scientific literacy* respectively, as these concepts have been defined in the PISA framework (OECD 1999, 2000).

Figure 1.1 Mean scores and distributions of reading literacy

Country	Mean	S.D.	Reading
Finland	546 (2,6)	89	
Canada	534 (1,6)	95	
New Zealand	529 (2,8)	108	
Australia	528 (3,5)	102	
Ireland	527 (3,2)	94	
Korea	525 (2,4)	70	
England	523 (2,6)	100	
Japan	522 (5,2)	86	
Sweden	516 (2,2)	92	
Austria	507 (2,4)	93	
Belgium	507 (3,6)	107	
Iceland	507 (1,5)	92	
Norway	505 (2,8)	104	
France	505 (2,7)	92	
United States	504 (7,1)	105	
Denmark	497 (2,4)	98	
Switzerland	494 (4,3)	102	
Spain	493 (2,7)	85	
Czech Republic	492 (2,4)	96	
Italy	487 (2,9)	91	
Germany	484 (2,5)	111	
Liechtenstein *	483 (4,1)	96	
Hungary	480 (4,0)	94	
Poland	479 (4,5)	100	
Greece	474 (5,0)	97	
Portugal	470 (4,5)	97	
Russian Federation *	462 (4,2)	92	
Latvia *	458 (5,3)	102	
Luxemburg	441 (1,6)	100	
Mexico	422 (3,3)	86	
Brazil *	396 (3,1)	86	

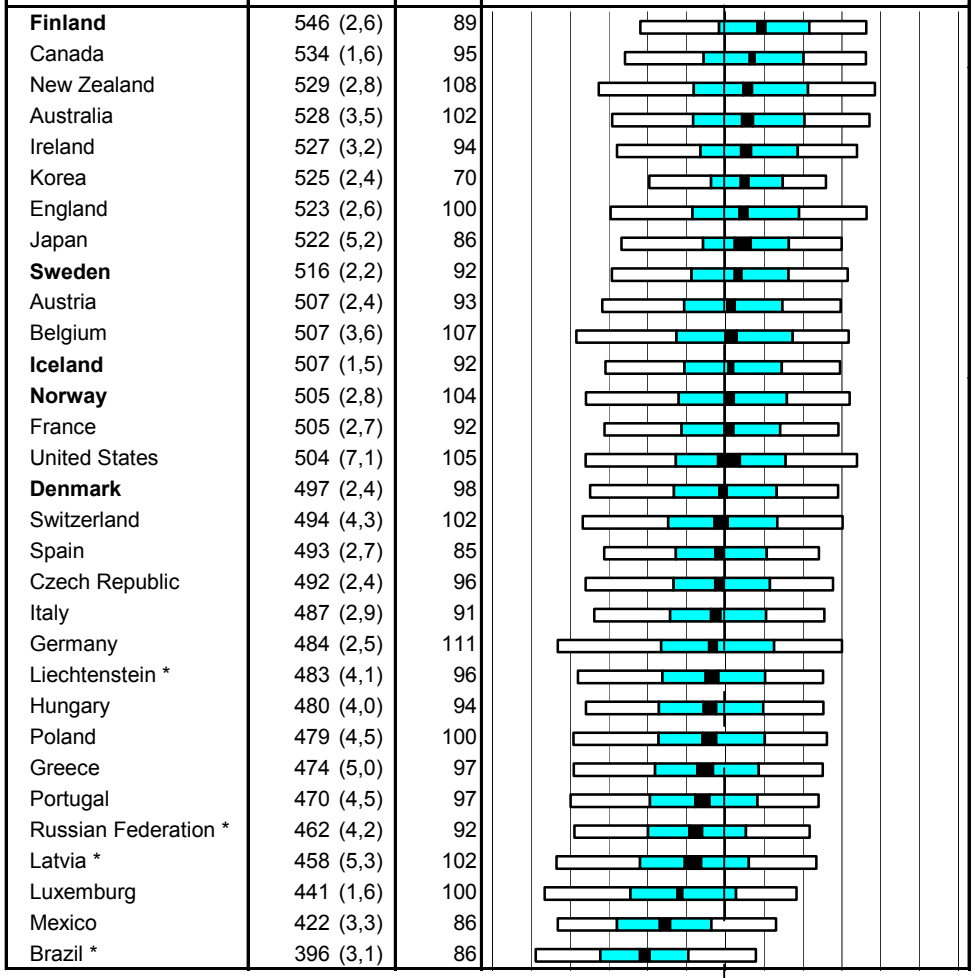


Figure 1.2 Mean scores and distributions of mathematical literacy

Country	Mean		S.D.	Mathematics
Japan	557	(5,5)	87	
Korea	547	(2,8)	84	
New Zealand	537	(3,1)	99	
Finland	536	(2,2)	80	
Australia	533	(3,5)	90	
Canada	533	(1,4)	85	
Switzerland	529	(4,4)	100	
England	529	(2,5)	92	
Belgium	520	(3,9)	106	
France	517	(2,7)	89	
Austria	515	(2,5)	92	
Denmark	514	(2,4)	87	
Iceland	514	(2,3)	85	
Liechtenstein *	514	(7,0)	96	
Sweden	510	(2,5)	93	
Ireland	503	(2,7)	84	
Norway	499	(2,8)	92	
Czech Republic	498	(2,8)	96	
United States	493	(7,6)	98	
Germany	490	(2,5)	103	
Hungary	488	(4,0)	98	
Russian Federation *	478	(5,5)	104	
Spain	476	(3,1)	91	
Poland	470	(5,5)	103	
Latvia *	463	(4,5)	103	
Italy	457	(2,9)	90	
Portugal	454	(4,1)	91	
Greece	447	(5,6)	108	
Luxemburg	446	(2,0)	93	
Mexico	387	(3,4)	83	
Brazil *	334	(3,7)	97	

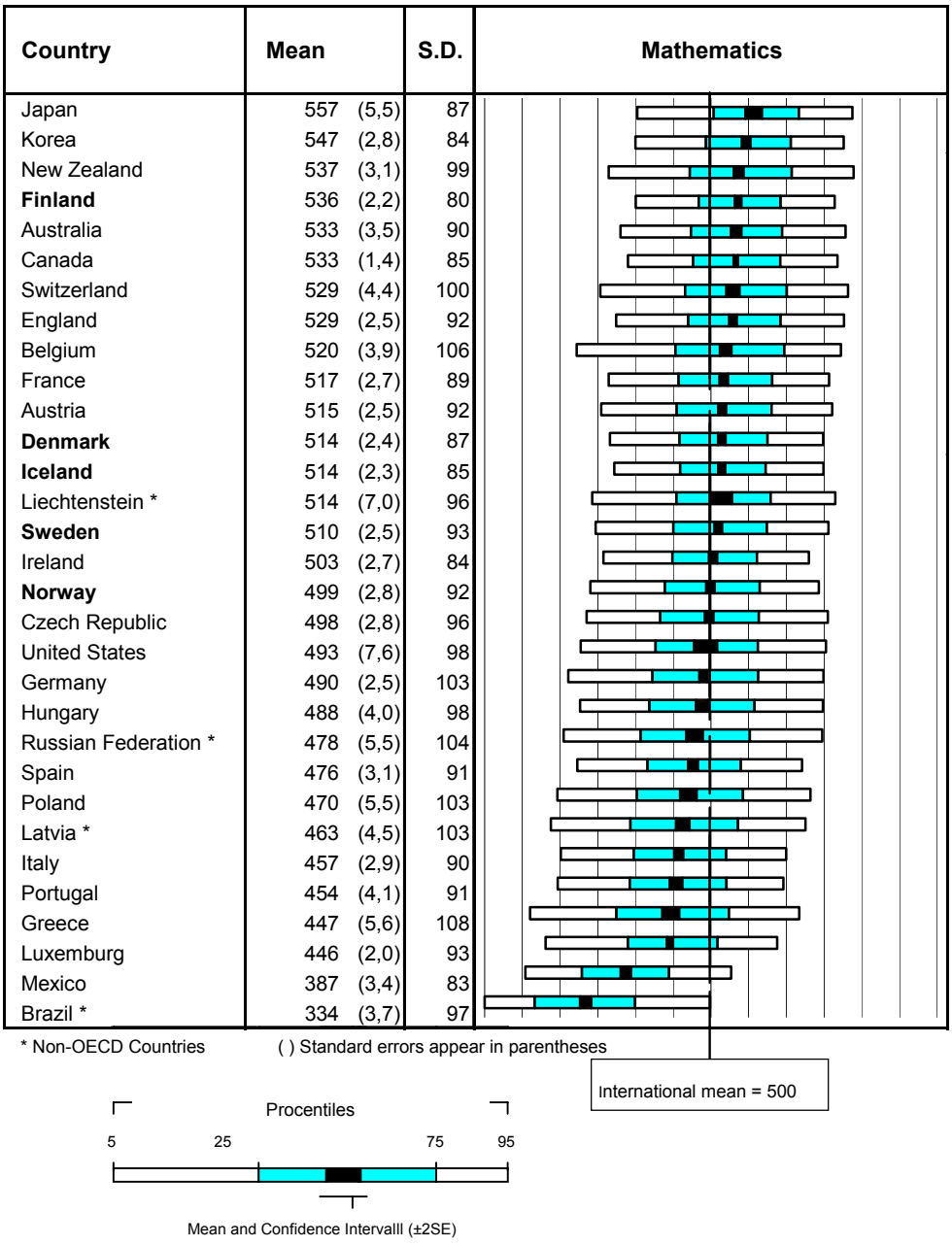


Figure 1.3 Mean scores and distributions of scientific literacy

Country	Mean	S.D.	Science
Korea	552 (2,7)	81	
Japan	550 (5,5)	90	
Finland	538 (2,5)	86	
England	532 (2,7)	98	
Canada	529 (1,6)	89	
New Zealand	528 (2,4)	101	
Australia	528 (3,5)	94	
Austria	519 (2,6)	91	
Ireland	513 (3,2)	92	
Sweden	512 (2,5)	93	
Czech Republic	511 (2,4)	94	
France	500 (3,2)	102	
Norway	500 (2,8)	96	
United States	499 (7,3)	101	
Hungary	496 (4,2)	103	
Iceland	496 (2,2)	88	
Belgium	496 (4,3)	111	
Switzerland	496 (4,4)	100	
Spain	491 (3,0)	95	
Germany	487 (2,4)	102	
Poland	483 (5,1)	97	
Denmark	481 (2,8)	103	
Italy	478 (3,1)	98	
Liechtenstein *	476 (7,1)	94	
Greece	461 (4,9)	97	
Russian Federation *	460 (4,7)	99	
Latvia *	460 (5,6)	98	
Portugal	459 (4,0)	89	
Luxemburg	443 (2,3)	96	
Mexico	422 (3,2)	77	
Brazil *	375 (3,3)	90	

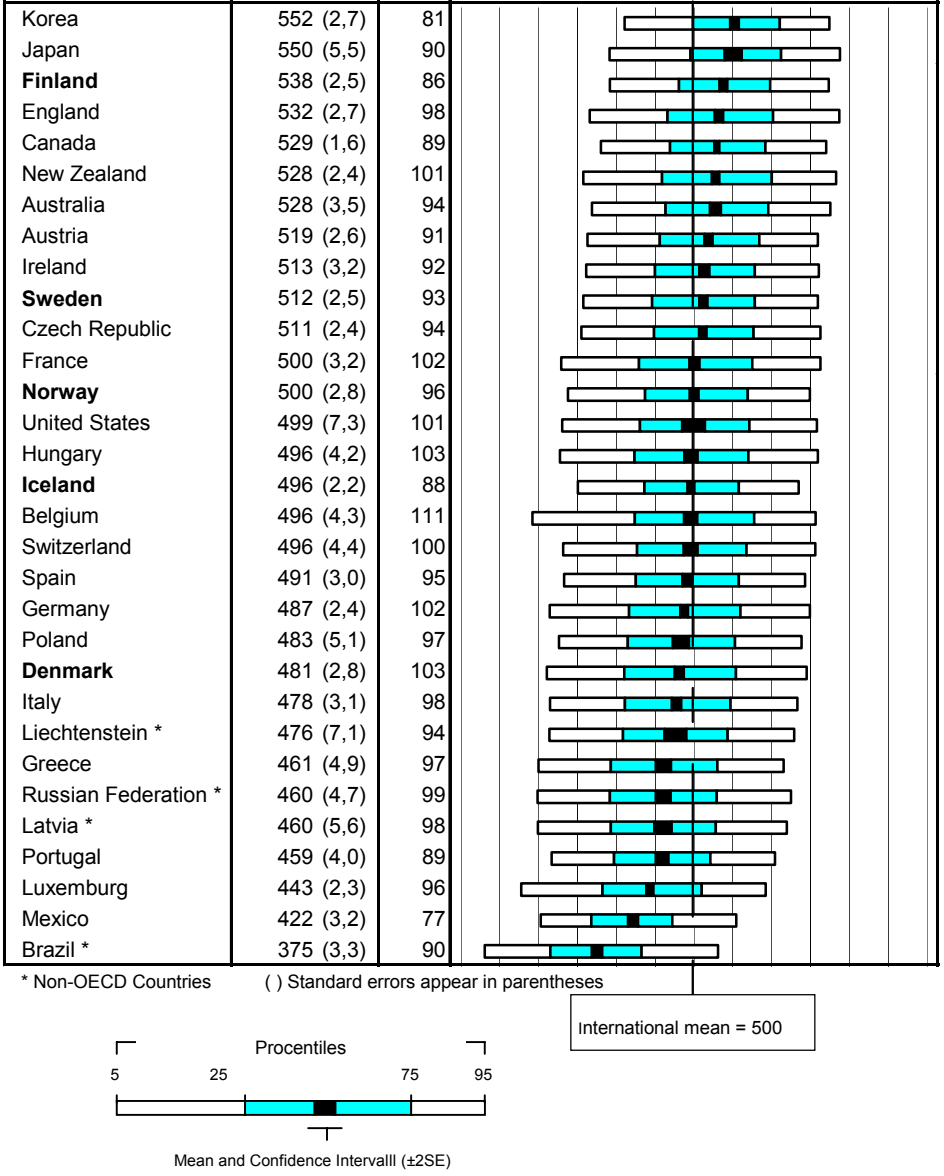
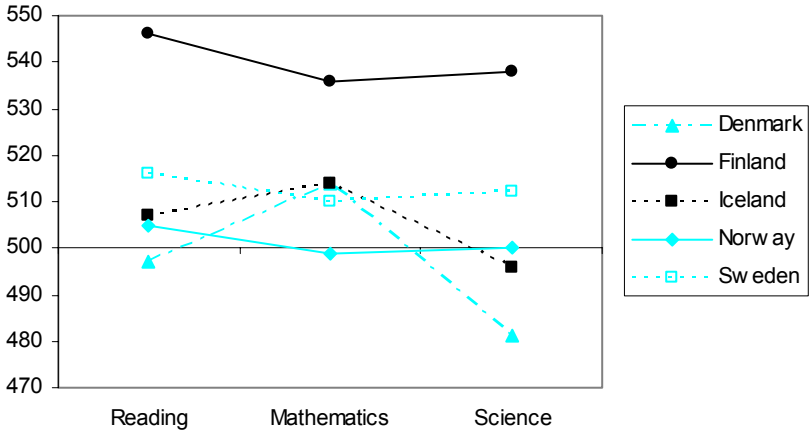
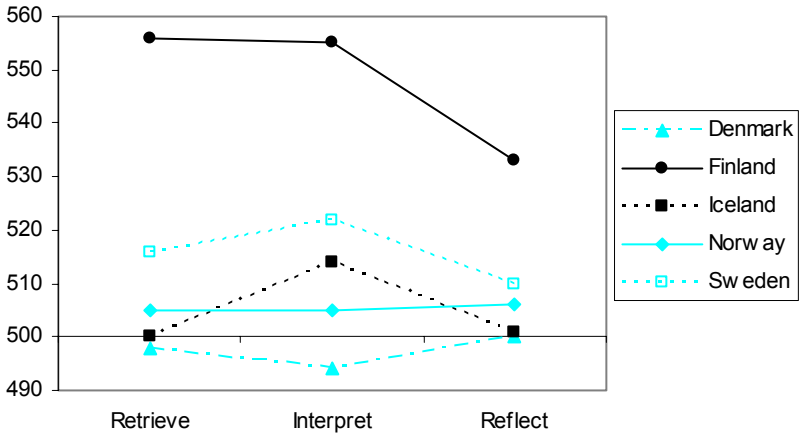


Figure 1.4 Mean scores in the three literacy domains for the Nordic countries



To further visualise the mean achievement for the Nordic countries in particular in the three domains, figure 1.4 presents an overview of the situation. Two features are immediately observed from this figure. First, the Finnish students score remarkably high and strongly outperform their Nordic peers. In fact, in reading the Finnish students significantly outperform students of all other participating countries, whereas the other Nordic students perform closer to the OECD mean. Secondly, the "profile" of Finland, Sweden and Norway is remarkably similar indicating higher performance in reading than in mathematics or science. This stands in contrast to students in Denmark and Iceland who have a pronounced relative strength in mathematics.

Figure 1.5 Mean scores on the three subscales for reading literacy



In PISA 2000 three subscales for reading literacy have been reported (for details, see OECD 2001, 2002), and figure 1.5 displays the mean achievement for the Nordic countries for each of the three: *the retrieving information*, the *developing interpretation* and the *reflection and evaluation* subscales. As can be seen from the figure, no distinct Nordic pattern emerges. Finland performs comparatively well in retrieving information and interpreting tasks; Iceland and Sweden perform best on the interpreting subscale; and there is virtually no difference in the aspect subscales in Denmark and Norway. A distinct feature, however, is the fact that the large gap between Finland and the other Nordic countries is less pronounced for the *reflection and evaluation* subscale than for the other two.

1.3.2 Equity measures

In the following we will discuss national standard deviations as well as differences between schools and between gender and socio-economic groups within each of the Nordic countries. From the columns marked S.D. in figures 1.1-1.3 it is seen that the standard deviations for the three achievement measures are remarkably small in Finland. Also in Iceland the spread of achievement is lower than in most other countries. On the other hand, the measures for Denmark and Norway are unexpectedly high for countries with no selection, tracking or streaming in the educational system up to the actual grade level.

An important focus in PISA is how cognitive score depends on various socio-economical variables. This theme will be discussed in some detail later in this report. Here we present the relationship in the Nordic countries between the students' reading literacy score and their International Socio-Economic Index (ISEI, for details, see OECD 2001, 221). This index is defined in PISA as a measure of the socio-economic status of the parents, based on information provided by the students on their parents' occupations, coded according to the International Standard of Classification of Occupations (ISCO). In table 1.1 we have shown this relationship in the form of the increased achievement scores associated with an increase in the ISEI index of one international standard deviation. Also shown in the table are the mean achievement scores and the standard deviations.

Table 1.1 demonstrates what can be regarded as ”good” and ”bad” results: Finland combines a high mean score in reading literacy with a low standard deviation and low dependence on the socio-economic index. The results for Iceland and Sweden are somewhat “poorer” in these respects, but still ”better” than for Denmark and Norway, which have the lowest mean scores and also the highest spread. In addition, these two countries have the strongest variation with ISEI, thus indicating that the school systems in these countries have not” succeeded in levelling out” differences in home background factors to the same degree (OECD 2001).

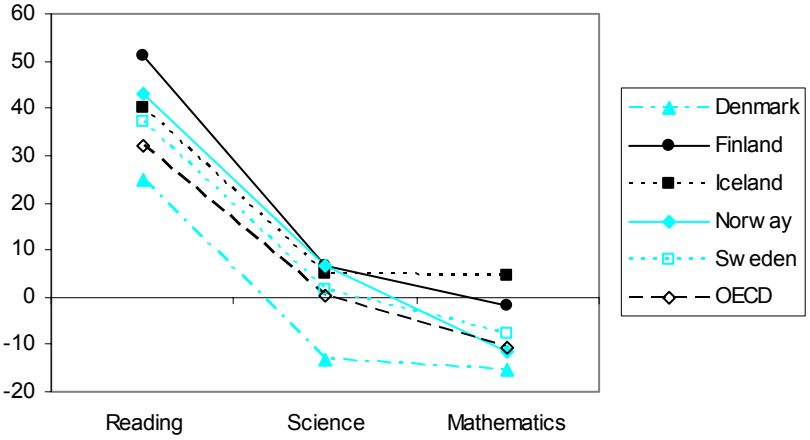
Table 1.1 Measures of reading literacy achievement: Mean scores, standard deviations and dependence on the ISEI index

	Mean	Standard deviation	Dependence on ISEI
Denmark	497	98	29
Finland	546	89	21
Iceland	507	92	19
Norway	505	104	30
Sweden	516	92	27
OECD mean	500	100	34

On the one hand we may choose to reject the statement that a high spread and strong dependence on home factors tell a story about lack of "success". After all, the dependence is definitely not a function of schooling alone; it is also one of the characteristics of the society, particularly how cultural and socio-economic factors are distributed among the population. On the other hand, the Nordic countries are basically similar in many respects, in particular by having rather homogeneous societies. Therefore, there are reasons to believe that the Danish and Norwegian educational systems are less "successful" than those of the neighbour countries in promoting equity with respect to home background factors. The relevance of this finding is significant for policy makers.

In figure 1.6 we have displayed the gender differences in the three content domains for the Nordic countries as well as for the OECD as a whole. A clear and similar pattern emerges from this figure. Most striking is the large gender gap in reading literacy in favour of girls, a finding that seems to be a common international feature. Nevertheless, in all Nordic countries except Denmark the gap is wider than in the OECD as a whole. Compared to the differences in reading, the gender differences in mathematics and science are much smaller, and are not always in favour of boys, as might be expected. Finally, by comparing gender differences across Nordic countries, a pattern emerges which shows relative female superiority clearly in Finland and partly in Iceland and Norway, while Denmark shows striking relative male advantages. One may wonder why the gender pattern is so different in the Nordic countries. One may also question specifically what lies behind the very large gender gap in reading, a finding that will be extensively explored in this report, particularly in the two next chapters.

Figure 1.6 Gender differences in score points within the three literacy domains
Positive differences are in favour of girls.



1.4 The present report

This report is the result of a co-operation between researchers in Denmark, Finland, Norway and Sweden. It will focus on the PISA results from a Nordic perspective, as the Nordic countries have many features in common, not only geographically, but also culturally, politically and in education systems. The report will present results from further analyses of the OECD PISA 2000 data, to throw light on differences and similarities between the Nordic countries. The aim is to go behind the ranking lists and try to impart a deeper understanding of some of the PISA findings, particularly those related to the principle of equity. All chapters have been written by Nordic researchers who have been involved in PISA 2000 in their countries.

In chapter 2 Astrid Roe and Karin Taube present reading results from a gender perspective. The PISA 2000 results show that there is a significant gender gap favouring girls in reading, more so, on average, in most of the Nordic countries than in the OECD. They try to find a gender specific pattern in reading achievement, and seek answers to questions like: What characterises reading tasks where boys are mostly outperformed by girls? Are there any reading tasks at all where boys actually outperform girls? And if so, what reading tasks are they? And finally, can the awareness of boys' strengths and weaknesses in reading performance help teachers to make boys better readers?

In chapter 3 Pirjo Linnakylä and Antero Malin investigate, using a two-level regression model, the possibilities of reducing the gender gap through examining the effect that various reading interests and activities have on the

students' reading literacy performance. They claim that if boys become as active and engaged in reading as girls, the gender gap in reading performance could be significantly reduced. Engaged readers provide themselves with self-generated learning opportunities that may be equivalent to several years of school education. Engagement in reading can compensate for low family income and poor educational background.

In chapter 4 Marit Kjærnsli, Astrid Pettersson and Are Turmo start out by discussing the definition of mathematical literacy in PISA. They then study the mathematics results from a Nordic perspective, focusing on different competence classes as defined by the mathematical framework in PISA. They also present detailed Nordic results from one of the units.

In chapter 5 Marit Kjærnsli and Bengt-Olov Molander focus on scientific literacy in PISA, a literacy that requires an understanding of scientific concepts as well as an ability to apply a scientific perspective as an intellectual skill. Compared to other international studies like to the IEA TIMSS, PISA has a much stronger emphasis on the science processes. In this chapter the results, and gender differences in particular, are discussed in relation to both processing skills and conceptual understanding.

In chapter 6 Kaisa Leino focuses on Nordic students' interest in and confidence and active engagement with the use of computers, as well as the relationship between using computers and reading literacy achievement. In the light of the PISA results she raises and answers the following questions: What effect does an active use of the Internet have on literacy skills? Are networks still a boys' playground as earlier studies have shown? What are the purposes teenagers use the Internet for? Are there differences in the use of computers between Nordic countries?

In chapter 7 Torben Pilegaard Jensen and Are Turmo use regression models to estimate the effects on reading literacy of selected social background variables. They demonstrate how the various socio-economic and cultural factors, one by one and also in combination, co-vary with reading achievement. Some remarkable differences between the Nordic countries are revealed in the results.

In chapter 8 Erik Knain and Are Turmo present results from the CCC (cross-curricular competencies) part of the PISA Student Questionnaire. CCC focuses on aspects considered important for lifelong learning. Firstly, a brief outline of the theoretical 'landscape' around CCC is presented. Secondly, descriptive statistics as well as correlation coefficients with score in reading literacy are presented and discussed.

In chapter 9 Rolf Vegar Olsen gives some results and reflections on classroom processes, based on both student and school questionnaire data. Results show that the differences between schools are relatively small in the Nordic countries. Central to our school systems have been equal opportunities to learn. The cluster of variables used for the analyses presented in the chapter is mainly related to how teachers' and students' behaviours in the schools affect learning - as perceived by the students and the principals respectively.

In chapter 10 Jouni Välijärvi and Antero Malin explore the PISA data to investigate how the socio-economic status affects literacy performance at individual and school level. By constructing a two-level model they show that the effect of the students' social background on their proficiency is divided into two components. There is the effect deriving from the whole school's social status, which can be interpreted as a 'bonus' the school brings to each student's performance. This effect proved to be clearly stronger in Denmark, Sweden and Norway than it was in Finland and Iceland. On the other hand, the social status of an individual student's family has a direct effect on the student's performance as well. Their results show that this direct effect was smallest in Iceland.

In chapter 11 Peter Allerup and Jan Mejding make a comparison between the PISA reading test and the IEA reading literacy study in 1991. By using test equating methods based on a test booklet that was only used in Denmark in addition to the nine PISA booklets, and which contained items from both PISA and IEA, they conclude that reading competence has generally decreased over the last decade in the OECD countries. This rather provoking result will hopefully inspire further studies on this important issue. It also highlights the importance of one of the goals of future phases of PISA, namely to measure trends in achievement.

In chapter 12 Svein Lie and Astrid Roe have explored similarities and differences in patterns of responses between the Nordic countries as well as between Nordic and all other participating countries. They have looked for characteristics of groups of countries with similar response patterns and tried to explain these characteristics in terms of linguistic and cultural factors.

In the final chapter the editors sum up some general findings and point to some further challenges for Nordic education.

References

Bourdieu, P. (1986). The forms of capital. In J. Richardson (ed.) *The handbook of theory and research for the sociology of education.* New York: Greenwood Press, p 241-258.

Brown, P. (1990). The third wave. Education and ideology of parentocracy. *British Journal of Sociology of Education* 11, p 65-84.

Fredriksson, U. (2002). *Reading skills among students of immigrant origin in Stockholm. Institute of International Education.* Stockholm University.

Hirvi, V. (1994). Vapauden ja tasa-arvon herkkä suhde [Sensitive relation of freedom and equity]. *Kasvatus* [Education] 25, 2, p 207-212.

Husén, T. (1974). *Learning society.* London: Methuen.

Husén, T. (1989). *Skolan i fokus*. Stockholm: Almqvist & Wiksell

OECD (1999). *Measuring Student Knowledge and Skills. A New Framework for Assessment.* Paris: OECD Publications.

OECD (2000). *Measuring Student Knowledge and Skills. The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy.* Paris: OECD Publications.

OECD (2001). *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD Publications.

OECD (2002). *Reading for Change. Performance and engagement across countries.* Paris: OECD Publications.

Rinne, R., Kivirauma, J., Aro, M. & Simola, H. (2000). *Liberal, conservative and Nordic. Opinions of the youth and the new educational policies of five post-industrial countries, in comparative perspective.*

Söderberg, S. (2001). Attitudes and expectations in relation to school: Swedish findings and some international comparisons. In *Schooling for tomorrow: What schools for the future.* Paris: OECD publications.

Välijärvi, J. (1994). Kaventuva tasa-arvo lisääntyvän vapauden hinta [Diminishing equality as a price for increased freedom] *Kasvatus* 25, 2, p 219-221.

2 READING ACHIEVEMENT AND GENDER DIFFERENCES

Astrid Roe and Karin Taube

2.1 Assessing reading literacy in PISA

Reading literacy in PISA is defined as more than just decoding written material or literal comprehension. It incorporates understanding and reflecting on texts and using written information to be able to function effectively in a knowledge-based society. The following definition of reading literacy is used in PISA:

> *"Reading literacy is understanding, using and reflecting on written texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society." (OECD 1999 p. 20)*

The PISA 2000 reading assessment employed 37 texts and 141 items[1], representing the kinds of reading literacy that 15-year-olds would require in the future. The selection of texts and construction of reading tasks is based on theories about reading as a dynamic process. Different readers have various ways of responding to a text, depending on their social and cultural background, their motivation and the context in which the reading takes place. Thus, many different text types with various contents are represented in PISA, and every item is based on a certain aspect of reading. The aim of this chapter is to present the reading results from a Nordic and particularly a gender perspective, with a focus on various text and item types.

2.2 Classification of texts and items

The texts in PISA are categorised by *structure, type* and *context* and the items are categorised by *aspect* and *format* (to be explained below). Table 2.1 shows the percentage distribution of text and item categories that will be presented in this chapter.

2.2.1 Texts

The main distinction between texts in the PISA assessment is between *continuous* and *non-continuous* texts. *Continuous* texts are formed of sentences

[1] Some few items were deleted either for all or for some countries. 127 items are represented in this presentation.

and arranged in paragraphs, and are supposed to be read from beginning to end. *Continuous* texts in PISA are defined as *argumentative*, *descriptive*, *expository*, *injunctive* and *narrative*. *Argumentative* texts are characterised by propositions, persuasions or arguments. They do not necessarily present facts, rather points of view. *Descriptive* and *expository* texts contain factual knowledge and information. They typically provide an answer to "what" or "how" questions. *Injunctive* or *instructive* texts provide directions on what to do or what behaviour is required in a certain situation. The reader has to understand the intentions of the rules or directions. *Narrative* texts are fictional texts, like fairytales, plays and short stories.

In PISA the *non-continuous* categories are *charts*, *forms*, *maps*, *schematics* and *tables.* They are not defined by content or intention, rather by structure, and the complexity of the structure is related to the reading strategy as well.

The texts in PISA can also be classified by *context*, meaning the use for which the text was written. Reading context includes reference to the people or objects that are connected with the reading situation. Thus, reading a letter will be associated with a *personal* situation while reading an advertisement will be associated with a *public* situation. Each text in PISA is defined by one of the four reading situations: *educational*, *occupational*, *personal* or *public*.

2.2.2 Items

The items in PISA are categorised by *aspect* and *format*. Each item was primarily defined by one of the following five aspects, each representing a certain way of reading and responding to a text: *retrieving information*, *forming a broad understanding, interpreting*, *reflecting on and evaluating the content* and *reflecting on and evaluating the form* (OECD 1999). Based on the pilot testing results, the five aspects were collapsed into the following three reporting scales, which will be used in this presentation:
Retrieve (retrieving information)
Interpret (forming a broad understanding + interpreting texts)
Reflect (reflecting on and evaluating the content + reflecting on and evaluating the form of a text).

There are five item formats:

- *Multiple choice* - four or five alternative answers are given, only one is correct.
- *Mixed multiple choice* - a series of statements is given. There are two alternative answers for each statement; for example: yes or no, true or false, included or not included.
- *Short response* - the item just requires a short answer without any further explanation.
- *Closed response* - the item defines the answer, for example a certain name or number which explicitly or implicitly occurs in the text.

- *Open constructed response* - the item requires a full sentence, often followed by an interpretation, explanation or evaluation of the form or content of the text.

The *Mixed multiple choice* category only includes five items, and in the following presentation *mixed multiple choice* and *multiple choice* will be collapsed into one *multiple choice* category. The *closed response* category only includes nine items. *Closed response* and *short response* will also be collapsed into one category: *short response*, as both require short responses. Table 2.1 shows the percentage of items within each text and item category.

Table 2.1 Classification of texts and items (within brackets the percentage of items within each text and item category)

Classification of texts	**Structure**	**Type**	**Reading context**
	Continuous (69) Non-continuous (31)	Argumentative (14) Descriptive (9) Expository (24) Injunctive (7) Narrative (14)	Educational (28) Occupational (15) Personal (21) Public (36)
		Charts/graphs (12) Forms (4) Maps (2) Schematics (4) Tables (10)	
Classification of items	**Aspect**	**Format**	
	Retrieving information (29) Interpreting (49) Reflecting (22)	Multiple choice (47) Short response (22) Open constructed response (31)	

2.3 The gender perspective

The gender perspective will be an essential part of this chapter. The main reason is that gender differences favouring girls seem to have increased radically in many countries over the last ten years. The IEA Reading Literacy Study in 1991 showed certain gender differences favouring girls in most countries, and the differences were larger for 9-year-olds than for 14-year-olds (Wagemaker 1996). A comparison of the IEA Reading Literacy study and the PISA study shows that gender differences were much larger in 2000 than they were in 1991, particularly in the Nordic countries, which all participated in both assessments (see table 2.2). The score points used in IEA and PISA are both based on a Rasch scale with 500 as the mean score and 100 as the international standard deviation.

Table 2.2 Gender differences in Rasch scale points favouring girls in 1991 and 2000

Country	IEA Reading Literacy 1991	OECD PISA 2000	Increase favouring girls
Denmark	4*	25	19
Finland	14	51	37
Iceland	13	40	27
Norway	4*	43	39
Sweden	15	37	22

*The difference is not statistically significant.

There are, however, certain problems comparing the two different assessments. One obvious problem is that the content and form of the reading material in PISA 2000 is different from the material used in 1991. Could the PISA material be more favourable to girls than the IEA material? Crawford and Chaffin (1986) suggest that there may be a gender factor in comprehension of texts. They claim that there is strong evidence that some kind of "gender schema" is likely to be activated in the process of linguistic comprehension. According to schema theory our experiences are summarized in an organizing structure, called a schema, which provides the framework necessary to make inferences and understand texts the way we do. Memory and comprehension are based on the same structures. What is recalled, according to schema theory, is not usually the actual sentences presented, but a reconstruction of what was understood. The schema often fills in information that is missing in the text. If our understanding is gender specific and if the reading material in PISA is more likely to activate feminine schemas than was the case with the IEA material in 1991, this could be one explanation of the seeming increase in gender differences. However, there is no evidence to support such a theory; in fact, the PISA material may very well be less favourable to girls than the IEA material.

A specific difference between the two assessments is that more than 30% of the tasks in PISA are open constructed response items, while there were no such items in the reported results from 1991. Open constructed items require writing skills and the ability to express oneself in written language. Motivation to write may also be of crucial importance. It could be that girls are better writers and/or are more motivated to do their best in a test situation than boys.

In an evaluation of writing competence Pajares and Valiante (1999) found that 13-15 year old girls to a larger extent than boys found it important to be able to write and that girls enjoy writing in learning contexts. Writing might be female-gendered and thus, boys, by avoiding it, miss opportunities to practice their writing competence. Another explanation for boys´ lower writing ability might be that writing is considered to require effort and involvement and that is why boys dissociates themselves from it (Jakobsson, 2000). Staberg (1992)

also found that girls think it is more important to write and enjoy writing in learning situations than boys do.

Boys may also try to manage with less than maximum effort, and general results from the PISA study show gender differences above the average for some of the longest texts - but also for some of the shortest ones.

Another difference between the two studies is that the PISA tasks emphasise reflective and evaluative reading more than the IEA tasks. Finally, the IEA Reading Literacy Study assessed 14-year-olds, while in PISA students were one year older. However, in the IEA Reading Literacy Study gender differences favouring girls were smaller for 14-year-olds than for 9-year-olds, which indicates that gender differences are likely to decrease when students grow older (Elley 1992). Consequently one should expect gender differences to be even smaller for 15-year-olds than for 14-year-olds.

2.4 Results by text and item format

In the first part of the following section we present the Nordic results and gender differences for various text- and item categories. Finally we focus on items that most prominently favour boys and girls. The focus in this report will not be on the three aspects, which are presented as *subscales* in the initial report (OECD 2001). All figures show average results for each Nordic country and average results for all 27 participating OECD countries as one group, which will be referred to as *the OECD average*.

The following general pattern (see chapter 1) will repeatedly emerge throughout the presentations, and will be referred to as the *normal* pattern. Any strong deviation from this pattern will be specifically commented upon.

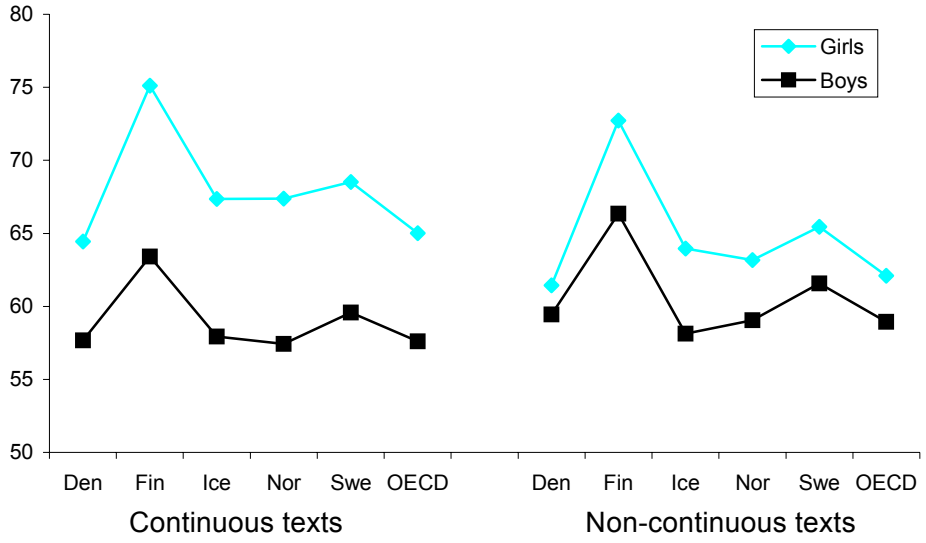
- Finish students generally outperform students in the Nordic countries as well as students in other OECD countries.
- Swedish students perform slightly better than Icelandic, Norwegian and Danish students.
- Icelandic, Norwegian and Danish students perform very near the OECD average.
- Gender differences in favour of girls are generally larger in Finland than in any other Nordic country as well as in the OECD, on average.
- Gender differences in favour of girls are generally larger in Norway, Iceland and Sweden than in the OECD, on average.
- Gender differences in favour of girls are generally smaller in Denmark than in the OECD, on average.

2.4.1 Text structure

Figure 2.1 shows that gender differences are generally less pronounced for items related to *non-continuous* than *continuous* texts in PISA. In the IEA study non-continuous and expository texts showed no significant gender

differences in any of the Nordic countries. Narrative texts showed significant gender differences in favour of girls for all countries, especially Finland, Iceland and Sweden (Taube & Munck 1996). Figure 2.1 below shows gender differences for items connected to different text structures in PISA.

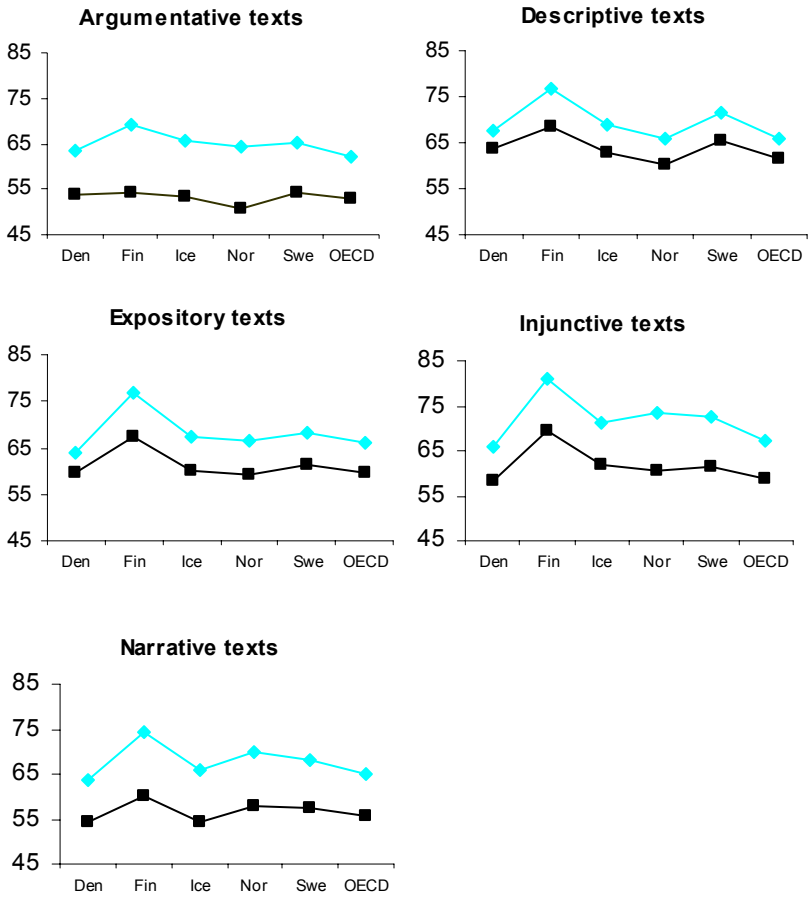
Figure 2.1 Mean results (percentage of correct answers) for girls and boys on items connected to continuous and non-continuous texts



2.4.2 Text type

Gender differences for each of the *non-continuous* text categories will not be given any special attention here. One reason is that gender differences for the *non-continuous* texts on the whole are comparably small; another reason is that the categories *maps*, *forms* and *schematics* each employ less than five percent of the total number of items (see table 2.1), which makes it difficult to generalise from these categories. Each of the *continuous* text types employ at least seven percent of the items, and gender differences for each *continuous* text type is presented in figure 2.2.

Figure 2.2 Mean results (percentage of fully credited answers) for girls and boys on items connected to the five continuous text categories. Girls are represented by the green line, boys by the black line



In all Nordic countries *argumentative*, *injunctive* and *narrative* texts favour girls to a greater extent than do *expository* and above all *descriptive* texts. However, gender differences are still smaller for *non-continuous* texts in general than for *descriptive* texts. This is all in accordance with the situation in the OECD. Both boys and girls seem to have most difficulties with the *argumentative* texts in PISA. Unexpectedly, Finnish boys do not outperform other boys' groups on *argumentative* texts.

How can the finding that gender differences are so much greater for *argumentative*, *injunctive* and *narrative* texts than for *descriptive* and *expository* texts be explained? The latter two carry factual knowledge and information, which may be easier for boys to handle than texts that contain propositions, persuasions or arguments, texts that give instructions and texts

that hold a lot of implicit information, metaphors and symbolic meaning. Furthermore, only 19% of the items connected to *argumentative*, *injunctive* and *narrative* texts require students to *retrieve information* (29% for all items), and 30% of these items are *reflecting* tasks (22% for all items). The general results show that boys are not outperformed to such an extent on the *retrieve* scale as they are on the *reflect* scale (OECD 2001). Thus, it is a moot point whether it is the *aspect* of the item or the *text type* that is the primary reason for the variations in gender differences. In the IEA Reading Literacy Study, in both populations, girls were generally most favoured by *narratives*. There was, as mentioned in 2.4.1 above, no significant gender difference in tasks connected to *expository texts* in any of the Nordic countries among 14-year-olds (Taube and Munck 1996).

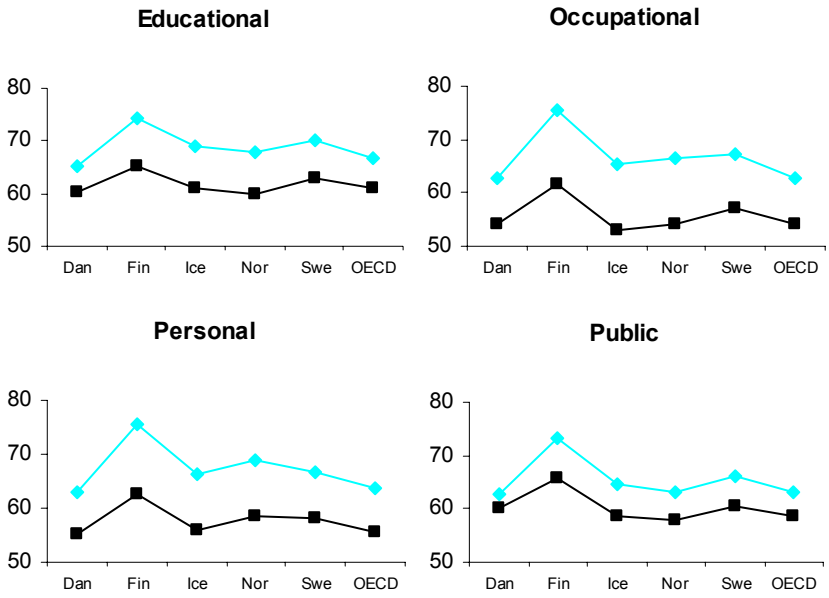
The general PISA results show smaller gender differences for the non-continuous texts than for continuous texts. Most non-continuous texts contain fewer words and sentences than continuous texts. On the other hand, they may contain advanced graphs and figures, and the tasks may demand an understanding of complex information. In the IEA Reading Literacy Study, maps, charts, etc. were labelled as *documents* and are thus equivalent to what is called non-continuous texts in PISA. The IEA results showed no significant gender differences for *documents*. It should be mentioned that the very few items where boys performed significantly better than girls were in most cases presented in connection with a map, chart or a table (Taube & Munck 1996). Thus, both the PISA and the IEA study have shown that documents/non-continuous texts are somewhat easier for boys than other kinds of texts.

2.4.3 Reading context

As seen in figure 2.3 the results from girls' and boys' achievements related to different reading contexts show a certain variation between the Nordic countries. Gender differences are generally largest for texts related to *occupational* situations. Ten of the 19 items related to *occupational* situations were also represented in the IALS in 1998 (International Adult Literacy Study), meaning they were initially meant for adult readers. If girls are more mature than boys at the age of 15, this may explain the fact that girls outperform boys to such an extent on these items. *Personal* reading situations also seem to favour girls noticeably. Most of the texts connected to *personal* situations are fictional texts, which traditionally favour girls (Taube & Munck 1996). The PISA results show that girls generally report a much higher frequency of reading fiction for pleasure than boys do (see Linnakylä & Malin in chapter 3). Gender differences are less pronounced for *educational* texts, which may be because boys and girls are exposed to the same *educational* texts at school. Boys are catching up with girls even more when it comes to *public* reading situations. It could be that boys are more oriented towards situations in the real world, and thus more attracted to *public* texts. This was also the tendency in the

IEA Reading Literacy Study in 1991 (Taube & Munck 1996). This issue is, however, a matter for further investigation.

Figure 2.3 Mean results (percentage of correct answers) for girls and boys on items connected to different reading contexts. Girls are represented by the green line, boys by the black line



2.4.4 Item formats

Figure 2.4 shows the gender differences for each of the three item formats. The Nordic countries are here collapsed into one group, and the figure shows the mean results for fully credited answers by percentage. It is obvious that for the Nordic countries as a group, as well as for the OECD average, gender differences are largest for *open constructed response* items and smallest for *multiple choice* items. Mean performance is generally highest in all Nordic countries for *multiple choice* items and lowest for *open constructed response* items. In the IEA study all items were multiple choice tasks, and there were no open constructed responses, which may explain part of the change in gender differences.

Figure 2.4 Mean performances for boys and girls on different item formats in OECD and in the Nordic countries

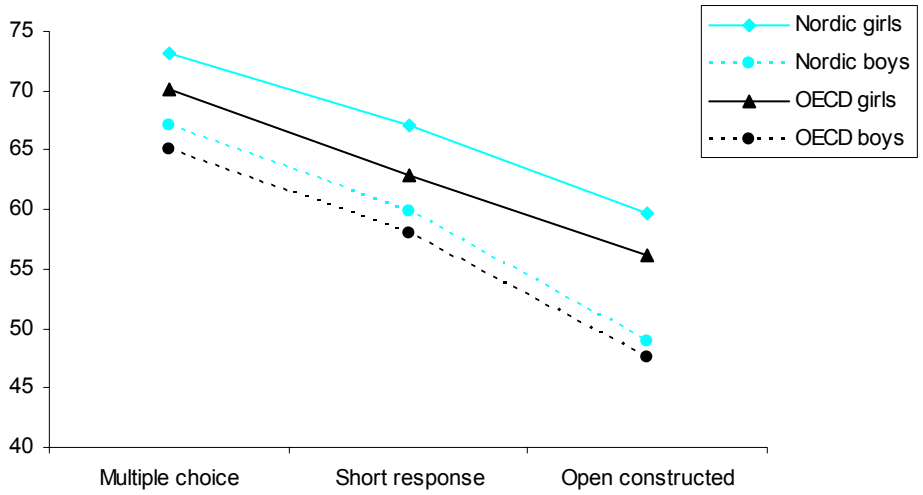
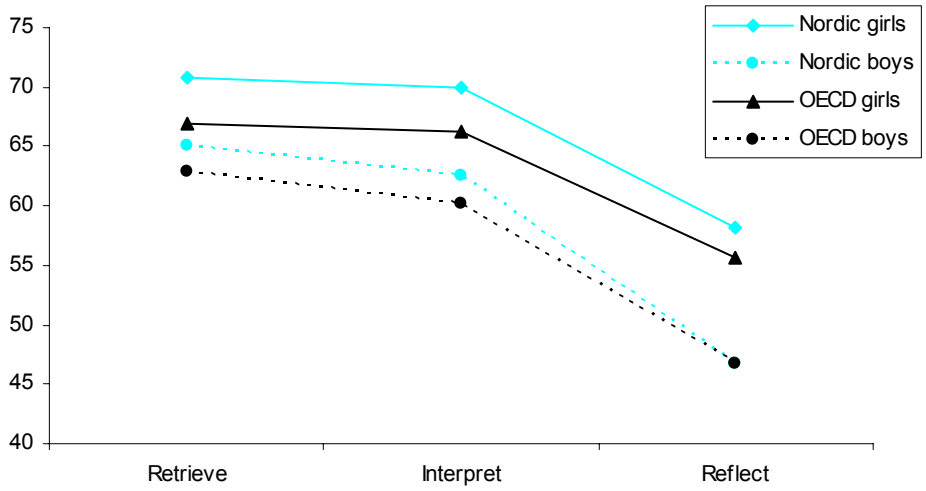


Figure 2.5 shows that mean performances are notably lower for both genders on *reflecting* tasks than on *retrieving* and *interpreting* tasks. Nordic boys, who as a group generally perform better than the other OECD boys, are at the same level as the OECD average when it comes to items that demand reflecting and evaluating the form or content of a text.

Figure 2.5 Mean performance in percentage for Nordic and OECD boys and girls on the three reading subscales



2.5 Gender differences at the item level

The general results have shown that girls outperform boys to a greater or lesser extent on all text and item categories in the Nordic countries as well as in the OECD. To get a more complete picture we need to ask the following questions: Are there any items at all that show gender differences in favour of boys, and if there are, what kind of items are they? And vice versa, what characterises items where gender differences in favour of girls are greatest? It would also be interesting to know if there are any noteworthy differences between the Nordic countries, or between the Nordic countries and the OECD regarding these item groups.

The method we used was simply to sort items by gender difference in percentage points. Statistical significance has not been calculated, as very small differences would prove to be statistically significant with such a large number of students. Furthermore, statistical significance will vary from country to country, and for the OECD average extremely small gender differences will prove to be statistically significant. In this case we have therefore decided that gender differences lower than 3 percentage points are too small to be given attention, and will hereafter be referred to as *no gender difference*.

Generally boys outperform girls on very few items in PISA. In Finland no single item shows a gender difference larger than 3 percentage points in favour of boys. In Iceland girls are outperformed by boys on three items, in Sweden on four items, in Norway on five items and in Denmark on eleven items. As the number of items is so low, we have chosen to search for characteristic features of the 25 items where girls are outperformed or that show no gender differences in the Nordic countries. These items will be named "Boys' items" here. We also present characteristics of the 25 items that show the greatest gender differences in favour of girls and name them "Girls' items". We will compare both categories with all the PISA items, which will be named "All items" in the following presentation.

2.5.1 Boys' items

Among the 25 Boys' items in each Nordic country 12 identical items are found in all countries. Most of the remaining 13 items are found among the 40 items that show the smallest gender differences in all the Nordic countries. A presentation and characterisation of Boys' items for each of the Nordic countries would be too detailed, and as the items to a large extent are identical in all Nordic countries, we have chosen to present a Nordic average here. The Nordic average will be compared with the OECD average.

Table 2.3 shows that in the Nordic countries grouped together as well as in all the OECD countries boys outperform girls on three items. If we look at what items they are in the two groups, we find that boys outperform girls on exactly the same three items in all OECD countries as in the Nordic countries. Two of these items are connected to *charts* and one is connected to a *map*. The

three tasks all require students to *retrieve information*. For 17 items gender differences are lower than 3 percentage points in the Nordic countries. In the OECD, gender differences are lower than that for 22 items. Figures 2.6 - 2.9 show the distribution of items within All items, Boys' items and Girls' items for different text and item categories.

Table 2.3 Number of items favouring girls, items favouring neither gender, and items favouring boys in OECD and in the Nordic countries

Items	OECD average		Nordic average	
	Number	Percentage	Number	Percentage
Gender dif. in favour of boys	3	2,4	3	2,4
No gender dif.	22	17,3	17	13,4
Gender dif. in favour of girls	102	80,3	107	84,2
Total	**127**	**100**	**127**	**100**

Figure 2.6 Percentage distribution of items by text structure for All items, Boys' items, and Girls' items in the Nordic countries

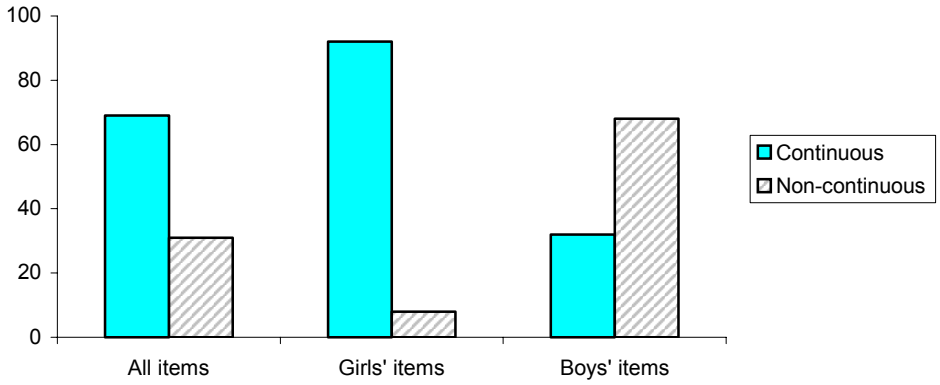
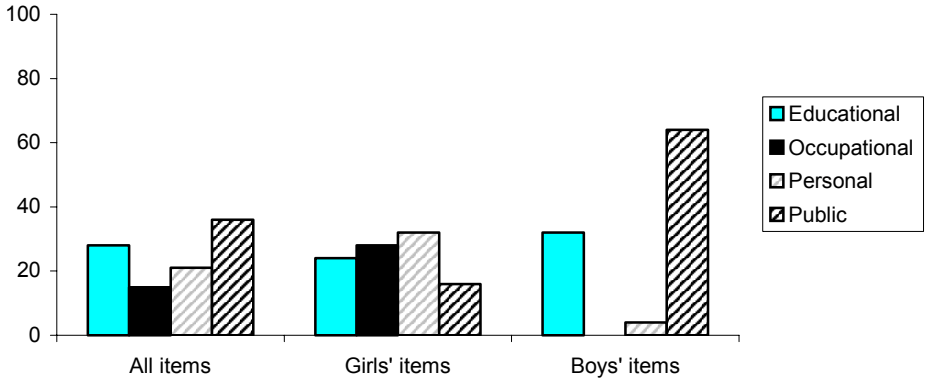


Figure 2.6 shows that the percentage distribution of items connected to *non-continuous* texts is clearly over-represented among the 25 Boys' items and correspondingly under-represented among the 25 Girls' items. A closer look at the *continuous* text types that are represented among Boys' and Girls' items reveals that among the Boys' 8 *continuous* text items 6 are connected to *expository* and 2 to *descriptive* texts. Among the 23 *continuous* Girls' items, 3 items are connected to *expository* texts, and no item is connected to *descriptive* texts. 20 items are connected to *argumentative*, *narrative* and *injunctive* texts.

Figure 2.7 Percentage distribution of items by reading context for All items, Boys' items, and Girls' items in the Nordic countries



One striking difference between Boys' and Girls' items in figure 2.7 is that Boys' items to a very large extent are connected to *educational* and *public* reading situations, while *personal* and *occupational* reading situations are slightly over-represented among Girls' items.

Figure 2.8 Percentage distribution of items by item aspect for All items, Boys' items, and Girls' items in the Nordic countries

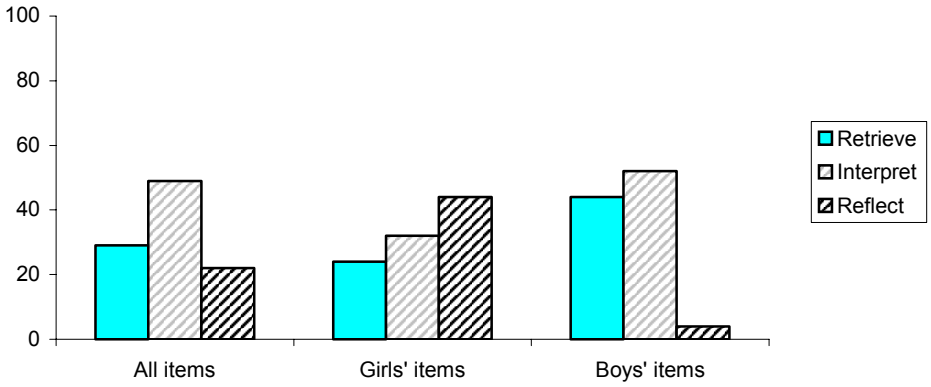
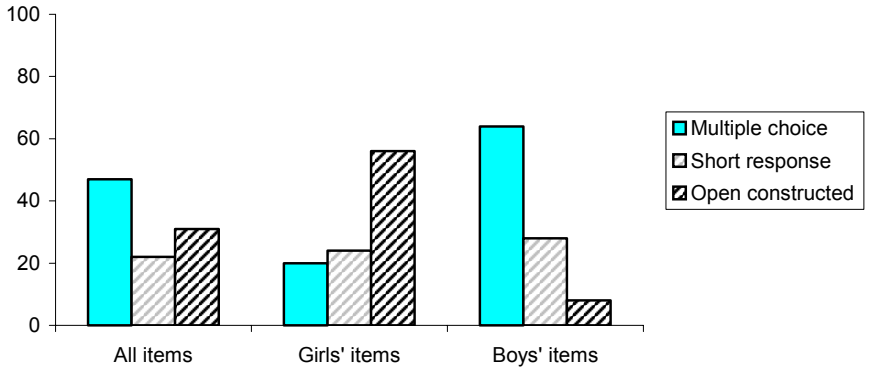


Figure 2.8 shows an under-representation of *reflect* items among Boys' items, whereas *reflect* items are over-represented among Girls' items. Furthermore, figure 2.9 shows a relatively large amount of multiple choice items among Boys' items, and a correspondingly high amount of open constructed items among Girls' items.

Figure 2.9 Percentage distribution of items by item format for All items, Boys' items, and Girls' items in the Nordic countries



Two texts are present in 8 of the 25 items among the Nordic Girls' items. These two texts only appear in the second half of the test booklets. This indicates that not only are reading ability and writing skills crucial for boys' performances, but effort and perseverance probably play an important part as well. According to Walkerdine (1989, 1998) hard work in school situations is strongly female-gendered and can be seen in contrast to rationality, which is strongly male-gendered.

An investigation of the 25 Boys' and Girls' items in all OECD countries shows a similar picture to the Nordic one, but generally gender differences are less pronounced in the OECD than in the Nordic average.

2.5.2 Are Boys' items easier than Girls' items?

One way of defining the difficulty of an item is by threshold values. In PISA a threshold value has been computed for each item on a sub-sample of 500 students per OECD country. The item threshold corresponds to the ability level which students have a probability of 0.5 of obtaining. The thresholds are in accordance with the PISA scale, which has a mean of 500 and a standard deviation of 100 for all OECD countries. The average threshold for Nordic Boys' items is 485, while the average threshold for Nordic Girls' items is 533. Thus one can conclude that the 25 Boys' items are a lot easier for all students

than the most typical Girls' items. However, thresholds only give information about how difficult an item is for all students, they do not reveal any gender difference or gender bias. The question is: Are Girls' items also generally more difficult for girls than Boys' items? One way of finding out is to look at the mean percentages of credited answers in both groups, for both genders. As shown in table 2.4, girls generally perform almost equally well on both kinds of items.

Table 2.4 Girls' and boys' percentage achievement on Boys' items and Girls' items in the Nordic countries

	Girls' mean results	Boys' mean results	Gender difference favouring girls
Boys' items	66,7	65,8	0,9
Girls' items	65,7	49,4	16,3

2.6 Summary and discussion

The texts and the items in PISA can be categorised in several ways; thus there are many possible ways of investigating gender differences. In this chapter we have given the results for boys and girls within most possible categories in PISA, trying to find patterns that can explain why girls perform so much better than boys in this assessment.

The results have, with very few exceptions, illustrated the pattern that was introduced early in this chapter and in chapter 1: Finnish students generally outperform all other students and gender differences are generally larger in Finland than in any other OECD country. Gender differences are generally larger than the OECD average in Norway, Iceland and Sweden, but smaller than the OECD average in Denmark.

A comparison with the results from the IEA Reading Literacy Study in 1991 suggests that gender differences favouring girls have increased strongly over the last ten years. In PISA the largest gender differences are found among items connected to *continuous* texts, especially *narrative*, *argumentative* and *injunctive* texts. Girls also seem to take advantage of *open response* items where they can express their understanding or reflection in their own words, which again points to the fact that writing skills may also be of some importance in the PISA assessment. Boys are not outperformed to such a large extent when it comes to *non-continuous* texts like *charts*, *maps* and *diagrams*. These texts have a relatively small amount of written information, but, on the other hand, they require an ability to understand and combine detailed information given in figures and tables. Items connected to *educational* and *public* reading situations are over-represented among the *non-continuous* texts, which could also explain why boys perform relatively well on such items.

Items connected to *personal* and *occupational* texts generally show large gender gaps favouring girls.

What should be the implications of these findings? Crawford and Chaffin (1986) suggest that there are gender-related factors that influence reading comprehension. If these factors make girls better readers and are biologically determined, one could be tempted to believe that there is not much to be done about it. Crawford and Chaffin, however, are careful to point out that gender schemata and the sense of masculinity and femininity they induce are not biologically determined. Rather, they are internalised social constructs that are difficult, but not impossible to change. According to Jakobsson (2000) language and related activities are more appreciated by girls and are seen as more female-oriented. Eccles (1987) found that high school girls are more positive towards English and boys to Mathematics (see also Skålvik, 2000; Lightbody et al. 1996). The engagement in learning might be less deep in areas which are considered to be more related to the opposite gender.

So what can schools, and above all, teachers do to narrow the gender gap in reading, apart from inspiring and encouraging boys to read more in their leisure time? Reading habits and interests play an important part, and will be thoroughly presented and discussed in the next chapter. The results presented in this chapter give strong indications that boys need to be more exposed to *continuous* texts that do not only contain factual knowledge with information to retrieve. It could be argued that interpreting, reflecting on and evaluating of texts require deeper understanding of the content of a text than retrieving information.

Boys definitely need to read more texts that require interpretation and reflection, for example *narrative* and *argumentative* texts. They also need to learn how various texts are constructed, how authors use the language to obtain a certain effect, how metaphors and implicit information are used and how to draw inferences from information given in various places in the text. They need to be given reading instruction to learn good reading strategies and to become conscious of what they read and for what purpose. Maybe all this might be accomplished if schools make a greater effort offer boys texts with a more male-gendered content.

References

Crawford, M. & Chaffin, R. (1986). The Reader's Construction of Meaning: Cognitive research on Gender and Comprehension. In Flynn E.& Schweikart P. P. (eds.). *Gender and reading. Essays on readers, texts and contexts.* Johns Hopkins University Press.

Eccles, J. (1987). Gender roles and achievements patterns: An expectancy value perspective. In J. M. Reinish, L.A. Rosenblum, & S.A. Sanders (Eds.)

Masculinity/femininity. Basic perspectives, 240-280. New York: Oxford University Press.

Elley, W. (1992). *How in the world do students read?* The Hague: International Association for the Evaluation of Educational Achievement.

Jakobsson, A-K (2000). *A gender perspective on motivation and learning: A study of students in theoretical programs in upper secondary school.* Doctoral dissertation. Gothenburg: University of Gothenburg.

Kimura, D. (1999). *Sex and cognition*. Cambridge Mass.: The MIT Press.

Lightbody, P., Siann, G., Stocks, R., & Walsh, D. (1996). Motivation and attribution at secondary school: The role of gender. *Educational Studies, 22*, 13-25.

OECD (1999). *Measuring students knowledge and skills. A new framework for assessment.* Paris: OECD Publications.

Pajares, F. & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology 24,* 390-405.

Skålvik, E.M. (2000). Selvoppfatning og motivasjon hos gutter og jenter. I G. Imsen (Ed.) *Kjønn og likestilling i grunnskolen,* 91-111. Oslo: Gyldendal Akademisk.

Staberg, E-M (1992). Olika världar, skilda värderingar. Hur flickor och pojkar möter högstadiets fysik, kemi och teknik. Umeå: University of Umeå.

Taube, K. & Munck, I. (1996). Gender differences at the item level. In Wagemaker, H. (ed.): *Are girls better readers? Gender differences in Reading Literacy in 32 Countries*. Amsterdam: International Association for the Evaluation of Educational Achievement.

Wagemaker, H. (1996). Part I – The Study. In Wagemaker (ed.): *Are girls better readers? Gender differences in Reading Literacy in 32 Countries*. Amsterdam: International Association for the Evaluation of Educational Achievement.

Walkerdine, V. (1989). Femininity as performance. *Oxford Review of Education, 15*, 267-279.

Walkerdine, V. (1998). *Counting girls out: Girls and mathematics* (new edition) (Studies in Mathematics Education Series: 8). London: Falmer press.

3 HOW TO REDUCE THE GENDER GAP IN READING LITERACY

Pirjo Linnakylä and Antero Malin

3.1 Revealing and reducing gender difference in reading literacy

The difference in educational outcomes between the gender groups has considerable significance in knowledge societies, which are increasingly concerned with the twin imperatives of quality and equity in education (Wagemaker 1993, 7). In reading, which is considered fundamental to life-long and life-wide learning as well as to knowledge-intensive work and everyday life, reducing the gender gap is a mission of great importance.

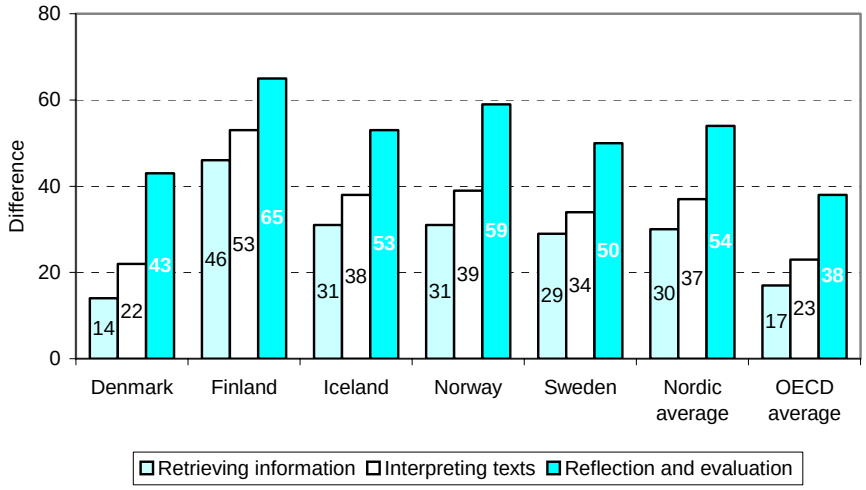
In the past, concerns about gender differences in educational outcomes have usually focused on the underachievement of girls. However, in the area of reading literacy girls have surpassed boys both in the IEA Reading Literacy Study (Elley 1994) and in the PISA study (OECD 2001). The initial PISA results showed a pattern of gender differences consistent across countries: in every country, on average, girls reached a higher level of performance than boys (OECD 2001, p.122; see also chapter 1). The initial PISA findings also assumed that by enhancing male students' interest and engagement in reading activities the gender gap in performance could be reduced (OECD 2001, p130).

In PISA, the gender gap in reading literacy performance was not only universal but also wide, much wider than in the IEA study in 1991 (Purves & Elley 1994; OECD 2001). However, there was considerable variation in gender differences between countries. Even in the Nordic countries, which are considered culturally relatively similar and have long emphasised the principle of equity in education, the gender differences between countries were significant. In Denmark, the gender gap was the smallest (25 points); in Finland it was about twice as large (51 points). The variation in gender differences suggests that the current situation is not inevitable. This gap can be closed or at least reduced (OECD 2001, p 125).

The initial PISA results also revealed that the gender gap in reading literacy differed in the three aspect domains. In all countries, the widest gender gap occurred on the reflection and evaluation scale (45 points), that is, on tasks requiring critical evaluation, argumentation and relating textual information to personal experiences, other texts, knowledge and ideas. The gap was narrower

on the interpretation and on the retrieving information scales (29 and 24 points respectively). A similar tendency was also found in the Nordic countries, where the gender difference on the reflection and evaluation scale was, on average, 54 points, on the interpretation scale 37 points and on the retrieving information scale 30 points (see figure 3.1).

Figure 3.1 Gender differences in the subscales of reading literacy

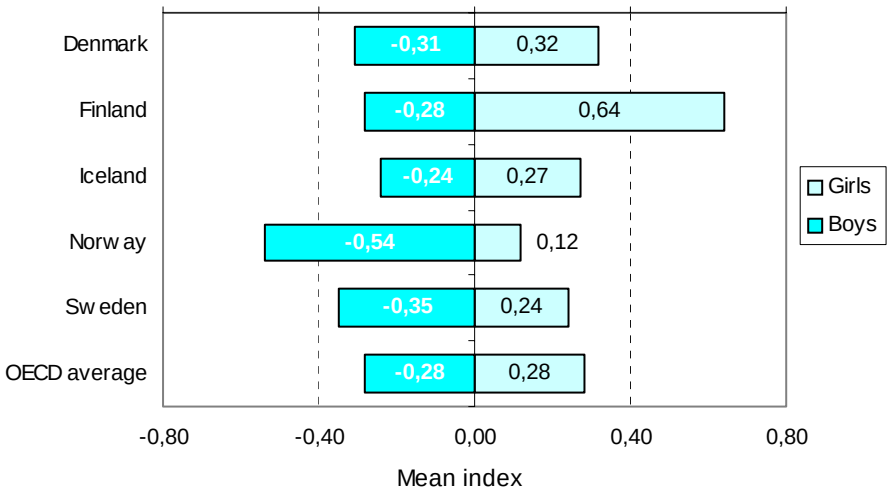


If we compare the PISA findings to the results of the IEA study, the widening gender gap may be related to domain specification: in PISA, reflection and evaluation played a more important role than they did in the IEA study, which focused more on searching information and developing interpretation.

3.2 Focus on how to even out reading interests and activities

The aim of this chapter is to investigate, in the light of the Nordic PISA data, possibilities for reducing the gender gap through examining the effect that various reading interests and activities have on students' reading literacy performance. Many previous studies have indicated that gender differences are strongly associated with motivation (Young 2000) and socio-cultural factors, especially factors such as reading interests and activities (Guthrie & Wigfield 2000; Kontogiannopoulou-Polydorides & Adamapoulou 1996; Purves & Elley 1994).

Figure 3.2 Engagement in reading by gender



(The mean of the index value for the OECD student population was 0 and the standard deviation was 1)

The initial results of PISA also suggested that voluntary reading activities and reading performance reinforce one another (OECD 2001, p129). The results pointed out, in particular, that gender differences in favour of females are associated with voluntary reading activities and reading materials. There appeared to be only limited engagement in reading among 15-year-old males beyond what is required of them at school. This was also typical of this group in the Nordic countries (see figure 3.2). The Norwegian boys, in particular, rarely engaged in voluntary reading. In all the Nordic countries, boys also reported spending much less time reading for enjoyment than girls did, as shown in figure 3.3. In Finland and Denmark the differences were widest in this respect. However, boys and girls differed in their reading activities not only with regard to their engagement or the time they spent reading, but also with regard to the materials they favoured. The largest difference could be detected in reading fiction (see figure 3.4). Girls were also more active than boys in borrowing books from the school or local library (OECD 2001, p 130-131.)

Figure 3.3 Students who read more than 30 minutes per day for enjoyment, by gender

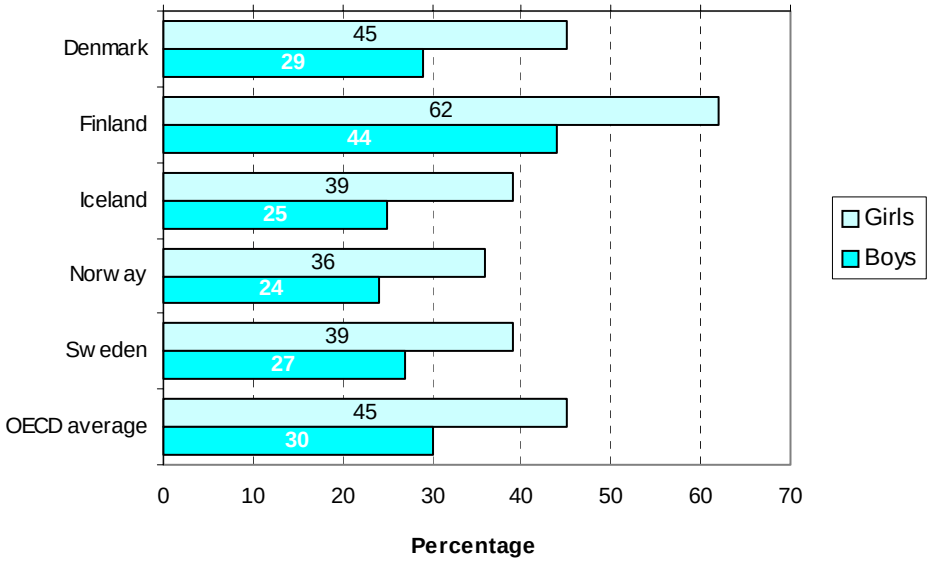
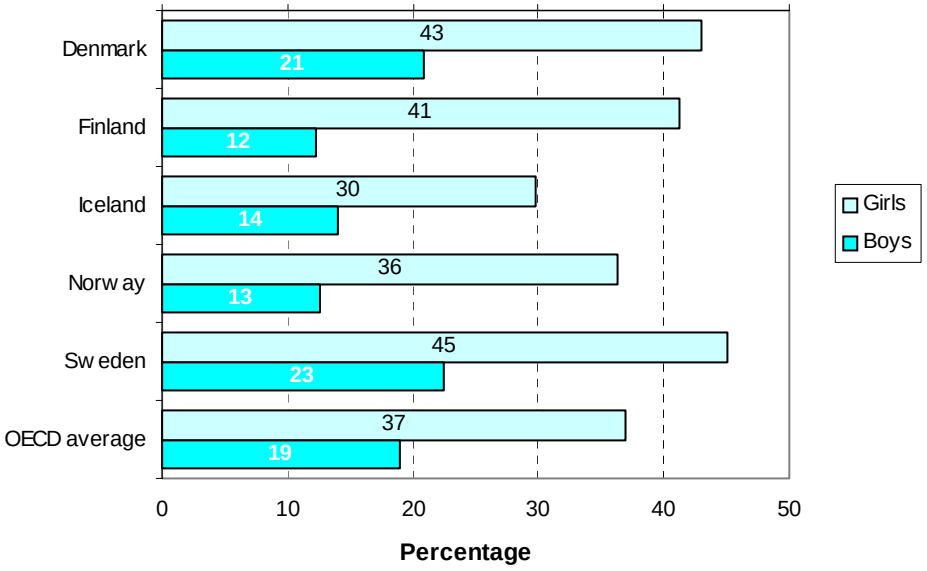


Figure 3.4 Percentage reading fiction several times per month by gender



With a view to revealing and, finally, to reducing the gender gap in reading literacy performance, we conducted some further analyses on the PISA data to examine *the change in the gender differences in reading literacy by controlling for those reading interest and activity factors which showed the strongest discriminating impact on reading proficiency*. We intend to provide some answers particularly to the following questions:

- Which are the reading interest and activity factors that are the most influential in reducing gender differences in reading literacy?
- How similar are the factors that reduce the gender gap most effectively in terms of the three reading literacy aspects - retrieving information, interpreting texts, and reflection on and evaluation of texts?
- To what extent do the gender differences decrease, when the most influential factors in reducing gender differences are controlled for simultaneously?

By answering these questions we also seek to find out whether a Nordic model could be established for reducing the gender gap by enhancing boys' reading interests and activities and to what extent this model would reflect similarities and differences among various Nordic countries and in comparison with the OECD average. In our view, identifying the factors that would most strongly reduce the gender gap will help us advance pedagogical development so as to promote equal opportunities for both genders. How far such a new pedagogy can be developed in all Nordic countries through co-operation is an interesting challenge that will be discussed at the end.

3.3 Building up a statistical model

The PISA data are hierarchically structured. They contain two levels - the school level and the student level. To take advantage of the information included in the data structure and to avoid the problems of intra-class correlated variables, the data were analysed with a multilevel modelling technique (Bryk & Raudenbush 1992; Goldstein 1987, 1995), using HLM software (Raudenbush et al. 2000).

In all the following analyses, the statistical method is a two-level regression model, with students as level 1 units and schools as level 2 units. In the models, reading literacy proficiency was used as a response variable. Gender was coded as 1 for girls and 0 for boys, in which case the coefficient connected with gender is an estimate of how much better (or worse) the girls perform in reading literacy compared with boys.

The interest and activity variables controlled for are of two types. The first type is a continuous variable, which is a combination of student responses to several questions and standardised so that the mean for the OECD student population is 0 and the standard deviation is 1. These variables are: *Engagement in reading, Interest in reading, Diversity of reading materials, Cultural communication, Self-concept in reading,* and *Effort and perseverance*. The second type is an ordinal variable with 4 to 7 categories. These variables

are: *Daily reading, Books at home, Reading fiction, Borrowing books from the library*. In PISA, *Engagement in reading* was a combined factor and derived from students' level of agreement (on a four-point scale) on the following nine statements: *I read only if I have to; Reading is one of my favourite hobbies; I like talking about books with other people; I find it hard to finish books; I feel happy if I receive a book as a present; For me reading is a waste of time; I enjoy going to a bookstore or a library; I read only to get information that I need;* and *I cannot sit still and read more than a few minutes*.

Graphical exploratory analyses revealed that there are linear dependencies between the ordinal controlling variables and reading proficiency, if the distances between the categories are considered of equal length. Thus, the regression coefficient of an ordinal variable is the change in the response variable associated with moving from one category in the explanatory variable to the next one, and for a continuous variable it is the change in the response variable associated with a change of one standard deviation in the explanatory variable.

Country mean adjustment was used in the Nordic average and OECD average models to exclude variation in the response variables due to differences between the country means. Only 21 OECD countries are included in the OECD average model, since some countries did not collect data on the variables Self-concept in reading and Effort and perseverance. To avoid the problems of multi-collinearity, only the most effective variables of the highly correlated controlling factors were chosen for the final models.

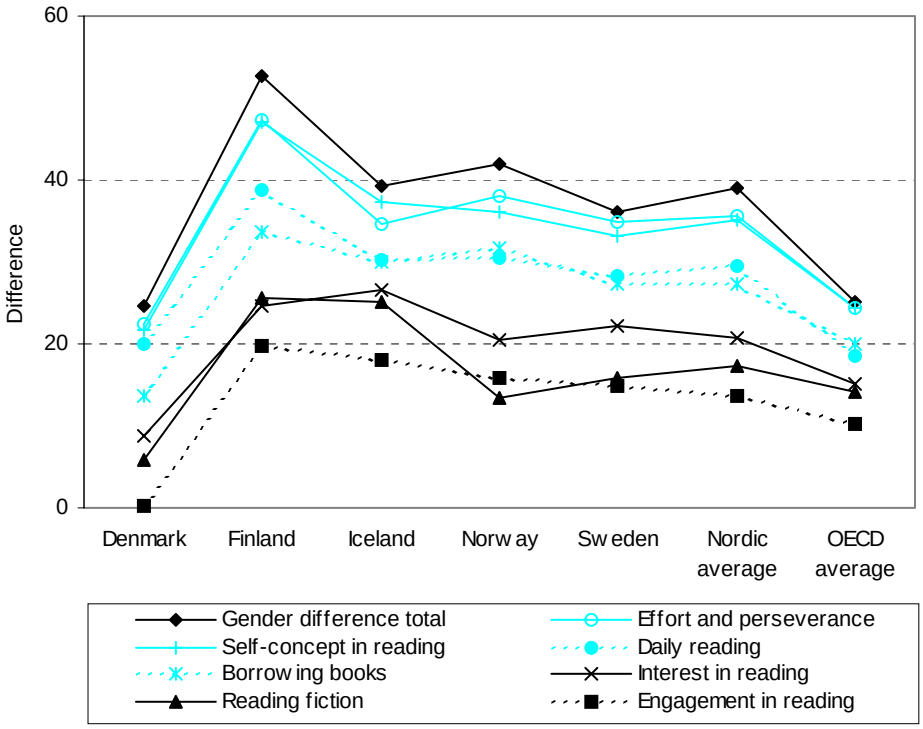
There may be some small differences in the numerical results of this study compared with the initial PISA results (OECD 2001). This is mainly due to the different estimation methods. In multilevel modelling, the intra-class correlation of the variables has an influence on the estimation results.

3.4 Results

3.4.1 Important interest and activity factors

When the effect of each single factor was examined separately by means of the two-level regression model, controlling for the factor Engagement in reading was what reduced the gender gap most significantly in each Nordic country, as shown in figure 3.5. When the effect of this factor is controlled for, that is, boys' and girls' engagement in reading is equalized statistically, the gender difference decreases an average of 25 points in the Nordic countries. In Denmark this would decrease the gender difference from 25 points to zero and thus close the whole gender gap. In Finland, controlling for Engagement in reading would reduce the gender difference by 33 points, from 53 to 20. In Iceland the reduction would be 21 points from 39 to 18; in Norway 26 points, from 42 to 16, and in Sweden 21 points, from 36 to 15 points. In the OECD countries, the average reduction would be 15 points from 25 to 10.

Figure 3.5 Gender differences in reading literacy when each interest and activity variable is controlled separately



The second strongest reducer of the gender gap was Reading fiction in leisure time. If this factor was controlled for, the gender difference in the Nordic countries decreased by 22 points, on average. The reduction was significant for each of the countries: in Denmark it was 19 points, in Finland 27 points, in Iceland 14 points, in Norway 28 points, and in Sweden 20 points. The third strongest factor reducing the gender gap was Interest in reading. When this factor was controlled for the difference in Nordic countries decreased by 18 points on average and, again, significantly for each country. Likewise, the Nordic average decreased by 12 points when the variable Borrowing books and by 9 points when Daily reading was controlled for.

Other factors reduced the gender difference by less than 9 points. These were, in descending order: Diversity of reading material, Cultural communication at home, Self-concept in reading, and Student's own effort and perseverance. Other factors related to reading interests and activities did not reduce the gender difference significantly.

In brief, the results indicate that in each Nordic country engagement in reading, reading fiction frequently, and interest in reading have a strong effect

on gender differences in reading literacy proficiency. If these factors could be equalised in real life; that is, if the boys were engaged in reading to the same extent as girls; if they read novels and short stories as often as girls and if they were similarly interested in reading as a hobby, the gender gap in reading literacy proficiency could be significantly reduced in the Nordic countries as well as in the OECD countries in general.

3.4.2 How do the effects differ for the three subscales?

The results were quite similar in all three aspects of reading literacy, retrieving information, interpreting texts and reflection on and evaluation of texts, when the most significant single factors were controlled separately. On all these aspect scales, controlling for Engagement in reading reduced the gender differences the mot, (see figure 3.6). Similarly, the order of the factors with the strongest reduction effects was the same as on the combined reading literacy scale. The next most powerful reducers were Reading fiction frequently on your own (see figure 3.7), Interest in reading, Borrowing books frequently from the library, and Daily reading.

Figure 3.6 Gender differences in the subscales of reading literacy when Engagement in reading is controlled

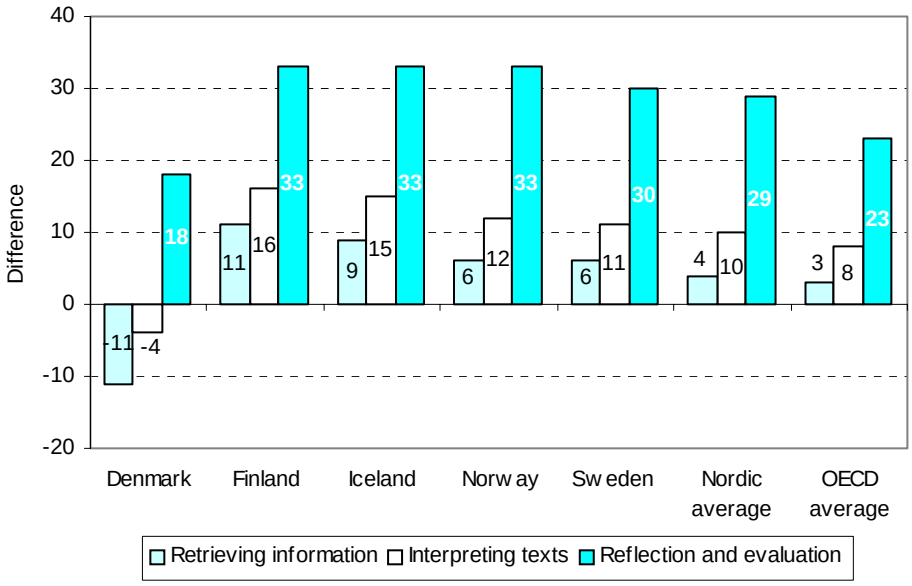
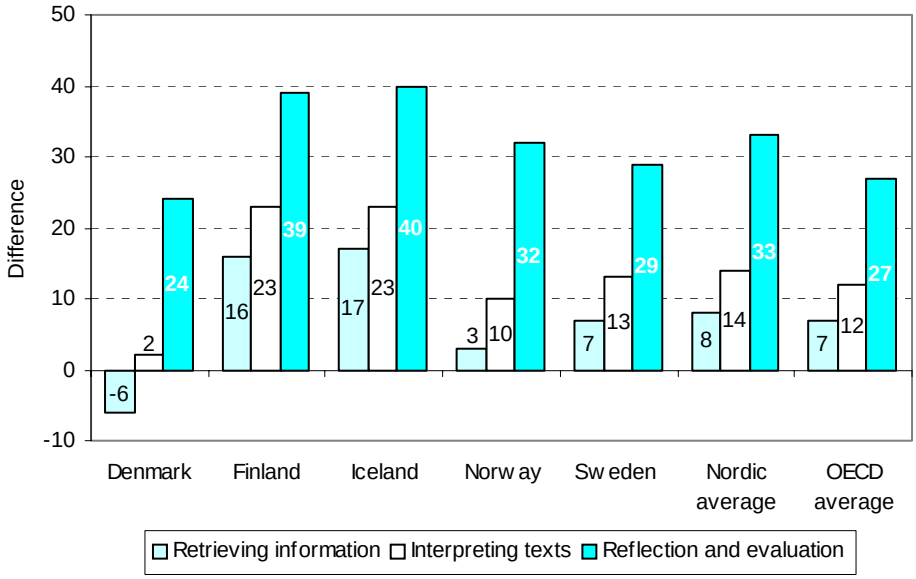


Figure 3.7 Gender differences in the subscales of reading literacy when Reading fiction frequently is controlled



Because the original gender gap was different in each sub-domain, the respective amounts of reduction varied as well. The original difference was smallest in retrieving information and largest in reflection and evaluation.

In *retrieving information*, controlling for Engagement in reading reduced, on average, the Nordic gender difference by 26 points, from 30 to 4. In Norway and in Sweden the significant gender difference disappeared altogether, and in Denmark the gender gap even reversed to favour boys. In Finland and in Iceland the reduction was also significant: in Finland 34 points and in Iceland 22 points. Likewise, controlling for the factor Reading fiction frequently closed the gender gap in Denmark and in Norway. In the same vein, controlling for the factors Interest in reading and Borrowing books from the library, respectively, removed the originally significant gender difference on the information retrieving scale in Denmark.

In *interpreting texts*, the situation was similar, except that the original gender difference was larger and controlling for Engagement in reading did not reduce the gender difference to the same extent as it did in retrieving information. The Nordic average gender difference decreased by 27 points, from 37 to 10, when Engagement in reading was controlled for. The second strongest reducer was, again, Reading fiction frequently and the third strongest was Interest in reading.

In *reflection and evaluation*, controlling for Engagement in reading reduced the gender difference about as much as on the other sub-scales. The Nordic

average decreased by 25 points, from 54 to 29. The next strongest reducers were Reading fiction and Interest in reading. Yet the gender gap did not disappear in any of the Nordic countries, mainly because the original gap was so wide.

To summarise, the findings reveal that by bringing boys' behaviour closer to that of girls' in terms of engagement in reading, frequent fictional reading, interest in reading and borrowing books for reading in leisure time, the gender gap in reading proficiency could be closed or reduced significantly, and fairly similarly for all three sub-scales included in the reading literacy. In retrieving information, where the gap in the beginning was smallest, the statistical evening out described above made it disappear almost totally. In interpreting texts the gap likewise reduced nearly as much, yet remained wider because it was wider originally. Similarly, in reflection and evaluation, the gender differences reduced, even though the gap still remained relatively wide in every Nordic country as well as in the OECD countries on average, due to the larger original difference as compared with the other sub-scales.

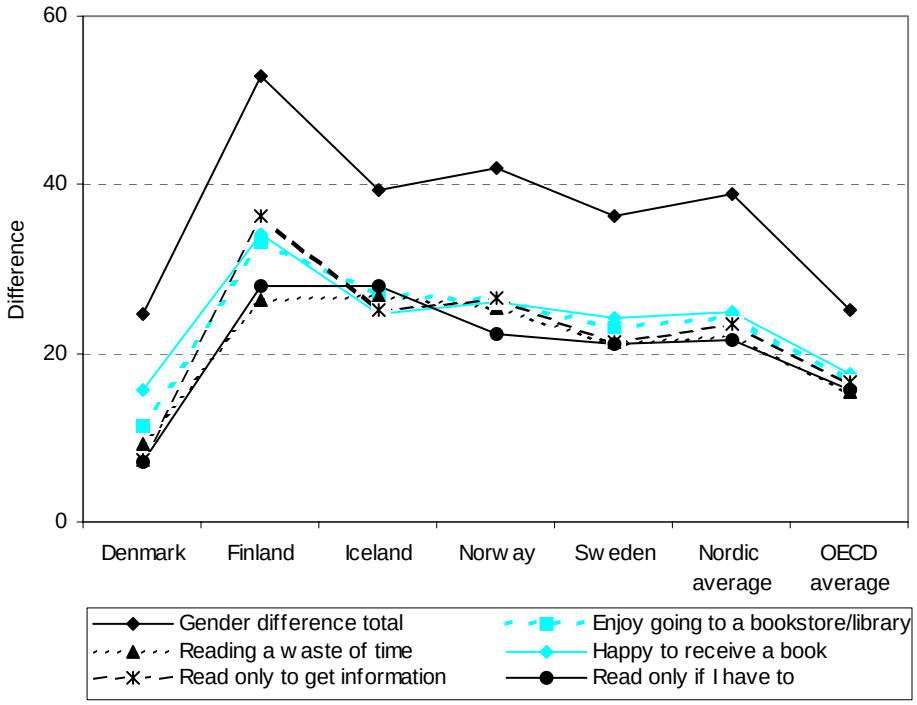
3.4.3 The most powerful components of engagement

In order to explore in more detail which components of engagement would have the strongest effect on reducing the gender difference, we also conducted a regression analysis on the original single variables. The results of this analysis are shown in figure 3.8. They indicate that the statements most strongly associated with the gender differences, while controlled for, were those revealing a strong negative attitude towards reading for enjoyment, such as *I read only if I have to; For me reading is a waste of time;* and *I read only to get information*. Especially in Finland, decreasing the negative attitude of boys would remove the gender gap.

In addition, in Iceland and Norway agreement with the statement *I feel happy if I receive a book as a present* clearly reduced the gender gap, and in Norway the statement *I enjoy going to a bookstore or a library* also had the same effect.

These results suggest that a reduction in the gender gap in the Nordic countries will not be achieved until we manage to change boys' negative attitudes towards reading, so that boys, too, find reading an enjoyable hobby, not just a waste of time or merely a means of searching for information. The results also suggest that it would be beneficial to include such elements as books as presents, use of libraries or visits to bookstores, and discussions about books in boys' culture. In contrast, the gender gap in reading proficiency had little to do with the student's ability to concentrate, as the gender differences could not be associated with variables such as *I cannot sit still and read more than a few minutes* or *I find it hard to finish books*.

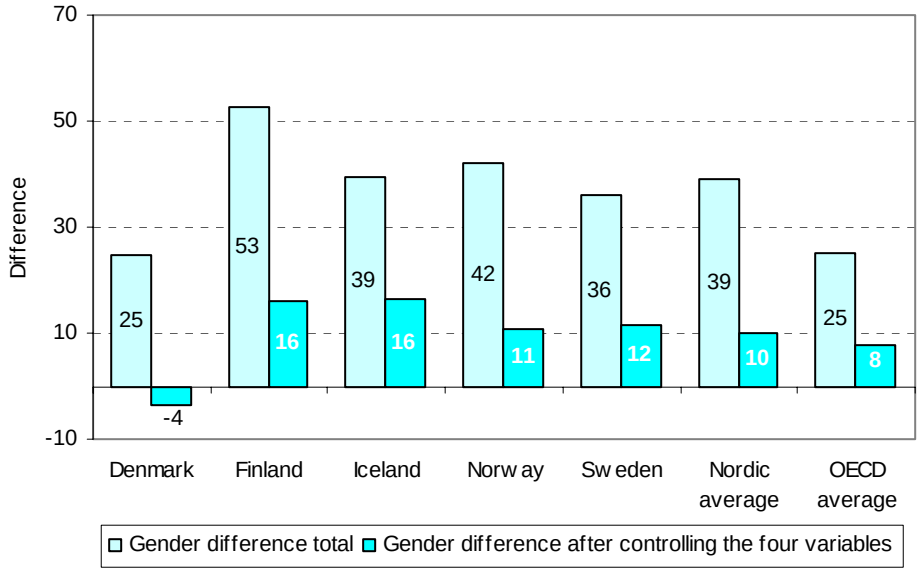
Figure 3.8 Gender differences in reading literacy when the components of engagement in reading are controlled



3.4.4 Controlling the strongest factors simultaneously

Having examined which factors have the strongest effect on reducing the gender gap, we also investigated how the situation would change if the factors with strongest effect were controlled for simultaneously. When selecting the control variables, we had to exclude some of the background factors to avoid multi-collinearity effects. The factors included in the combined model were as follows: *Engagement in reading, Reading fiction frequently, Self-concept in reading,* and *Student's effort and perseverance*. Controlling for these factors simultaneously, we constructed statistically an imaginary situation where boys and girls were evenly engaged in reading; boys and girls read fiction equally as frequently in their leisure time; boys' self-concept was as strong as that of girls; and boys and girls showed equal effort and perseverance with regard to reading. The residual gender differences, after controlling for all these four background factors simultaneously, are shown in figure 3.9.

Figure 3.9 Reducing gender differences in reading literacy when the four most significant activity factors (Engagement in reading, Reading fiction frequently, Self-concept, Effort and perseverance) are controlled for simultaneously



The findings show that in this imaginary situation the gender gap disappeared in the case of Denmark, or at least reduced significantly as was the case in all the other Nordic countries, when the four activity and interest factors were controlled for at the same time. In Denmark, the difference after controlling for the factors was even slightly in favour of boys. However, this difference was not statistically significant. In Norway, the gender difference narrowed down to 11 points, in Sweden to 12 points, and in Finland and Iceland to 16 points. In Norway and Finland the reduction was relatively larger than in Iceland and Sweden. The Nordic average gender gap decreased to 10 points, the OECD average to 8 points. The reduction was significantly larger in the Nordic countries than in the OECD countries on average.

This model with four controlling variables explained 20 to 28% of the variance in the Nordic countries at the student level and 18 to 27% of the variance at the school level. In the OECD countries on average, the percentages were slightly lower: 13% for the student level variance and 18% for the school level variance. These percentages are remarkably high, considering that only four activity and interest variables were included in the model. This shows that reading engagement, activity and enjoyment of reading fiction, self-concept and effort are important aspects both of reading proficiency in general and of the related gender differences in particular.

3.5 Conclusions

The models for explaining and reducing gender differences produced similar results in all the Nordic countries. Although gender differences in reading proficiency varied considerably across both the OECD and the Nordic countries, reading interest, activity and engagement factors seemed to have very similar effects on these differences in all the countries. In Denmark, the gender gap could be closed statistically by controlling the effect of one single factor: Engagement in reading. In the other Nordic countries, not even controlling for all four factors could make the difference disappear entirely, although it did decrease the difference to about one third of the original size. In relative terms, the controlling procedure reduced the difference the most in Finland and Norway, but least in Iceland. In all, the four-factor model reduced the gender difference more in the Nordic countries on average than it did in the OECD countries in general. This implies that reading activity and engagement factors are more powerfully associated with reading proficiency and gender differences, in particular, in the Nordic countries than in all OECD countries on average.

The modelling outcomes are to a great extent similar for the different Nordic countries, both when it comes to explaining reading proficiency in general and gender differences more specifically in the three aspects of reading. Considering the individual variables related to gender differences, the factor Reading fiction turned out to be most strongly correlated in Norway, whereas in the other Nordic countries this was true for Engagement in reading. Reading fiction, however, also proved to be surprisingly important in all these countries on the retrieving information sub-scale. On this scale, controlling for the effect of Reading fiction made the gender difference disappear not only in Denmark but also in Norway.

In terms of the interest and activity factors involved in reducing the gender gap, the differences between the Nordic countries were so small, after all, that a joint pedagogical effort to enhance boys' engagement in reading and thereby to improve both the equality and quality of students' reading literacy achievement is desirable. As students become engaged readers, they provide themselves with self-generated learning opportunities that may be equivalent to several years of school education. Engagement in reading can compensate for low family income and poor educational background (Guthrie & Wiegfield 2000, p 404).

At the same time as providing a stimulus for a co-operative Nordic pedagogy, the PISA findings (OECD 2001) draw attention to the link between cognitive and affective elements of reading. Poor readers need both affective and cognitive support. The affective support is increased through real-world interaction, interesting and exciting texts, personal choice and significance of literary experiences and collaboration with peers. These instructional attributes will support the development of effective cognitive strategies (Guthrie & Wiegfield 2000).

The PISA results also suggest that the Nordic countries could learn something, particularly with regard to boys' reflective and evaluative reading, from Anglo-American pedagogy, especially the pedagogies of Australia, Ireland, New Zealand and the United States. The PISA results for these countries showed smaller gender gaps than in most of the Nordic countries, particularly with regard to reflective and evaluative literacy (OECD 2001).

In the light of these findings, it is apparent that, in the Nordic countries, both interest and engagement in reading and reading fiction are seen as features of feminine culture. If boys' attitudes towards reading are truly so negative that they do not read anything unless they have to and consider reading just a waste of time, there is certainly a need for cultural change. We should invest heavily in attitudinal development in our pedagogy; schools should likewise provide more literature and reading materials that would interest boys, such as science fiction and fantasy stories. Ideally, this would help boys realise that reading fiction can be enjoyable and interesting. Introducing male authors and their works would, perhaps, prompt boys to take up reading. Parental involvement should also be encouraged, and parents should be informed about the significance of reading as a leisure activity. Fathers in particular should be made highly conscious of the role model they provide for their sons as regards reading. We should get young people to realise that even 'a real man' reads books, including fiction.

References

Bryk, A., & Raudenbush, S. (1992). *Hierarchical Linear Models: Application and Data Analysis Methods*. Newbury Park: Sage.

Goldstein, H. (1987). *Multilevel Models in Educational and Social Research*. London: Griffin.

Goldstein, H. (1995). *Multilevel Statistical Models*. London: Edward Arnold.

Guthrie, J.T. & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P.B. Mosenthal, P.D. Pearson & R.Barr (Eds.) *Handbook of reading research. Volume III*, p 403-422. New York: Longman.

Kontogiannopoulou-Polydorides, G. & Adamapoulou, A. (1996). The influence of culture, society and educational systems. In Wagemaker, H., Taube, K., Munck, I., Kontogiannopoulou-Polydorides, G. & Martin, M. 1996. *Are girls better readers? Gender differences in reading literacy in 32 countries*, p 77 – 98. Amsterdam: IEA.

OECD (2001). *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD Publications.

Purves, A. & Elley, W.B. (1994). The role of the home and student differences in reading performance. In Teoksessa W. B. Elley (Ed.*) The IEA study of reading literacy: achievement and instruction in thirty-two school systems*, p. 89-121. Oxford: Pergamon & IEA.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2000). *HLM 5. Hierarchical Linear and Nonlinear Modeling*. Scientific Software International.

Wagemaker, H., Taube, K., Munck, I., Kontogiannopoulou-Polydorides, G. & Martin, M. (1996). *Are girls better readers? Gender differences in reading literacy in 32 countries.* Amsterdam: IEA.

Young, J. 2000. Boy talk: Critical literacy and masculinities. *Reading Research Quarterly 35*, 3, p.312-337.

4 MATHEMATICAL LITERACY AND COMPETENCY CLASSES

Are Turmo, Marit Kjærnsli, and Astrid Pettersson

4.1 Mathematical literacy in PISA

Mathematical literacy is defined in PISA as

> *"the capacity to identify, to understand, and to engage in mathematics and make well-founded judgements about the role that mathematics plays, as needed for an individual's current and future private life, occupational life, social life with peers and relatives, and life as a constructive, concerned, and reflective citizen" (OECD 2000).*

To transform this definition into an assessment of mathematical literacy, three broad dimensions have been identified:

- *Processes.* The focus here is on students' abilities to analyse, reason and communicate ideas effectively by posing, formulating and solving mathematical problems.
- *Content.* PISA emphasises broad mathematical themes such as change and growth, space and shape, chance, quantitative reasoning, and uncertainty and dependency relationships. In the first cycle of PISA, only the themes *change and growth* and *space and shape* were assessed. In the PISA study in 2003 all the themes are included in the assessment.
- *Context.* An important aspect of mathematical literacy is doing and using mathematics in a variety of situations, including personal life, school life, work and sports, local community and society.

Concerning the processes dimension, the items are categorised by the kind of competency that is mainly required for answering correctly. Three *competency classes* have been defined:

- *Competency class 1* includes factual knowledge, the ability to recognise mathematical objects and the ability to do routine procedures and standard algorithms.
- *Competency class 2* includes the ability to make connections between different domains of mathematics. It involves the use of representations, to realise the relationship between definitions, mathematical proof, examples and claims, and the use of formal mathematical language.
- *Competency class 3* involves the ability to recognise mathematics in different contexts in real life, and the ability to use mathematics for

problem solving. These processes include critical thinking, and the ability to analyse and to reflect.

The overall results in mathematics for the Nordic countries looked at from an international perspective are briefly described in chapter 1. We notice that with the exception of Norway, the four other Nordic countries score significantly above the OECD mean in PISA (OECD 2001). Furthermore, Finland significantly outperforms the other Nordic countries, and only three countries in the study have a higher average score than Finland.

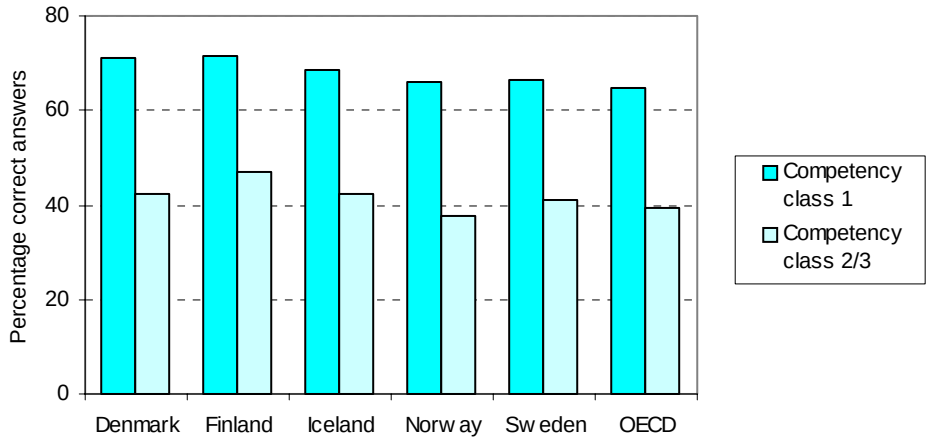
4.2 Achievement results by competency class

In this section we will present results for the different competency classes (see above). In PISA 2000 there were only two items in competency class 3, and we have therefore combined this competency class with competency class 2, which consists of 19 items. In competency class 1 there were 10 items. Since there are no individual achievement subscales for the different competency classes, we will in this section compare the achievement of different groups of students by comparing the percentages of correct answers.

4.2.1 Results for all students

Figure 4.1 shows the results for each of the two competency classes in the Nordic countries as well as for the OECD mean.

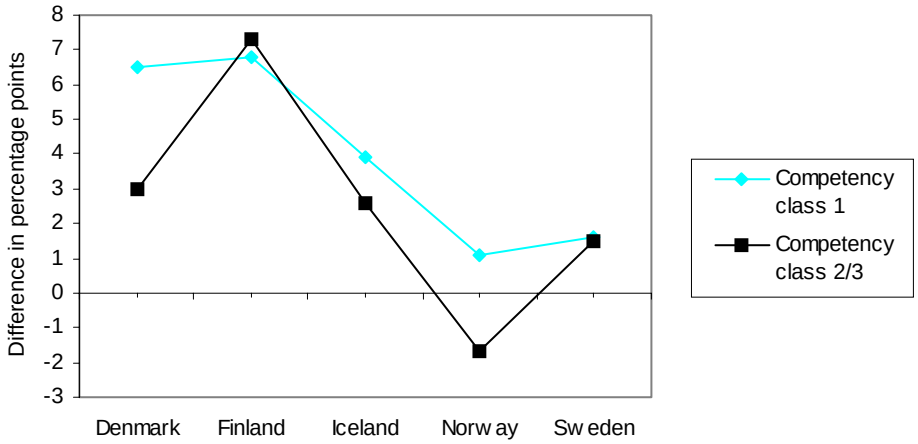
Figure 4.1 Results (average percent correct) for each of the two competency classes



In order to highlight the differences between countries, we have further displayed in figure 4.2 the differences between the percentages of correct answers in the two competency classes. In other words; the figure shows how much higher or lower the percentages of correct answers are in the Nordic

countries compared to the OECD mean. From figure 4.2 it can be seen that the average for Norwegian students is lower than the OECD mean in competency class 2/3. All the other Nordic countries perform better than the OECD mean in this competency class. The results for Denmark and Iceland are very similar in this competency class. In competency class 1, all the Nordic countries perform better than the OECD mean. The results for Denmark and Finland are very similar in this competency class. Norway and Denmark have the largest differences between the two competency classes while the students in Sweden and Finland perform equally well compared to the OECD mean in both the competency classes. Another feature that emerges from figure 4.2 is the following: Whereas the differences between the Nordic countries concerning factual knowledge and simple mathematical routines (class 1) are relatively small, the differences are more pronounced when it comes to more abstract mathematics, with Finland and Norway as the two extreme cases.

Figure 4.2 Achievement (percent correct) relative to the OECD mean



4.2.2 Gender differences

Figure 4.3 shows the differences between the international Rasch scale scores of girls and boys in mathematical literacy in the Nordic countries. A positive value means that the difference is in favour of girls. On average, the gender difference in the OECD countries participating in PISA is –11, in other words, the difference is 11 score points in favour of boys. Figure 4.3 shows that boys score higher than girls in mathematics in all the Nordic countries, except Iceland. Denmark has the largest gender gap among the Nordic countries. In Norway too boys score significantly higher than girls, and the difference in Norway equals the OECD mean difference. In Finland and Sweden the differences are small and not significant.

Figure 4.3 Gender differences in score points in mathematics. Positive values are in favour of girls

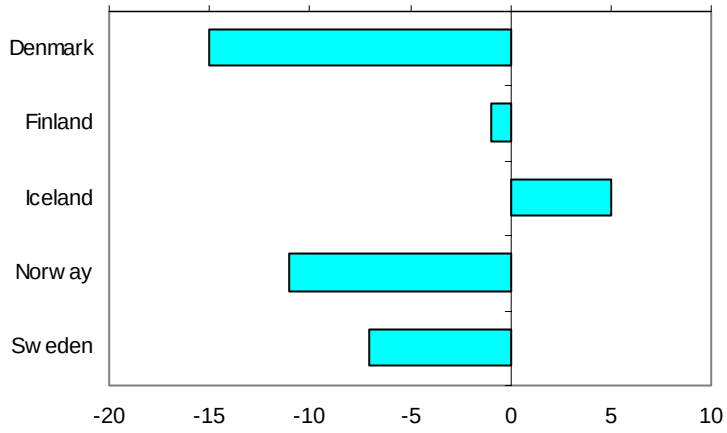


Figure 4.4 shows the differences in achievement between girls and boys in the different competency classes. Again, due to the relatively few items in competency class 3, this competency class has been combined with competency class 2. And as explained earlier (for figure 4.1), gender differences are given as differences in average percent correct answers.

Figure 4.4 Gender differences in achievement (percentage points) within the two competency classes

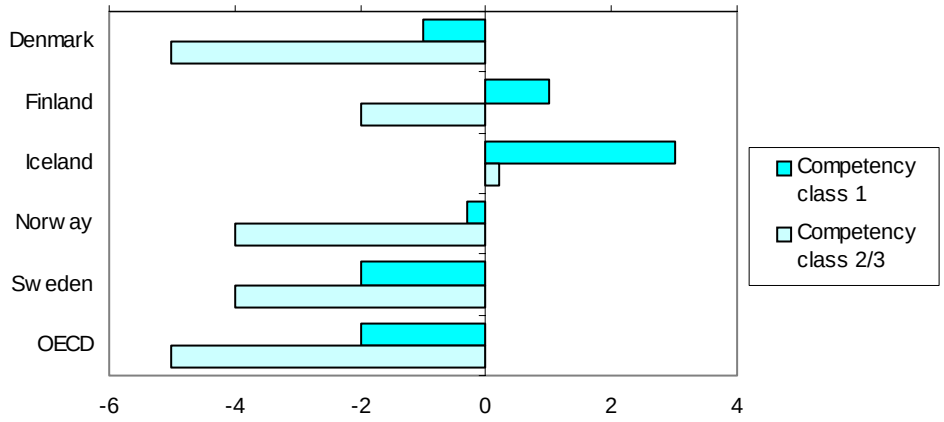


Figure 4.4 shows that generally, boys score higher than girls in competency class 2/3 in the Nordic countries, with Iceland as the single exception. For competency class 1 the picture is somewhat more differentiated: girls

outperform boys in Iceland and Finland, and all differences in favour of boys are smaller than for class 2/3. A common feature of all the Nordic countries is that girls perform relatively better in competency class 1 compared to competency class 2/3. This is similar to findings from previous studies. In TIMSS (Third International Mathematics and Science Study, Beaton et al. 1996), girls performed relatively better on items in the categories "Number" and "Algebra" (Lie et al.1997), which can be argued to be more "algorithmic" in nature than other topics. It is also interesting to compare these results with the findings reported in chapter 5.3.2 on gender differences in the two kinds of scientific literacy, conceptual understanding and process skills. This comparison does not invite a simple explanation of gender-specific learning styles.

4.3 Focus on one item

So far we have presented and discussed results for the mathematical literacy achievement scores and for each of the competency classes. We will now study the results for one single item in detail. This will illustrate the kind of mathematics items the students are exposed to in PISA. It also illustrates how double-digit coding can be used to make diagnostic comparisons between countries. Table 4.1 presents the results for the unit "Continent area" in the Nordic countries, which is a unit that has only one item. In this respect the unit is not typical, because most of the units in PISA have more than one item. The unit is classified in the content category *Space and shape* and can be viewed on the OECD's PISA website (*www.pisa.oecd.org*).

The students are asked to estimate the area of Antarctica using the map scale (a map of Antarctica is given, together with the scale in the form of a line with marks labelled by 0, 100 km, 200 km ….). They are explicitly asked to show their work and explain how they made the estimate. The challenges for the students in this item consist of:

- using an adequate method to estimate the area, *and*
- applying the map scale to calculate the correct "real" area.

The item has been coded using a double-digit coding scheme explained below. The coding criteria are based on the two separate challenges incorporated in the item. The codes can be sorted into four main categories; full credit, partial credit, no credit and non-response.

To get full credit, the student has to use a correct method and also get the correct answer. The first digit gives the number of score points, whereas the second digit is used to differentiate between different approaches or answers, see table 4.1. Code 21 is used when the students estimate the area by drawing one square or rectangle. If the student has estimated the area by drawing one circle, the code 22 is used. Code 23 is used for responses where the student has added areas of several geometric figures. Code 24 is given when the student estimates the area by the use of another correct method. Finally, code 25 is used if the student only gives the correct answer, without any work shown.

The partial credit codes (1 for first digit) are for responses using a correct estimation procedure, but giving an incorrect or incomplete answer, usually by incorrect application of the map scale. The second digit indicates the different approaches, matching the second digit of the full credit codes. At the no credit level, only two different codes are used. Code 01 is applied when the student has calculated the perimeter instead of the area, and code 02 is used for other incorrect responses. In table 4.1 the non-responses are sorted into two categories, "Reached" and "Not reached". The results for the Nordic countries and the OECD mean are shown.

Table 4.1 Results for the unit "Continent area". Percent distribution of responses

Codes and descriptions		Denmark	Finland	Iceland	Norway	Sweden	OECD mean
Full credit		12	21	10	6	10	10
21	Square/rectangle	5	7	4	4	3	4
22	One circle	1	2	1	1	2	1
23	Several figures	4	7	4	1	3	3
24	Other correct	1	2	1	0	2	1
25	No work shown	1	1	0	0	0	1
Partial credit		16	22	20	18	25	19
11	Square/rectangle	9	13	13	12	15	12
12	One circle	1	1	1	2	3	1
13	Several figures	5	5	4	3	4	4
14	Other method	1	3	2	1	3	2
No credit		20	19	19	20	17	19
01	Perimeter	3	1	1	2	2	2
02	Other incorrect	17	18	18	18	15	17
Non-response		51	41	51	57	49	52
	Reached	47	38	47	52	45	48
	Not reached	4	3	4	5	4	4

Table 4.1 shows that there are pronounced differences between the Nordic countries when it comes to the percentages of fully correct responses (codes 21-25) to this item. Finland has the highest percentage (21%), and Norway the lowest (6%). On the partial credit level (codes 11-14), the differences between the countries are small. The same is the case at the no credit level (codes 01 and 02). The most striking results in the table are the high percentages of non-responses. In Norway the percentage is as high as 57 percent, which is somewhat higher than the OECD mean. However, even in Japan and Korea, the two highest-scoring countries in mathematical literacy, the percent non-responses were well above 50%. It can also be seen from the table that most of the students who did not respond to the item did actually reach it, so the challenges in the item itself have provided the major hurdle.

Table 4.1 also invites a further discussion on students' problem solving strategies for handling this challenge. Among students who successfully attacked the first challenge of finding the area (regardless of scale) and

received full or partial credits, the most common approach for estimating the area was to draw one square or rectangle (codes 21 or 11), but adding the sum of several squares or rectangles (codes 23 or 13) also emerged as a relatively popular strategy. Each method was used by at least *some* students in each of the Nordic countries. Since the shape of the continent of Antarctica clearly approximates to a circle, one could have expected that more students would actually have applied this procedure. It is assumed that the more complicated formula for the area of a circle was the main reason that this did not happen.

The questions arise: Why did so many students reach the item but leave it blank? Or even: Why is this item so difficult at this grade level throughout the world? Two other studies can illuminate this discussion. In a recent follow-up study by Haugsten (2002), Norwegian 15-year olds were asked to find the area of the two American states Wyoming (rectangular) and Texas (complicated shape), first on the map in cm^2, and then in km^2 using the scale. By splitting the challenge into two separate steps, the non-response rate was much lower for the first part (about 30%), but still close to 50% for the last step. In the 1999 field trial of PISA, the unit "Continent area", in addition to the item discussed here, also included an introductory multiple choice item asking students to estimate the real distance between two points on the map, using the map scale. For this item the non-response rates were "normal", below 10% in most countries, and correct responses amounted to around 50% or higher. It therefore seems that the main hurdle for the challenge in the PISA item has to do with converting areas on paper into "real" measures. The use of a linear scale cannot be applied in a straightforward way when areas are to be calculated.

4.4 Concluding remarks

PISA tries to assess what students need to know in order to be "prepared for life" as informed and reflective citizens. The test is not based on the countries' curricula but on what is regarded as important in this lifelong context. This fact distinguishes mathematics in PISA from the TIMSS study (Beaton et al. 1996). Nevertheless, the differences between the Nordic countries in PISA can be compared with TIMSS results. In TIMSS 1995 (Beaton et al. 1996), the Swedish 13-year-old students clearly outperformed students in Norway, Denmark and Iceland, while Finland did not participate. On the other hand, Finland was the only Nordic country that participated in TIMSS 1999 (Mullis et al. 2000), testing students from the 8^{th} grade. In this study Finland scored well above the international average, in spite of their students being at the lower end of the defined age range. In the light of these features, the PISA results for the Nordic countries are not very surprising. It seems to be a consistent finding that Finnish students seem to achieve better than their Nordic peers in mathematics.

Comparison by competency classes reveals that all the Nordic countries perform above the OECD mean on the items in competency class 1, which includes mainly routine problems and algorithms. For items in competency

class 2/3, which include more advanced mathematical thinking, only the Norwegian students perform lower than the OECD mean.

The gender differences in mathematical literacy are small compared to the dramatic differences in reading literacy (see chapter 2). In Denmark and Norway, boys score significantly higher than girls, whereas there were no significant gender differences in the other Nordic countries. We have also noticed that girls seem to have a relative advantage compared to boys in competency class 1.

In PISA 2000 there are only 32 mathematics items, covering only a limited part of the framework for mathematical literacy (OECD 2000). It is important to study in more detail how the different countries teach mathematics and how the students solve different types of items. As an example of an analysis of the last type, we have included an analysis of Nordic student responses to one particular item. We can carry out a greater variety of analyses following the next PISA study in 2003, when mathematics will be the main subject.

References

Haugsten, B. (2002). *Hvilken betydning har kontekst i matematikkoppgaver? Et studie av matematikkoppgaver i PISA-undersøkelsen.* Master degree thesis in mathematics education. Oslo: Department of Teacher Education and School Development, University of Oslo.

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

Lie, S., Kjærnsli, M. & Brekke, G. (1997). *Hva i all verden skjer i realfagene?* (Norwegian national TIMSS report). Oslo: Department of Teacher Education and School Development, University of Oslo.

Mullis, I.V.S. et al.(2000). *TIMSS 1999. International Mathematics Report. Findings from IEA´s Repeat of the Third International Mathematics and Science Study at the Eighth Grade.* Chestnut Hill, MA: TIMSS International Study Center, Boston College.

OECD (2000). *Measuring Student Knowledge and Skills. The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy.* Paris: OECD Publications.

OECD (2001). *Knowledge and Skills for Life. First results from PISA 2000.* Paris: OECD Publications.

5 SCIENTIFIC LITERACY: CONTENT KNOWLEDGE AND PROCESS SKILLS

Marit Kjærnsli and Bengt-Olov Molander

5.1 Scientific literacy in PISA 2000

5.1.1 Framework and over-all results

In this chapter we will look more closely into the science achievement in PISA 2000 for the Nordic countries. As we have already seen in chapter 1, it is remarkable how much better the students in Finland perform than students in the other Nordic countries. Sweden is also in the group of countries where students score significantly above the OECD average, but the scores are still far lower than in Finland. Students in Norway and Iceland score almost at the OECD average, while Denmark's result is below the average. Compared to the results in mathematics, where the differences between all the Nordic countries except Finland are quite small, we see much larger differences in science.

The aim of PISA 2000 regarding science was to assess to what extent 15-year-olds acquire knowledge in science, which is judged as important in order to be able to participate in a society that to a great extent relies on science. Thus, the focus was not on curricular content in a narrow sense, but rather on scientific literacy. In PISA, scientific literacy is defined as:

> *"the capacity to use scientific knowledge, to identify questions, and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity" (OECD 1999 p.60).*

PISA 2000 was developed around three dimensions of scientific literacy: scientific concepts, situations and processes. Scientific concepts are related to a number of themes, such as biodiversity, human biology and energy. Scientific situations describe frameworks in which knowledge is being applied, such as questions concerning global matters, the local community, or of a more direct personal nature. While for scientific processes, five defined competencies were examined in PISA 2000: *recognition of scientific questions, identification of evidence*; *drawing and evaluating conclusions*; *communicating valid conclusions*; and *demonstration of understanding of scientific concepts*. In this chapter we will use the term process skills for the first four of these, whereas the last will be called conceptual understanding. For a more elaborate

presentation of scientific literacy and dimensions as defined in PISA 2000, we refer to OECD publications in 1999, 2000 and 2001.

5.1.2 Distribution of student performance

The spread of the scores on the scientific literacy scale is lowest in Finland and highest in Denmark. Students in Sweden, Finland and Norway perform almost equally well in science as they do in mathematics. On the other hand, students in Iceland and Denmark perform relatively much better in mathematics than they do in science.

Table 5.1 Distribution of student performance in science literacy. Nordic countries compared to OECD mean

Percentiles	10th	25th	75th	90th
Denmark	347	410	554	613
Finland	425	481	598	645
Iceland	381	436	558	607
Norway	377	437	569	619
Sweden	390	446	578	630
OECD average	**368**	**431**	**576**	**631**

Table 5.1 shows as percentiles the distribution of student scores in the Nordic countries. It can be seen from the table that students in Finland, Norway, Iceland, and Sweden all have a relatively high lowest level compared to the OECD mean.

For all percentiles the results for Finnish students are well above the OECD average while for Danish students they are below this average. The pattern for Norway, Iceland and Sweden is different. The top 25% of students in Sweden achieved roughly the same result as the OECD average, while the corresponding students in Norway and Iceland achieved results slightly lower than the average. However, students at the lower end of the scale in these three countries perform better than the OECD average.

5.2 Conceptual understanding vs. process skills

Scientific literacy requires an understanding of scientific concepts as well as an ability to apply a scientific perspective. The tasks required students to understand certain key scientific concepts and to show that they could acquire, interpret and act on evidence in situations where science can be applied. Compared to other international studies like TIMSS (Beaton et al. 1996), PISA has a much stronger emphasis on what we here call science process skills.

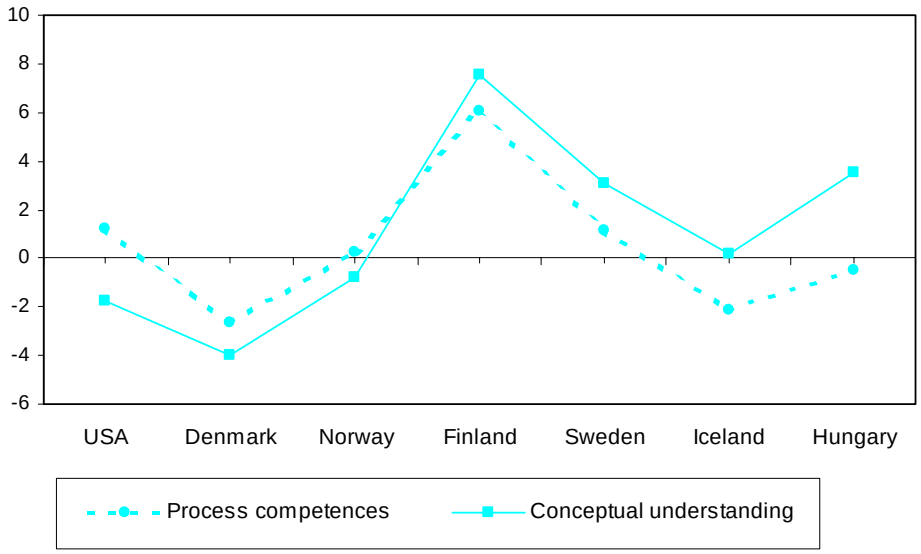
The definition of scientific literacy, together with the dimensions mentioned above (see also OECD 1999, 2000), will be used as a basis for presentation of the results in this chapter. *Scientific processes* in PISA are the intellectual processes that are involved in addressing a question or issue (such as

identifying evidence or explaining conclusions). On the other hand, the content dimension in PISA is defined as the scientific knowledge and *conceptual understanding* that are required in using these processes. We have categorized the items into two groups depending on whether they measure mainly scientific process skills or conceptual understanding.

The results in all countries show the same pattern: the items focusing on conceptual understanding have a higher score. We may interpret this as showing that the "process items" have turned out to be more difficult, a finding that in itself is not worth much consideration. What is interesting, however, is to look more closely at the results and see if some countries perform relatively better in one sub-domain than others. Since there are no separate subscores for each of the two categories, we have calculated the average percent correct responses for each group of items and students.

In figure 5.1 we present the differences in percentage points in relation to the OECD average for each of the two categories. In addition to the Nordic countries we have included the results for the USA and Hungary. As extreme examples, these two countries represent two very different traditions; English-speaking countries seem to put more pressure on the process aspect in their education, while we often see that East European countries give more emphasis to the conceptual understanding.

Figure 5.1 Achievement (average percent correct) in science process skills and conceptual understanding. Percentage points above or below the OECD mean



In figure 5.1 we have ordered the countries by the differences in their results in these two categories. We can see that students in the USA perform relatively

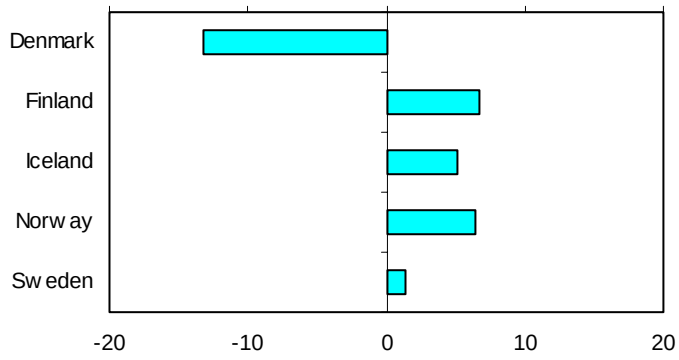
better in the science process competencies (the results from the United Kingdom are the same), while students in Hungary achieve relatively better in the domain of conceptual understanding. Students in Denmark and Norway perform relatively better in the process than in the conceptual understanding categories while Iceland, Sweden and Finland do the opposite. It is, however, important to point out that the differences are small.

5.3 Gender differences

5.3.1 General differences in science literacy

All countries seek to reduce gender differences with regard to achievement in science. In PISA the gender differences in both mathematical and scientific literacy tend to be much smaller than the large differences (in favour of girls) in reading literacy (see chapter 1). Most of the countries show no significant gender differences in science performance. In Korea, Denmark and Liechtenstein there were significant differences in favour of boys, whereas in Latvia, the Russian Federation and New Zealand there were significant differences in favour of girls. These results contrast strongly with those in TIMSS where gender differences in science performance among grade 8 students were much larger, and almost always in favour of boys (Beaton et al. 1996).

Figure 5.2 Gender differences in scientific literacy in the Nordic countries. Positive score point differences in favour of girls



In figure 5.2 the gender differences are displayed for the Nordic countries only. In all Nordic countries, except Denmark, girls perform better than boys, but the differences are small and not significant. The most striking feature is the fact that Denmark deviates from the general pattern, since boys perform significantly better than girls. This is in line with the result from TIMSS where Denmark, together with Israel, was the country with the largest gender differences in science in favour of boys (Beaton et al 1996; Kjærnsli & Lie

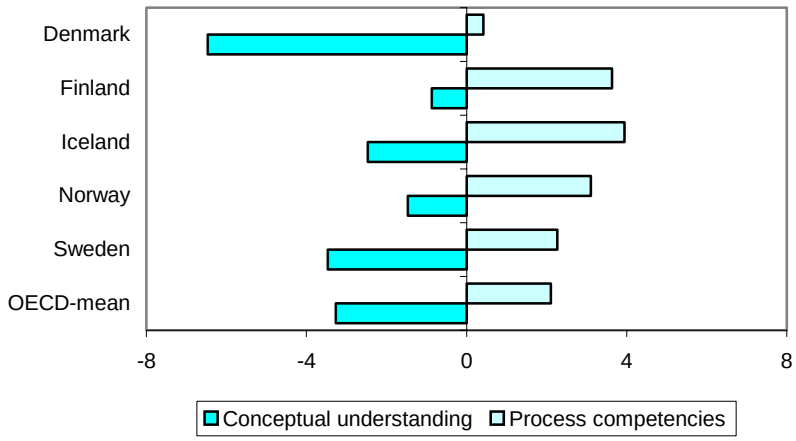
1999). It is interesting to note that in mathematics and reading literacy too the Danish girls fell somewhat behind their Nordic counterparts (see chapter 1).

Do the results presented in figure 5.2 mean that the gender differences in science have been reduced, or do the PISA items in some way favour the girls? There are various possible answers to this question. In PISA, compared to TIMSS, more of the items deal with biology. We know from the TIMSS results that girls perform relatively better in biology compared to physics and chemistry. Furthermore, the PISA items demand extensive reading, which also tends to favour the girls. Finally, the items in PISA have greater emphasis on science processes than previous assessments. Below we will discuss this in more detail.

5.3.2 Conceptual understanding vs. process skills

In the following section we will examine how girls and boys achieve in the domains of conceptual understanding and process competencies. The gender differences (measured as differences between average percent correct responses) are displayed in figure 5.3.

Figure 5.3 Gender differences in achievement within the two categories of items. Positive differences in average percent correct are in favour of girls



It can be seen from the figure that in all Nordic countries, as well as in the OECD countries on average, results show that girls tend to achieve better than boys in scientific process skills. The difference in favour of girls is highest in Iceland and lowest in Denmark. In conceptual understanding, on the other hand, boys obtain better results than girls, particularly in Denmark. To put it in somewhat simplistic terms, girls seem to demonstrate higher scientific skills, whereas boys seem to have better conceptual understanding in science. It is interesting here to compare these results with what could be regarded as the

opposite finding in chapter 4.2: in mathematics boys tend to achieve relatively better within competency class 2/3, which represents “higher order skills”, whereas girls do relatively better in the more routine, competency class 1 tasks. These partly contradictory findings in science and mathematics provide a note of caution against drawing any simplistic conclusions about gender-specific ways of thinking.

5.4 Science and reading

In PISA science tasks, students are required to read and interpret quite complex texts, to understand scientific problems embedded in the various types of texts, and to evaluate and use scientific concepts and processes in presenting responses. This task demands more than the understanding of scientific concepts in a narrow sense. It demands a proficient literacy skill and knowledge of the specific language used in science to describe concepts and processes. PISA 2000 is the first large study in which the same pupils have been tested for competency in reading, mathematics and science. For this reason, PISA presents a unique opportunity to examine possible connections between reading and science. Table 5.2 shows the correlations between the scores in reading and mathematical and scientific literacy for students in the Nordic countries.

Table 5.2 Correlations between scores in reading, mathematical and scientific literacy in the Nordic countries

	Reading and science	Reading and mathematics	Science and mathematics
Denmark	0.87	0.84	0.77
Finland	0.83	0.72	0.74
Iceland	0.82	0.78	0.68
Norway	0.87	0.78	0.78
Sweden	0.87	0.83	0.77

As seen in the table there are strong correlations between the different types of literacy tested in PISA. The correlation between the results in reading and scientific literacy is particularly strong and almost equally strong in all the Nordic countries. The results can partly be explained by the nature of the items in PISA. To be able to respond to science items students are required to read and understand extensive texts. Consequently, a weaker reading ability is likely to be connected to weaker results in science. In fact, students have to understand the “language of science” to be able to produce answers that are judged as correct (Schoultz et al. 2001).

This line of reasoning is supported by a comparison of the distribution of results in reading and science. Students with the 25% lowest, 50% intermediate and 25% highest results in reading in the Nordic countries have been compared

to see how these groups are distributed regarding results in science (using the same distribution categories of low, intermediate and high results). The comparison shows that practically no students who are low-achieving in reading obtain results categorized as high in science, and no students who are high-achieving in reading produce results in the low-achieving category in science. Such findings can be used as the basis for a discussion on the relationship between reading and scientific literacy in PISA as well as from a wider perspective.

5.5 Concluding remarks

In this chapter we have presented some of the Nordic results obtained in PISA 2000. Students in Finland achieved strikingly good general results. The results for Swedish students were slightly above the OECD average, whereas students in Norway and Iceland achieved a score equal to the OECD average. Only Denmark of the Nordic countries performed significantly below the OECD average. It is beyond the scope of this report to suggest explanations for the differences between the countries. However, we wish to comment on some of the findings.

There has been a pronounced change regarding general gender differences compared to earlier international assessments in science (TIMSS and earlier IEA studies). In contrast to the earlier assessments, in PISA 2000 girls actually perform better than boys in a majority of the participating countries. This holds true for all the Nordic countries, except Denmark. We have presented some possible explanations for why girls perform relatively better in PISA than before. One explanation concerns the subject areas tested, as more of the questions relate to the field of biology and also to health and environmental issues. Another explanation might be that the tasks in PISA to a greater extent focus on the relevance of scientific knowledge in a functional context. Thus, since girls perform relatively better on “process items”, girls may have been favoured. A third explanation refers to girls’ good performance in reading literacy, which may explain why girls are better able to handle language, in terms of both the PISA items and the responses, in science as well.

In Finland, results are well above the OECD mean for all percentile groups of students. There is a difference regarding the students with the best results in the Nordic countries, corresponding to the general results mentioned above. For the students with lower results, however, the picture is somewhat different. In Sweden, Norway and Iceland, the lower 10% and 25 % of students achieved results clearly above, in the case of Sweden, or slightly above the OECD average. Thus, there are fewer students with low results in these countries, which is a comforting result. What one would wish to see, however, is a greater number with excellent results.

Earlier science assessment projects such as TIMSS focused strongly on conceptual knowledge. PISA 2000 puts a greater emphasis on tasks requiring

scientific skills ("processes"), which include the recognition of scientific *questions*, the identification of *evidence*, the drawing of *conclusions*, and the *communication* of these conclusions. To the extent that we regard PISA 2000 as a valid measurement of scientific literacy according to the definition given above, we might say that students in Finland are indeed very well prepared for using science in their future lives as useful and critical members of a democratic society. If we want students in the other Nordic countries to acquire scientific knowledge and skills for use in their future lives, it is reasonable to suggest that instruction needs to emphasize science as a specific culture and a specific way of reasoning and thinking. This includes time being spent on interpretation, reflection and discussions on how science relates to society as a whole.

References

Beaton, A. E., Martin, M. O., Mullis, I. V. S., Gonzales, E. J., Smith, T. A. & Kelly, D. L. (1996). *Science Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Kjærnsli, M. & Lie, S. (1999). Kjønnsforskjeller i realfag: Hva kan TIMSS fortelle? I Imsen, G. (red.): *Kjønn og likestilling i grunnskolen*. Oslo: Gyldendal Akademisk.

OECD (1999). *Measuring Student Knowledge and Skills. A New Framework for Assessment*. Paris: OECD Publications.

OECD (2000). *Measuring Student Knowledge and Skills. The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*. Paris: OECD Publications.

OECD (2001). *Knowledge and Skills for Life. First results from PISA 2000*. Paris: OECD Publications.

Schoultz, J., Säljö, R. & Wyndhamn, J. (2001). Conceptual knowledge in talk and text: What does it take to understand a science question? *Instructional Science* 29: p 213 – 236.

6 COMPUTER USAGE AND READING LITERACY

Kaisa Leino

6.1 Focus on computer usage among the Nordic youngsters

Lifelong learning and the development of technologies make new demands on reading literacy and broaden its concept. Reading websites or hypertexts were not assessed in PISA 2000, although electronic texts were included in the definition of written texts in the PISA framework. The students' interest in and usage of computers were, however, surveyed by means of a student questionnaire.

Electronic texts are very common among young people. The multitude of information and the ease of transferring and downloading texts inspire many youngsters to use the Internet for retrieving information and communication. At the same time, the Internet offers a way to publish one's own writings and opinions. The significant role played by the Internet in everyday life has raised questions about the relation between reading literacy and the use of computers and information networks. Some have foreseen the destruction of traditional literacy (e.g. Birkerts 1996) but many also see the stability and new possibilities of the texts, even though the medium has changed (e.g. Nunberg 1996; Reinking et al. 1998; Cope & Kalantzis 2000).

What effect does an active use of the Internet have on literacy skills? Are networks still a boys' playground as earlier studies have shown? What are the purposes teenagers use the Internet for? Are there differences in the use of computers between Nordic teenagers? The focus of this article is on Nordic students' interest in, and confidence and active engagement with the use of computers, as well as the relationship between computer usage and reading literacy proficiency.

6.2 Towards multiliteracy

Literacy in the digital age can be seen as multiliteracy (e.g. Tyner 1998; Cope & Kalantzis 2000; Wade & Moje 2000). To be a member of the information society a student should be *computer literate*, which means s/he has a basic knowledge of and skills in, for instance, word processing and the use of operating systems, spreadsheets, graphic and drawing programs and

information networks. Information networks, on the other hand, require *network literacy* skills, which enable the reader to navigate fluently through www-sites, transfer files from FTP-servers, exploit the services of online stores and have conversations with other Internet users (Gilster 1997; Tyner 1998; Smith 2000; Kapitzke 2001).

Reading electronic texts requires mostly the same reading strategies as reading printed texts. Some skills, however, are emphasized and some strategies need to be implemented in new ways. *Information literacy* skills are important, and a person should be capable of determining the amount of information needed, finding the information effectively, evaluating the information and its source critically and relating new information to prior knowledge. Finding information requires, for example, a knowledge of information structures, such as the basis for classification, and an understanding of the hierarchical structure of files. An information literate person can exploit information in an everyday task and understands the economic, legal, ethical and social effects of using that information (OECD 1999; Tyner 1998; ACRL 2000; Kapitzke 2000; Smith 2000).

Several search engines and functions help to locate the information needed, but the variety of different kinds of text and styles sometimes makes interpreting texts a challenge. Readers themselves create the structure of the text by using hyperlinks. It may be difficult to perceive what constitutes the whole text . One piece of information may be located on one page and another piece of information on another page, created by a totally different person, who is probably unaware of the other page. However, these pages are linked like chapters in a book. The reader needs to compare the texts and outputs to recognise the writers' intentions and to identify similarities and differences in the texts. Electronic communication particularly changes the kind of interpretation skills that are needed, as electronic texts have a great large number of acronyms and graphics, such as "smileys" (e.g. Danet, Ruedenberg & Rosenbaum-Tamari 1998; Laihanen 1999). It may be difficult to interpret whether the writer is serious, sarcastic or joking. Reading between the lines may even reveal that the writer means something completely different from what s/he says.

Reflecting and evaluating texts acquires a new meaning when it comes to electronic texts, because almost anyone can make a seemingly professional website, whose content, however, may be inaccurate or outdated. Readers must constantly evaluate the value, relevance and reliability of the texts. Furthermore, at school more attention should be paid to the evaluation of the quality and relevance of the content of web pages. In addition, ethical issues should be considered when electronic texts by other writers are quoted, in the same way as when printed texts are referred to.

In general, boys and girls use computers and the Internet in different ways. Boys use a computer to explore its possibilities. They want to know how things work, both as far as hardware and software are concerned. They enjoy being able to make computers do something that others cannot do, proudly presenting

their discoveries. Girls, on the other hand, see computers as a tool for contacting others, writing a short story and placing it on the web for everyone to read, and also as a source of discussion topics (Leino 2001).

6.3 Nordic students as active computer users

In the context of the PISA 2000 assessment, 20 of the 32 participating countries asked their students about the use of computers. In the Nordic countries, Denmark, Finland, Norway and Sweden used the questionnaire as an international option. Students were asked about their interest in computers, their self-assessment of their computer literacy skills, and their reasons for and frequencies of using computers.

In the Nordic countries, the Swedish and Norwegian teenagers were the most active computer users at home. More than 70 per cent of them used computers at home at least a few times per week. On the other hand, the Danish teenagers were the most active when it came to using computers at school. (OECD 2001.)

Figure 6.1 Percentage of the Nordic students who use the computer at least a few times per week for different activities

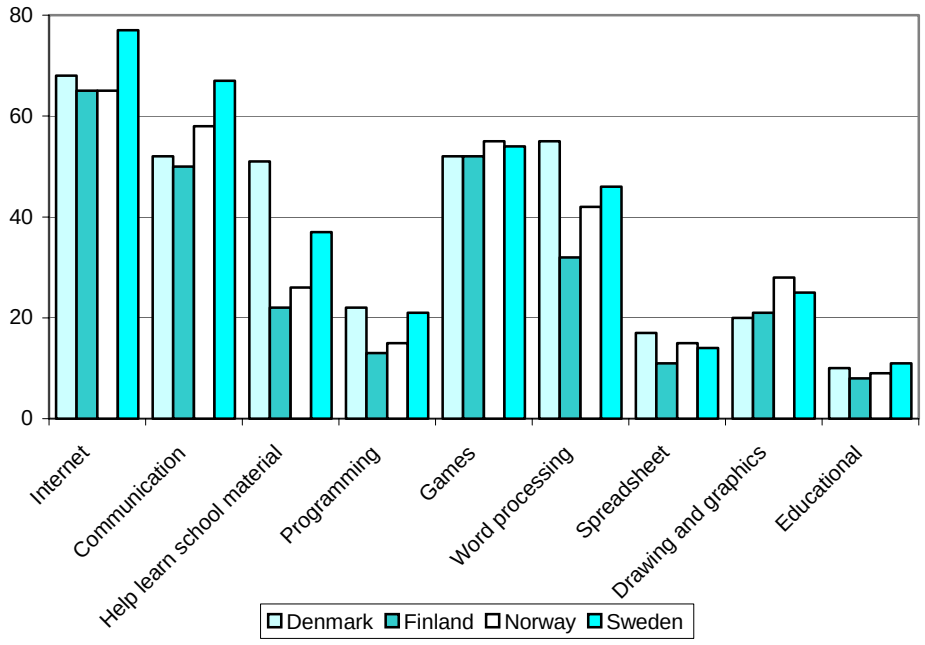


Figure 6.1 shows the percentage of the Nordic students who used the computer at least a few times per week for different activities. The Swedish teenagers were the most active Internet users. The percentage of those students who used the Internet almost every day was 48 in Sweden, 37 in Norway, 33 in Denmark, and 26 in Finland. However, if we consider those students who used the Internet at least a few times per week, the percentages covered a smaller range: from 65 in Finland and Norway to 77 in Sweden. In the weekly usage category, the average for all participating countries was only 23%, which shows that students in the Nordic countries are very active Internet users. So does the fact that the percentage of those who never use the Internet was only a few per cent in all the Nordic countries, whereas the average of all participating countries was 16%.

In addition to using the Internet, students are keen on using computers for communication, playing games and word processing. If computers are actively used at school, there is an effect on the frequencies of the activities that students undertake using computers. The Danish teenagers were the most active users of those computer activities which are part of teaching, such as word processing and spreadsheets. In Denmark, the students also used the computer more to help learn school material; linked to that, they also used more educational software than their counterparts in the other Nordic countries.

6.4 Gender and attitudes towards computers

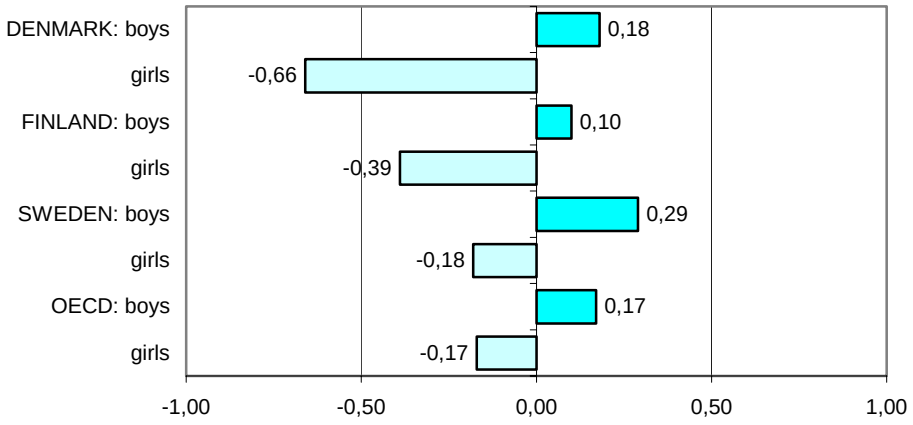
The index of *interest in computers* was derived from the students' responses to the following statements:

It is very important to me to work with computer.
To play or work with computer is really fun.
I use computer because I am very interested in this.
I forget the time, when I am working with computer.

The index is constructed with the average score across all countries set at 0 and the standard deviation set at 1. This means that a negative value does not mean a negative attitude, but that interest in computers is below the OECD average (OECD 2001). Figure 6.2 compares boys' and girls' interest in computers in Sweden, Finland and Denmark. (The actual questions were not asked in Norway.) In all those three Nordic countries, but also in most of the other participating countries, boys were more interested in computers than girls. Only in the United States and Mexico were girls (and only slightly) more interested in computers than boys. The gender difference in interest in computers was largest in Denmark, whereas Swedish boys and girls were those most interested in computers and their gender difference was also the smallest of the participating Nordic countries. The interest in computers of girls in all the Nordic countries was below the OECD average.

Many things can have an influence on how interested one is in computers. In the questionnaire, interest in computers was ascertained by means of only four questions. However, for example, an active use of and familiarity with computers may change computer usage from mere entertainment to more professional activity. In the same way, one's attitude to computers may change. Novelty and excitement diminish and the user will begin to notice problems and difficulties he may use the computer because it helps with projects and may even be an obligatory tool, yet it does not mean s/he is excited about it. For those who have not used computers often or who have only used them for a short time, the new tool and medium may still have the charm of novelty; there is so much to learn and so much one can do with it.

Figure 6.2 Boys' and girls' interest in computers in the Nordic countries



(OECD 2001)

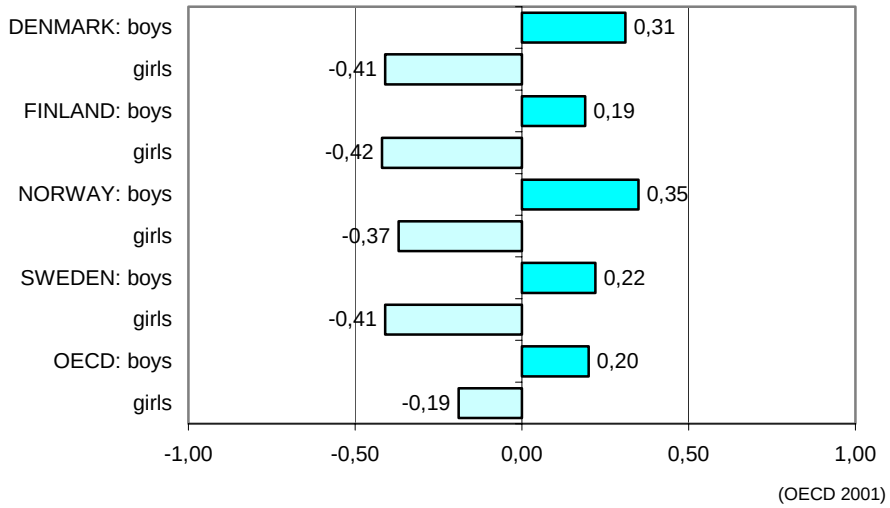
The index of *comfort with and perceived ability to use computers* was derived from students' responses to the following self-assessment questions:

How comfortable are you with using computer?
How comfortable are you with using a computer to write a paper?
How comfortable are you with taking a test on a computer?
If you compare yourself with other 15-year-olds, how would you rate your ability to use a computer?

This index is constructed with the average score across all countries set at 0 and the standard deviation set at 1 (OECD 2001). Figure 6.3 compares boys' and girls' comfort with and perceived ability to use computers in the Nordic countries. Boys, who were more interested in computers, were also more confident about their ability to use computers and more comfortable with them. This was the situation in every participating country. The Swedish and Norwegian students were the most active computer users; however, Norwegian students were more confident with their computer usage. On the other hand, the gender difference in confidence with computers was also largest in Norway

(and the same as Denmark), whereas in Finland the difference was the smallest of the Nordic countries but boys and girls did not feel as comfortable with using computers as in the other Nordic countries. In all the Nordic countries, girls' comfort with and perceived ability to use computers was below the OECD average, whereas for boys' it was the same or higher.

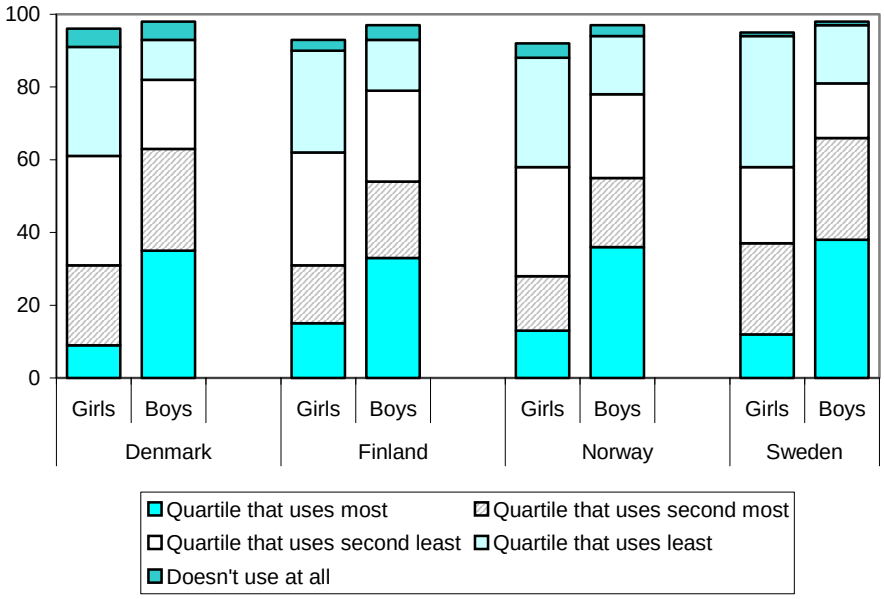
Figure 6.3 Comfort with and perceived ability to use computers among Nordic boys and girls



6.5 Computer usage and reading literacy

In order to be able to compare boys' and girls' computer usage and their results in reading literacy proficiency, a sum score of the use of computer (based on the activities illustrated in figure 6.1) was formed. The values of the variables were recoded (values *almost every day* and a *few times per week* recoded as 2, that is, the most active computer users; values *1 - 4 times per month* and *less than once a month* recoded as 1, relatively active; value *never* recoded as 0, not active) and a sum score was formed. With the help of this sum score, the students were divided into quartiles that represent activity in the use of computers. In addition, those who did not use computers at all and therefore did not answer the questions about the amount of computer usage were added to the analysis as an independent (fifth) group. As expected, boys were much more active computer users than girls. However, in both gender groups almost everyone used computers at least sometimes. The distributions for both genders are displayed in figure 6.4. (Due to some missing responses to individual questions the bars do not go to 100 %.)

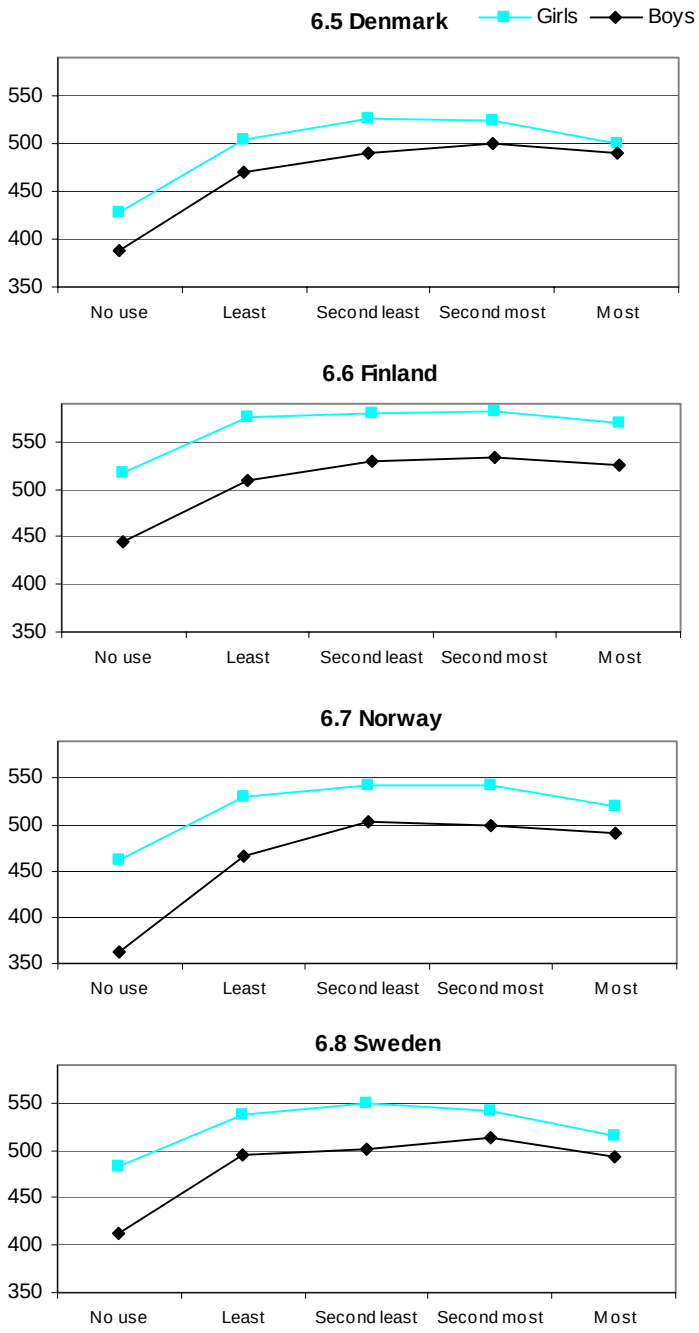
Figure 6.4 Distribution of computer usage for boys and girls in the Nordic countries



An examination of computer usage and the mean scores of reading literacy proficiency shows that those students who used computers performed better in the reading literacy assessment than those who did not use computers at all (see figures 6.5 – 6.8; the categories of computer usage are the same as in figure 6.4). The best performers, however, were those who used computers to a reasonable extent (not the most or the least actively). Those who never used computers had the lowest mean score. The same kind of result was presented in the Second International Adult Literacy Survey (SIALS) in Finland (Linnakylä et al. 2000: p 96 - 100). In addition, those boys who use computers to a reasonable extent had slightly better mean scores in retrieving information than the average (Leino 2002).

Figures 6.5 – 6.8 compare boys' and girls' mean scores in reading literacy relative to computer usage. In Finland, the difference between boys and girls was almost the same irrespective of how much they used computers. In Sweden, Norway and especially in Denmark, however, the smallest difference between the mean scores of boys and girls was found among those who were in the quartile that used the computer the most. In all the four Nordic countries, the difference in the mean scores between those who were in the quartiles that used the computer the most and the second most actively was larger among girls than boys. As for the mean scores of the boys, there was only a difference of approximately 10 points between the second least, the second most and the most active users. Likewise, in all these countries, the difference between boys and girls was largest among those who did not use computers at all.

Figure 6.5 – 6.8 Computer usage and reading literacy mean scores



6.6 Discussion

The results show that computers and the Internet are part of a 15-year-old Nordic student's everyday life. Nordic students, particularly the boys, were especially active in using computers. In addition to using the Internet and the communication possibilities it offers, teenagers also use computers actively for word processing and playing games. Although boys in general are more active computer users, the results in Finland show that girls are more active email users than boys and as active as boys in chatting and using other discussion forums. Boys use activities that demand more technical skills (Leino 2001).

Nordic boys were very interested in computers and confident and comfortable with their ability to use them. However, the interest of Nordic girls, as well as their self-assessment of their comfort with and perceived ability to use computers, was below the OECD average. The gender differences in these two aspects were generally larger in the Nordic countries than in the other participating countries on average. Naturally, the amount of usage is connected with confidence, but the gender difference may also be accounted for by the fact that computers and networks are still seen as boys' territory. There is an old prejudice according to which girls cannot use and understand computers. The self-assessment measure may be somewhat unreliable, as Nordic people in general are quite humble when it comes to self-assessing their skills.

The results show that there is a positive relationship between moderate computer usage and reading literacy skills. This can be seen in all the participating Nordic countries. The most active computer usage, however, seems to be related to a somewhat lower reading literacy score, particularly among girls. The difference between boys' and girls' mean scores was smallest among those who used computers the most actively. This can be specifically seen in Denmark, Norway and Sweden. The largest difference was to be found among those who did not use computers at all. However, these findings will need further investigation.

The most alarming situation is among those teenagers, be they boys or girls, who do not use computers at all. Their mean reading literacy score was clearly much lower than those who used computers at least sometimes. As some previous studies have shown, it is clearly not just a matter of using computers; those who did not use computers also read newspapers, magazines, comics, non-fiction and in most cases fiction less frequently than computer users (Leino 2002). Teenagers who do not use computers and therefore have poor multiliteracy skills are in danger of being marginalized in the information society, because in today's labour force technological knowledge and skills are in ever greater need. Workplace activities require traditional literacy skills as well as the multiliteracy skills necessitated by technology.

The increasing gender gap in reading literacy may be explained by the increase in computer use among boys leading to a decrease in reading of fiction. On the other hand, the purposes for which computers are used vary:

using the Internet and electronic communication channels has a more positive relationship to literacy proficiency than playing games or programming (Leino 2002). We need to create a pedagogy which enables everyone to use computers at least at school and, even more importantly, which also supports the literacy skills of those who use computers only now and then as well as of those who use them almost every day.

Electronic texts should not only be part of students' free time, they should also be a part of school material in every country, and should include not just word processing activities, but also reading, interpreting and evaluating authentic texts on web pages and in chat rooms. Teachers should prepare tasks for students involving retrieving information and critically evaluating the information in class or as a network discussion. Using electronic texts in teaching may be one way to get computer "nerds" interested in reading and to look beyond the surface of the texts.

References

ACRL (2000). *Information Literacy Competency Standards for Higher Education*. Association of College and Research Libraries. Online <URL: http://www.ala.org/acrl/ilintro.html> (18.4.2002)

Birkerts, S. (1996). *The Gutenberg Elegies. The Fate of Reading in an Electronic Age.* Faber and Faber, Inc.

Cope, B. & Kalantzis, M. (eds.) (2000). *Multiliteracies – Literacy Learning and the Design of Social Futures*. Routledge.

Danet, B., Ruedenberg, L. & Rosenbaum-Tamari Y. (1998). Hmmm…Where's That Smoke Coming From? Writing, Play and Performance on Internet Relay Chat. In F. Sudweeks, F., McLaughlin, F. & Rafaeli, S. (eds.): *Network & Netplay. Virtual groups on the Internet*. The MIT Press.

Gilster, P. (1997). *Digital Literacy*. John Wiley & Sons, Inc.

Kapitzke, C. (2001). Information literacy: The changing library. *Journal of Adolescent & Adult Literacy*, 44 (5), p 450 - 456.

Laihanen, K. (1999). *Suomenkielisten IRC-keskustelujen piirteitä. Kirjoittamalla välitettyä puhetta*. Master's thesis, University of Jyväskylä, Department of Finnish.

Leino, K. (2001). *Network Literacy Practices of 15-year-old Finnish Students - A functional view*. Presentation, European Conference on Educational Reseach 5. - 8.9.2001 Lille, France.

Leino, K. (2002). Tietotekniikan käyttö ja lukutaito. In Välijärvi, J. & Linnakylä, P. (eds.): *Tulevaisuuden osaajat. PISA 2000 Suomessa*. Kirjapaino Oma.

Linnakylä, P., Malin, A., Blomqvist, I. & Sulkunen, S. (2000). *Lukutaito työssä ja arjessa. Aikuisten kansainvälinen lukutaitotutkimus Suomessa*. ER-paino.

Luke, C. (1996). Ekstasis@cyberia. *Discourse*, 17 (2), pp. 187 - 208. Online <URL: http://www.gseis.ucla.edu/courses/ed253a/Luke/CYBERDIS.html> (16.4.2002)

Nunberg, G. (ed.) (1996). *The Future of the Book*. University of California Press.

OECD (1999). *Measuring Student Knowledge and Skills. A New Framework for Assessment. Paris: OECD Publications.*

OECD (2001). *Knowledge and Skills for Life. First results from PISA 2000*. Paris: OECD Publications.

Reinking, D., McKenna, M. C., Labbo, L. D., Kieffer, R. D. (eds.) (1998). *Handbook of Literacy and Technology. Transformations in a Post-Typographic World*. Lawrence Erlbaum Associates.

Smith, D. (2000). *Directory of Online Resources for Information literacy: Definitions of information Literacy and Related Terms*. Online <URL: http://nosferatu.cas.usf.edu/lis/il/definitions.html> (14.1.2002)

Tyner, K. (1998). *Literacy in a Digital World. Teaching and Learning in the Age of Information*. Lawrence Erlbaum Associates.

Wade, S. E. & Moje, E. B. (2000). The Role of Text in Classroom Learning. In Kamil, M. L., Mosenthal, P. B., Pearson, P. D. & Barr, R. (eds.): *Handbook of Reading Research. Volume III*. Lawrence Erlbaum Associates.

7 READING LITERACY AND HOME BACKGROUND

Torben Pilegaard Jensen and Are Turmo

7.1 Economic, cultural and social capital

7.1.1 Introduction

The relationship between socio-economic status and school performance has been given a lot of attention by researchers within the field of the sociology of education. Ho & Willms (1996) claim that perhaps the most enduring finding in the sociology of education is that schooling outcomes are related to the socio-economic status (SES) of the child's parents. According to Ho and Willms, most of the work in this field has been directed at determining the processes that contribute to this relationship, including structural processes at the level of the school, community, or larger society and processes at the micro-level associated with individual and group actions.

The concept of socio-economic status is also a strong focus of the PISA study. Socio-economic status is regarded as one of the strongest predictors for achievement in schools, and several questions in the student questionnaire aim to tap information about the students' home social background. The definition of SES in the PISA study (OECD 2001) is based on three sub-concepts, *economic capital*, *cultural capital* and *social capital*.

7.1.2 Economic capital

One common view is that differences in economic capital create differences in school performance. This view implicitly assumes that education is related to costs, and that well-off parents to a larger extent are able to cover such costs for their children. In the PISA study, the effect of financial resources cannot be assessed directly. However, the student questionnaire contains questions about the appearance of different objects in the student's home, and this information is used as an indication of economic status (OECD 2001). Information about the income levels of the parents can also be obtained indirectly from the parents' professions, as given by the students. Income and educational levels are derived from the professions using a system created by Hauser and Warren (1997). Previous research has suggested that economic resources are not among the most important explanatory factors for differences in school performance in

modern welfare states like the Nordic countries. The main reason for this is the obvious fact that family expenses related to the child's education are limited, due to a large degree of public financing.

7.1.3 Cultural capital

One of the other resources that is in focus in the PISA study, is often referred to as cultural capital, a concept borrowed from the French sociologist Pierre Bourdieu. According to Bourdieu's (1977, 1984) cultural reproduction theory, one would anticipate strong direct links between the parents' cultural backgrounds and student performance in many countries. One of the key aspects of cultural capital is language. According to Bourdieu, language is not simply an instrument of communication. It also provides, together with a richer or poorer vocabulary, a more or less complex system of categories, whether logical or aesthetic. The child's language depends partly on the complexity of the language transmitted by the family. This can be described as the students' language heritage (Bourdieu & Passeron 1990). Bourdieu draws the distinction between what could be labelled working-class language, which is often referred to as vulgar or common, and bourgeois language, which by contrast is said to be pure and correct. Bourgeois language is characterised by abstraction, formalism, intellectualism and euphemistic moderation. The authority of language is strongly related to pedagogic authority in schools. Bourgeois language practices are also strongly related to other high status cultural expressions like knowledge of classical literature and music. This knowledge together constitutes a person's cultural capital, and cultural capital is strongly related to values and preferences. The concept of "taste" is used by Bourdieu to describe the preferences of different social groups (Bourdieu 1984). As a consequence of this, an affinity for high status cultural expressions could be used as an indication of cultural capital, as is done in the PISA study. According to Bourdieu's theories, a lack of cultural capital is assumed to distance students from academic and school culture, which can often have consequences for the students' school careers and for the future of students facing exclusion and selection processes within the education system.

7.1.4 Social capital

A third type of resource is social capital. Coleman (1988) uses the concept of social capital as part of a general theoretical strategy, involving taking rational action as a starting point, but rejecting the extreme individualistic premises that often accompany it. According to Coleman, social capital is defined by its function. Like other forms of capital, social capital is productive, making possible the achievement of certain ends that in its absence would not be possible. Willms (2001) claims that, during the past decade, theorists have stressed that learning societies also depend on relationships among people, within both communities and organisations. They have invoked the term "social capital" to embody the nature of relationships among people, and how

these facilitate collective action, the strength of social networks, and the norms and values of a community. Social capital refers to resources in the form of social ties that can be used in different situations for different purposes, for example in relation to the children's school career. The traditional hypothesis about social capital is that students do better in school if they have a close social network surrounding them where parents, children and teachers collaborate and know each other well. Social capital is assumed to be particularly important for individuals who possess relatively little economic and cultural capital, in some way compensating for their relative disadvantage.

7.2 Measuring socio-economic status

In PISA, several constructs related to socio-economic status have been derived from the student questionnaire. For all these composite variables the internal consistency measured by Cronbach's alpha is at or above 0.70. Cronbach's alpha is a measure of internal consistency based on item covariances (Crocker & Algina 1996). In table 7.1 the constructs are grouped according to their relationship to the three forms of capital outlined above. This classification should be regarded as one possible suggestion, and it is clear that in certain ways the validity of classification could be questioned, as will be touched upon later. However, factor analysis of the constructs also supports the classification presented.

Table 7.1 Classification of constructs in relation to the three forms of capital

Cultural capital	Social capital	Economic capital
Parental education	Home social capital	Home economy
Highest family socio-economic index		Highest family socio-economic index
Home cultural possessions		
Home educational resources		
Books at home		
Home cultural competence		
Student's cultural activity		

All the concepts in table 7.1 are measured using composite variables, except "Books at home", which is a single variable, and "Parental education", which is a simple combination of two single items. In the table, the construct "Highest family socio-economic index" appears twice. This is due to the fact that the International Socio-Economic Index (ISEI) used, represents a combination of education and income, and education and income are related to cultural and economic capital, respectively.

Parental education

The student questionnaire contained explicit questions about the parents'

education. It could reasonably be argued that the parents' education theoretically should be more strongly related to student achievement than the parents' professions. There are reasons to believe that parents with a high income level, but a relatively limited education, will place less value on education as being important for their children's working career than parents with a high level of education, but limited financial resources. Parents' education is strongly related to the concept of cultural capital. In PISA, the ISCED (International Standard Coding of Education) system has been used to make international comparisons possible. The six categories have been translated into the proper terms in all the participating countries.

Highest family socio-economic status

The concept of "Highest family socio-economic index" needs further clarification. The highest family socio-economic index is derived from the parental occupations given in the student questionnaire. These questions are asked about the mother's profession (the same questions are asked about the student's father):

What is your mother's main job? (e.g. school teacher, nurse, sales manager). If she is not working now, please tell us her last main job. Please write in the job title.

What does your mother do in her main job? (e.g. teaches high school students, cares for patients, manages a sales team). If she is not working now, please tell us her last main job. Please use a sentence to describe the kind of work she does or did in that job.

The classification of the parents' occupations is done using a system called ISCO (International Standard Classification of Occupations). This is a system designed for classification of occupations across countries. Ganzeboom & Treiman (1996) developed an algorithm where each ISCO category is related to a numerical value for socio-economic status on the so-called ISEI scale (International Socio-Economic Index). The ISEI index gives a metric for the sum of education level and income level for each occupational category. This has been quantified based on international empirical data. The ISEI scale ranges from 0 to 90, and the level of socio-economic status increases with increasing values. The different profession categories in the ISCO system have been given ISEI values ranging from 16 to 90.

Home cultural possessions

This construct consists of three single items:

In your home, do you have: classical literature (e.g. Shakespeare)?/ books of poetry?/works of art (e.g. paintings)? (Yes or *No)*

This construct focuses on possessions in the student's home as a measure of cultural capital. All the possessions are strongly related to the concept of cultural capital, and the construct could be characterised as a very valid measure of this concept.

Home educational resources

This construct consists of these single items:

In your home, do you have: a dictionary/ a quiet place to study/ a desk for study/ text books? (Yes or *No).*
How many calculators do you have at home? (None, One, Two, Three or more)

This construct focuses on home resources that are directly useful for the student's schoolwork, and in table 7.1 the construct was classified in the category "Cultural capital". These aspects of the home environment could be said to indicate an academic orientation, and the construct could therefore reasonably be classified under the label "Cultural capital".

Books at home

As stated before, this is not a construct, but a single item:

*How many books are there in your home? (*From *None* to *More than 500)*

In previous studies, the number of books at home has been shown to be strongly related to student achievement, for example in the Third International Mathematics and Science Study (TIMSS) (Lie et al. 1997). This is the case even though there is no differentiation between different kinds of books. To be consistent with the concept of cultural capital, the type of books should be significant, i.e. popular literature vs. classical literature.

Home cultural competence

The construct "Home cultural competence" consists of three single items:

In general, how often do your parents: discuss political or social issues with you/ discuss books, films or television programmes with you/ listen to classical music with you? (Never or hardly ever, A few times a year, About once a month, Several times a month, Several times a week).

As the title of the construct indicates, this construct is intended to measure the level of cultural competence in the student's home. In a later phase of the process of developing constructs, this construct was renamed "Parental Academic Interest". As is evident, the construct also incorporates a significant element of social interaction and communication between the student and the parents. In this way the construct differs from another construct related to cultural capital that will be presented later, "Home cultural possessions", which is more distinctly related to "Cultural capital". The construct "Home cultural competence" could therefore also be said to measure a component of social capital. It could be argued that this construct is not very precisely linked to Bourdieu's concept of cultural capital. Listening to classical music is the only activity that is a typical indication of cultural capital. By contrast, discussing TV programmes is not a very typical indication of cultural capital in Bourdieu's sense. This single item is perhaps a better indication of social capital, as defined earlier in the chapter.

Student's cultural activity

The construct "Student's cultural activity" consists of three single items:

During the past year, how often have you participated in these activities:

*visited a museum or art gallery/ attended an opera, ballet or classical symphony concert/ watched live theatre? (*From *Never or hardly ever* to *Several times a week)*

This construct focuses on the student's own cultural activities. All the activities that are included in the construct, are typical activities related to Bourdieu's concept of cultural capital. Even though the construct clearly focuses on the student's own activities, there are strong reasons to believe that the tendency among 15-year-olds to attend these kinds of activities is strongly linked to parental preferences and practices. Therefore, it could be argued that the construct is also a proper measure of home cultural capital.

Home social capital

The construct "Home social capital" consists of three single items:

In general, how often do your parents: discuss how well you are doing at school/ eat the main meal with you around a table/ spend time just talking to you? (From *Never or hardly ever* to *Several times a week)*

This construct is designed to explicitly measure home social capital. One of the single items "Discuss how well you are doing at school?" particularly relates to the findings of Ho & Willms (1996) as presented earlier in the chapter. They found that discussing school activities and helping children plan their school programmes had the strongest relationship to academic achievement.

Home economy

One construct is used to measure home economic capital. The construct "Home economy" consists of these single items:

In your home, do you have: a dishwasher/ a room of your own/ educational software/ a link to the Internet? (Yes or *No)*

How many of these do you have at home: cellular phone/ TV/ computer/ motor car/ bathroom?(None, One, Two, Three or more).

7.3 Empirical results

7.3.1 Introduction to the analyses

Analysis of the relationship between student reading literacy and the different constructs measuring aspects of social background may be carried out in different ways. The simplest is to look at each construct separately. The individual empirical relationships between the SES constructs and reading literacy are given under each of the three forms of capital. As several of the constructs are highly correlated (see correlations in the appendix table), the effect of a single construct can easily be overestimated. Therefore, after presenting the relationship between each construct and reading literacy, multiple regression will be used to estimate the total effect of all constructs together. It has to be stressed that the relationships presented are statistical relationships that do not necessarily imply cause-effect relationships.

Nevertheless we expect that the results may contribute to the discussion about which mechanisms that are at work in the process of social reproduction.

Some of the constructs included in the analyses measure "objective" aspects of the cultural and economic capital in the family: parental education, highest family socio-economic index and home economy. Other constructs could be considered as transmission factors, e.g. the mechanism by which the capital is transmitted from parents to children. These constructs are: home cultural possessions, home educational resources, books at home, home cultural competence, home social capital and student's cultural activity. Other social background variables could have been included, for example language spoken at home and type of family.

The results will be presented in the following section. The overall impression is that between the Nordic countries there are common features, but also interesting disparities. We will look at the relationship between the different constructs and reading literacy in the Nordic countries for cultural capital, social capital and economic capital. We will focus on differences and similarities between the Nordic countries. Finally we will analyse the relationship between reading literacy and all SES constructs taken into account simultaneously.

7.3.2 Cultural capital

What differences or similarities do we find, looking at cultural capital? To summarise, cultural capital consists of the following variables: parental education, highest family socio-economic index, home cultural competence, student's cultural activity, home cultural possessions, home educational resources and books at home.

We start out with parental education and reading literacy. Table 7.2 shows the mean, the 5th and 95th percentile for parental education and its correlation with reading literacy. The table also contains the regression coefficient, which shows the effect of an increase of one step in parental education (six categories) on the reading literacy score.

Table 7.2 Reading literacy and parental education (six categories)

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	5.1	3.0	6.0	0.34	28.1
Finland	4.5	2.0	6.0	0.18	11.7
Iceland	4.7	2.0	6.0	0.17	12.2
Norway	5.2	3.0	6.0	0.14	12.5
Sweden	5.3	3.0	6.0	0.14	12.1
OECD	4.7	2.0	6.0	0.28	20.6

It appears that parental education is positively related to students' reading skills in all countries, particularly in Denmark, where the correlation between

parental education and reading literacy is high (0.34), and the regression coefficient is as high as 28.1. This means that an increase of one category in parents' education (out of six categories) increases the students' reading score by 28.1 points. This is more than the OECD average, and much more than in the other Nordic countries, where the effect of a one-step increase is about 12 points.

In the Nordic countries, and in OECD countries on average, education plays a large role in the type of work people do (see the correlation matrix in the appendix table). The socio-economic index indicates where the parents are placed in the hierarchical occupational structure. And where the parents are placed also predicts their childrens' reading skills, with the correlation being best in Norway, Sweden and Denmark, where the effect is nearly at the same level as the OECD average. This is shown in table 7.3.

Table 7.3 Reading literacy and highest socio-economic index

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	49.7	25.0	73.0	0.28	1.6
Finland	50.0	23.0	74.0	0.21	1.2
Iceland	52.7	25.0	77.0	0.20	1.1
Norway	53.9	30.0	78.0	0.25	1.7
Sweden	50.6	26.0	77.0	0.28	1.6
OECD	48.9	23.0	74.0	0.32	1.9

The mean effect of an increase of one unit in the socio-economic index – the values being from 0 to 90 – in the OECD is 1.9 scale points on the reading literacy scale. The corresponding effect is lower in all the Nordic countries, particularly in Iceland and Finland.

Table 7.4 shows how home cultural possessions can predict reading literacy. Here we see the strongest relationship in Norway, where an increase of one in the measure of cultural possessions means an increase of 27.8 scale points on the reading scale. The smallest effects are found in Finland and Iceland, where the corresponding figure is about 20 scale points. Although cultural possessions are more common among well-educated parents (see the appendix table), less well-educated parents may also possess this cultural capital and in this way may stimulate their children. The point is that there is obviously a freedom to act.

Among the Nordic countries we find great differences in the amount of cultural possessions (see table 7.4, "Mean"). In Denmark the level of cultural possessions is lower than the OECD average, whereas it is much higher in Iceland. Families with many cultural possessions also tend to have many home educational resources (see the appendix). Here also the large differences between the Nordic countries are striking. In Denmark the mean home educational resources are much lower than in the other Nordic countries.

Table 7.4 Reading literacy and home cultural possessions

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	-0.11	-1.65	1.15	0.24	24.4
Finland	0.12	-1.65	1.15	0.22	19.8
Iceland	0.67	-0.62	1.15	0.17	20.5
Norway	0.14	-1.65	1.15	0.28	27.8
Sweden	0.05	-1.65	1.15	0.26	24.2
OECD	0.00	-1.65	1.15	0.25	25.3

Home educational resources seem to affect reading literacy in different ways in the Nordic countries (see table 7.5). In Norway in particular home educational resources are seen to be an important predictor for reading literacy, as an increase of one unit increases the reading literacy score by 31.7 scale points, compared with an average of 22.9 in the OECD. In Denmark the effect is similar to the OECD average, whereas it is much lower in the other Nordic countries.

Table 7.5 Reading literacy and home educational resources

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	-0.22	-2.00	0.76	0.21	21.5
Finland	0.00	-2.00	0.76	0.12	11.3
Iceland	0.20	-1.33	0.76	0.10	10.2
Norway	0.10	-1.54	0.76	0.29	31.7
Sweden	0.03	-2.00	0.76	0.13	12.7
OECD	0.00	-2.00	0.76	0.23	22.9

“Books at home” (the variable consists of seven categories) has in all countries a significant co-variation with reading skills, but there are small differences between the Nordic countries as illustrated in table 7.6.

Table 7.6 Reading literacy and books at home

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	4.58	2.00	7.00	0.33	20.7
Finland	4.34	2.00	7.00	0.24	15.5
Iceland	5.05	3.00	7.00	0.25	16.5
Norway	4.87	2.00	7.00	0.29	19.2
Sweden	4.88	2.00	7.00	0.32	20.2
OECD	4.48	2.00	7.00	0.35	22.1

The effect of "books at home" is similar to the OECD average in Sweden, Denmark and to some extent in Norway, and weakest in Finland and Iceland. "Books at home" acts as a relatively strong indicator for cultural capital at home. It has strong relationships to other indicators for cultural capital, e.g. home cultural possessions, whereas there is no strong relationship between "books at home" and, for example, social capital (see the appendix table), measured here by social communication in the family.

Home cultural competence plays an important role (see table 7.7). And it is worth noting here that there is no strong correlation between parental education, socio-economic status and home cultural competence (see the appendix). This means that factors other than parental education may explain the level of home cultural competence.

Table 7.7 Reading literacy and home cultural competence

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	0.11	-2.20	1.36	0.32	31.4
Finland	-0.01	-2.20	1.13	0.23	24.5
Iceland	0.08	-2.20	1.61	0.19	17.7
Norway	-0.22	-2.20	1.13	0.27	27.7
Sweden	-0.14	-2.20	1.13	0.22	21.8
OECD	0.00	-2.20	1.36	0.20	19.3

In Denmark and Norway an increase of one unit rises the reading literacy score by about 30 scale points, much more than in OECD countries on average. In Iceland and Sweden the values are only about 20, close to the OECD average. These results indicate that the Danish and Norwegian schools do not succeed particularly well in levelling the influence of differences in cultural competence between families.

Students' cultural activity is also positively related to reading literacy, mostly so in Denmark (see table 7.8). In Sweden and Finland the effect is smaller and below the OECD average. When we look at mean levels of student's cultural activity it is interesting to see the differences between the Nordic countries, e.g. that the level is half a standard deviation higher in Denmark than in Norway. This indicates that cultural activity among students is much higher in Denmark than in Norway and in the OECD on average.

Table 7.8 Reading literacy and students' cultural activity

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	0.31	-1.28	1.55	0.22	23.9
Finland	-0.16	-1.28	1.29	0.16	14.8
Iceland	0.21	-1.28	1.55	0.21	20.9
Norway	-0.21	-1.28	1.29	0.18	19.0
Sweden	-0.13	-1.28	1.29	0.15	13.8
OECD	0.00	-1.28	1.55	0.18	17.2

The above results have shown the simple relationships between the single constructs included in cultural capital and reading literacy. As several of the constructs are highly correlated, it is easy to overestimate the effect of a single construct. Therefore we have used a multiple regression analysis which includes all the constructs of cultural capital. This makes it possible to evaluate in general how much cultural capital predicts reading literacy in the different Nordic countries (see table 7.9).

Table 7.9 Explained variance: Cultural capital

Country	R^2
Denmark	0.20
Finland	0.11
Iceland	0.12
Norway	0.17
Sweden	0.15
OECD	0.18

As will become apparent in the following section, cultural capital is a much more important factor in explaining variations in reading literacy than economic and social capital. In Denmark and Norway the level of correlation is about the OECD average, whereas cultural capital in Finland and Iceland predicts reading literacy considerably less well.

7.3.3 Social capital

In the analysis presented here there is only one construct in this category, "Home social capital". Students are asked how often their parents discuss how well they are doing at school, eat the main meal with them, and spend time just talking to them. It is open to discussion whether this variable describes social capital in the way it is defined by Coleman (1988). Table 7.10 shows a positive relationship between social capital and reading performance.

Table 7.10 Reading literacy and home social capital

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	0.20	-1.16	1.20	0.19	20.4
Finland	-0.20	-1.34	1.20	0.06	6.4
Iceland	-0.09	-1.53	1.20	0.11	10.9
Norway	-0.01	-1.53	1.20	0.14	15.2
Sweden	-0.04	-1.34	1.20	0.05	4.9
OECD	0.00	-1.53	1.20	0.10	9.4

We find large disparities between the Nordic countries, both concerning the mean value of social capital and how it predicts reading literacy. Between Finland and Denmark there is a difference of nearly half a standard deviation in the mean value of home social capital. This means that Danish parents seem to be communicating much more with their children than Finnish parents do. The amount of home social capital is around the OECD average in Iceland, Norway and Sweden. As can be seen in table 7.10, the effect of home social capital is much stronger in Denmark and Norway than in the other countries.

Compared to the effect of home cultural competence on reading literacy, the influence of home social capital is much smaller. It looks as if home social capital, defined by aspects of the social communication in the family, does not stimulate reading ability to the same extent as the cultural communication in the family. The small effect of home social capital, especially in some Nordic countries, may be caused by the fact that social communication includes the question of how often the parents discuss how well the student is doing at school. If this is more common in families where the students have problems in school, we have one explanation of the weak relationship between home social capital and reading literacy.

7.3.4 Economic capital

One could expect a strong co-variation between socio-economic status, parental education and the home economy (indicated, for example, by having a dishwasher, a room of your own, and a link to the Internet, televisions, computers and motor cars). But in the Nordic countries this is not the case, and we see no strong relationship between home economy and reading skills as we do in the OECD countries as a whole (see table 7.11). Both the correlations and the coefficients are small in the Nordic countries, and in Iceland they have a small negative value. Compared to the OECD average it is striking how weak the relationship between wealth and reading skills is in the Nordic countries. The table also shows that the level of home economy in the Nordic countries is high.

Table 7.11 Reading literacy and home economy

Country	Mean	5 %	95 %	Correlation	Regression coefficient
Denmark	0.49	-0.74	1.84	0.09	11.0
Finland	0.22	-1.13	1.46	0.09	11.3
Iceland	0.53	-0.67	1.84	-0.05	-5.7
Norway	0.56	-0.60	1.84	0.03	3.8
Sweden	0.65	-0.60	1.84	0.07	7.9
OECD	0.00	-1.84	1.46	0.19	18.7

In the definition of economic capital we have included the construct highest family socio-economic index (see table 7.1). Here also we have used a multiple regression analysis which included the two constructs: home economy and highest family socio-economic index. Table 7.12 shows the extent to which economic capital in total predicts reading literacy in the different Nordic countries.

Table 7.12 Explained variance: Economic capital

Country	R^2
Denmark	0.08
Finland	0.05
Iceland	0.05
Norway	0.07
Sweden	0.08
OECD	0.11

Including both the highest socio-economic index and home economy in the regression model analysis of the effect of economic capital on reading literacy confirms that economic capital predicts reading literacy less well in the Nordic countries than in the OECD on average.

7.3.5 A total model

After presenting the effects of the three types of capital on reading literacy, we will summarise the results using a total regression model, which includes all the home background constructs. As shown in table 7.13, the explanatory power (the percentage of explained variance) of the model varies from 11% to 20% in the Nordic countries, compared with the OECD average at 18%. Home background has approximately the same effect on reading literacy in Denmark and Norway as in the OECD countries on average, whereas the effect is minor in the other Nordic countries, especially in Finland. There are many possible explanations for these differences between the Nordic countries; one possible explanation is differences in the school systems in the Nordic countries. The

relationship between home background and reading skills is thus mediated by the structure of the school system.

Table 7.13 Reading skills among 15-year-old students explained by home background (cultural, social, and economic)

Country	R^2
Denmark	0.20
Finland	0.11
Iceland	0.13
Norway	0.18
Sweden	0.15
OECD	0.18

7.4 Conclusion

The results have shown that there are important differences regarding how much of the variance in reading literacy the three dimensions can explain. In statistical analysis it is obvious that co-variances between two or more variables do not necessarily imply direct cause-effect relationships. However, if theory and previous research tell us that there is a mechanism for a cause-effect relationship between the variables, then there is strong support for interpreting the findings in this way. The theoretical framework presented in the introductory parts of this chapter implies that there should be a cause-effect relationship between cultural capital and student levels of reading literacy. The same should be the case for social capital. However, it has to be stressed that the mechanisms at work will be different for different indicators of the three forms of capital. In the case of economic capital, one would not anticipate a strong relationship with reading literacy in modern welfare states like the Nordic countries. In the PISA data, we have seen that there are strong relationships between cultural and social capital and levels of reading literacy, while the relationships with economic capital are weak. In this respect, the findings are consistent with the theories presented in the introductory parts of the chapter.

In general, the Nordic societies are often perceived as countries with a high level of equality. Based on this, one would anticipate that the relationship between SES and levels of reading literacy would be weak when viewed from an international perspective. The results from PISA do not support such a picture. In many of the Nordic countries there are surprisingly strong relationships, strongest in Denmark and Norway and weakest in Finland and Iceland, between students' reading literacy and their overall SES.

In Denmark and Norway *cultural capital* has nearly the same effect on students' reading literacy as in the OECD as a whole, while the effect is smaller in Finland and Iceland (table 7.9). In general the individual constructs within *cultural capital* can predict reading literacy about equally well. Some of

them will be commented on briefly here. In the OECD countries on average, *parents' education* is positively related to students' reading skills. In the Nordic countries we see the same tendency, although the effect is somewhat stronger in Denmark. The parents' type of work – *socio-economic index* – is not as good a predictor of reading literacy in the Nordic countries, particularly in Finland and Iceland, as in the OECD countries generally. The availability of *educational resources* is a good predictor of reading skills in Norway, better than in the OECD countries on average. In Denmark the co-variation is at the average level for the OECD, whereas it is lower in the other Nordic countries. In all the Nordic countries except Iceland, the indicator *cultural competence* has a stronger effect on reading literacy than in the OECD countries combined. The effect is strongest in Denmark but in Norway it also plays an important role. *Students' cultural activities* are positively related to student literacy skills in both OECD on average and in the Nordic countries. In Denmark and Iceland, where the level of cultural activity is also high, the effect is particularly strong.

As there is no determinant relationship between the more objective constructs of cultural capital (parents' education and socio-economic index) and the behaviours of the family (e.g. cultural possessions, home cultural competence), the conclusion is that the way the family acts has its own influence on the students' reading literacy. Cultural possessions, books at home and cultural communication apparently stimulate the students' ideas and motivation for reading and through that their reading skills.

S*ocial capital* generally has a much weaker relationship with reading literacy than cultural capital. In Denmark and Norway, however, social capital plays a relatively large role, and only in Finland and Sweden is the effect of social capital on reading literacy smaller than in the OECD countries on average.

Economic capital is also positively related to reading literacy. However, this relationship is weak in the Nordic countries, particularly in Finland and Iceland.

What do these findings mean for reading education in the Nordic countries? If we assume that the main source of student reading literacy is instruction in school, reading education in several of the Nordic countries seems to be better adapted to students from homes with high cultural capital than students from homes with low cultural capital. An unequal distribution of reading literacy levels between different groups of the population can be a serious problem for democracy. An important question arises: How can we in reading education better meet the needs of students with a home culture that is not consistent with the school culture?

According to Bourdieu's (1977, 1984) theories, lack of home cultural capital distances students from academic or school culture. The mechanisms at work can be formulated as follows: Congruent worlds support smooth transitions between the cultures, different worlds require transitions to be managed, diverse worlds lead to hazardous transitions, and highly discordant

worlds cause students to resist transitions which therefore become virtually impossible.

The empirical results presented in this chapter show that there are important issues about how to motivate and meet the needs of students from families with lower levels of cultural capital in particular in the Nordic countries as elsewhere.

References

Bourdieu, P. (1977). Cultural reproduction and social reproduction. In Karabel, J. & A.H. Halsey (eds.). *Power and ideology in education.* Oxford University Press.

Bourdieu, P. (1984). *Distinction: A Social Critique of The Judgement of Taste.* Harvard University Press.

Bourdieu, P. & J.C. Passeron (1990). *Reproduction in Education, Society and Culture. Second Edition.* Sage Publications.

Coleman, J.S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94, 95-140.

Crocker, L. & J. Algina (1996). *Introduction to Classical and Modern Test Theory.* Harcourt Brace Jovanovich College Publishers.

Ganzeboom, H.B.G. & D.J. Treiman (1996). Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25, 201-239.

Hauser, R.M. & J.T. Warren (1997). Socioeconomic Indexes for Occupation: a Review, Update and Critique. *Sociological Methodology*, 27, 177-298.

Ho, E. & J.D. Willms (1996). The effects of parental involvement on eight grade achievement. *Sociology of Education*, 69, 126-141.

Lie, S.; M. Kjærnsli & G. Brekke (1997). *Hva i all verden skjer i realfagene?* Dept. of Teacher Education and School Development, University of Oslo.

OECD (2001). *Knowledge and skills for life. First results from PISA 2000.* Paris: OECD Publications.

Willms, D. (2001). Three hypotheses about community effects. In Helliwell, J. (ed.): *The contribution of investment in human and social capital to sustained economic growth and well-being.* Human Resources Development Canada.

Appendix table: Correlations between the constructs

Denmark **Finland** **Iceland** **Norway** **Sweden** **OECD**	Cultural compe-tence	Cultural activity	Cultural poss-essions	Home educa-tional resources	Books at home	Highest socio-econo-mic index	Social com-muni-cation	Home eco-nomy
Cultural activity	0.37 0.33 0.31 0.32 0.33 0.33							
Cultural possessions	0.33 0.27 0.25 0.32 0.27 0.30	0.34 0.29 0.26 0.28 0.29 0.35						
Home educational resources	0.26 0.16 0.18 0.24 0.16 0.18	0.20 0.14 0.16 0.16 0.14 0.17	0.37 0.30 0.30 0.43 0.30 0.29					
Books at home	0.31 0.23 0.21 0.26 0.23 0.24	0.32 0.27 0.22 0.22 0.27 0.28	0.47 0.46 0.41 0.48 0.46 0.46	0.34 0.27 0.22 0.34 0.27 0.30				
Highest socio-economic index	0.23 0.14 0.15 0.19 0.14 0.18	0.23 0.19 0.23 0.19 0.19 0.22	0.34 0.29 0.22 0.34 0.29 0.29	0.24 0.17 0.12 0.21 0.17 0.20	0.35 0.33 0.21 0.31 0.33 0.36			
Social commu-nication.	0.41 0.39 0.47 0.43 0.39 0.43	0.18 0.13 0.20 0.14 0.13 0.19	0.17 0.14 0.19 0.21 0.14 0.20	0.18 0.14 0.23 0.25 0.14 0.19	0.13 0.07 0.14 0.17 0.07 0.13	0.14 0.04 0.12 0.11 0.04 0.07		
Home economy	0.05 0.06 0.06 0.06 0.06 0.07	0.05 0.06 0.07 0.04 0.06 0.07	0.18 0.19 0.17 0.20 0.19 0.20	0.29 0.29 0.29 0.29 0.29 0.31	0.16 0.20 0.12 0.17 0.20 0.27	0.26 0.26 0.16 0.19 0.26 0.32	0.06 0.05 0.10 0.09 0.05 0.02	
Highest education (years)	0.24 0.14 0.17 0.17 0.14 0.15	0.17 0.17 0.23 0.15 0.17 0.19	0.29 0.23 0.26 0.28 0.23 0.27	0.26 0.15 0.20 0.18 0.15 0.19	0.32 0.27 0.26 0.28 0.27 0.36	0.44 0.47 0.48 0.41 0.46 0.47	0.13 0.06 0.14 0.10 0.06 0.05	0.16 0.21 0.16 0.17 0.21 0.28

8 SELF-REGULATED LEARNING

Erik Knain and Are Turmo

In this chapter, results from the CCC (Cross-curricular competencies) on aspects considered important to lifelong learning are presented. A brief outline of the theoretical 'landscape' around CCC will be presented first. This theoretical frame will be used in the discussion of the results.

8.1 Background

By the end of the 1980s it was recognised that the OECD's indicators on educational outcomes in science, mathematics and reading only covered *some* important parts of the students' learning outcomes. In most countries, one of the important aims of schooling is to develop competencies that go beyond the specific knowledge involved in the different school subjects. Against this background, the OECD found it interesting to develop indicators of Cross Curricular Competencies (CCC). The CCC instrument that became part of PISA has the concept of "self-regulated learning" as a part of lifelong learning as an overarching theme. To be able to learn for the rest of their life, students need certain motivational, cognitive and metacognitive dispositions.

Defining 'self-regulated' learning is not a straightforward matter. As will be discussed below, self-regulated learning has been studied in several research traditions, and it is hard to find an agreed definition. Normally, definitions of self-regulated learning (usually 'SRL' hereafter) tend to focus on factors that are important for SRL. Pintrich has offered the following definition:

> *Self-regulated learning [is] an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features in the environment (Pintrich 2000, p 453).*

SRL is a dynamic process whereby students orchestrate knowledge, skills and attitudes when learning in specific contexts. The student needs a range of learning strategies in order to do this, but also experience in using them in various situations, and the ability to reflect on their effectiveness. But it is not enough to know the strategies if the student is not motivated to use them, or does not see their relevance. SRL is therefore also related to the students' self-concepts, in particular to efficacy beliefs and how successes and failures are experienced and explained. SRL extends beyond a matter of developing individual skills. Students develop learning strategies in a social setting and develop self-concepts and identities as learners as part of the social interaction.

In this chapter data on these issues will be presented and discussed for the Nordic countries. A more detailed discussion of Norwegian data is given in Knain (2002). We will present both descriptive statistics for the constructs and their correlations with reading literacy. We have chosen to focus only on the score in reading literacy, and not on the scores in science and mathematics. There are two reasons for this. In PISA 2000 reading literacy was the main domain and the largest part of the test time was used for this domain. The reading literacy score is therefore the most reliable of the three score values. Analyses of the results also show that there are in general only marginal differences between the three domains regarding the empirical relationships with the CCC constructs.

Each of the CCC constructs presented in this chapter is derived from 3-5 single items in the student questionnaire. The items that require a statement of how often something happens have the answer alternatives "almost never", "sometimes", "often" and "always". The questions where the student is asked to state to what extent s/he agrees or not have the answer alternatives "disagree", "disagree somewhat", "agree somewhat" and "agree".

CCC measures certain psychological dispositions and attitudes. When comparing means between countries, one should always be aware of cultural influences; any difference in mean between countries may reflect different traditions of how one responds to such questions rather than a real difference in the traits that are measured. This is certainly a cause for concern for the data presented in this chapter, but we believe that our discussion of the data overcomes the problem for three reasons: Firstly, the Nordic countries are more homogeneous culturally than would be expected if five countries taking part in the CCC were sampled at random. Secondly, we will focus on the largest differences between means. Thirdly, measures of means will be supplemented with correlations between constructs. Correlations are expected to be less culturally sensitive than the mean, and are also important for connecting CCC to its theoretical 'environments'. Nevertheless, the cultural factor must be kept in mind.

8.2 The CCC aspects

8.2.1 Aspects and constructs

The constructs related to Cross-Curricular Competencies in PISA can be grouped into four categories, or aspects: learning strategies, motivation, self-concept and learning style. The constructs within each aspect are presented in table 8.1 below. Whereas the constructs are meaningful entities both theoretically and empirically, the aspects should be considered more as 'buckets' of constructs that are more or less related, and each focus on a certain aspect that is considered important to SRL.

Table 8.1 Constructs within aspects

Aspect	Construct
Learning strategies	Memorisation Elaboration Control strategies
Motivation	Instrumental motivation Interest in reading Interest in mathematics Effort and perseverance
Self-concept	Self-efficacy Verbal self-concept Self-concept in mathematics Academic self-concept
Learning style	Co-operative learning Competitive learning

Although a comprehensive model of SRL does not exist, there are certainly relationships between constructs and across aspects that both theoretically and empirically are considered important in the SRL field. Boekaerts (1999) offers a three-layered model of self-regulated learning where each layer represents the contribution from the different schools of thought that have contributed to the current understanding of SRL. Constructs in the CCC instrument can be placed in the model in all three layers, and the model is therefore described below (from Boekaerts 1999, p 447–454), and then related to CCC.

1. The search for learning or processing styles. A key issue in this tradition is the students' ability to select, combine and coordinate cognitive strategies in an effective way. The goal has been to identify ways students process academic knowledge. For instance, some students were found to mainly rehearse and memorise the study material ('shallow style'), while others spontaneously related ideas and arguments expressed by others to their own experiences and evidence ('deep-processing style'). These are essentially the 'Memorisation' and 'Elaboration' constructs that are learning strategies in the CCC. Identifying these strategies is important, but not sufficient. It is also important that students experience that they have a choice among different strategies, and are able to fit a given strategy to particular learning tasks.
2. Students' ability to steer and direct their learning processes is dependent on metacognitive skills such as orienting, planning, executing, monitoring, evaluating, and correcting. This tradition holds that successful learners are able to swiftly transfer knowledge and strategies acquired in one situation to new situations, modifying and extending strategies on the way. Being able to regulate one's learning in one situation (e.g., foreign language learning) does not mean that one can do it in a different context (e.g.,

mathematics). The dependency on external support may also vary between situations. In CCC there is one construct in this domain, 'Control strategies'. Layers 1 and 2 describe the cognitive and metacognitive aspects of SRL as they relate to the 'Learning strategies' aspect of CCC. Motivation and issues related to the self-system are also important. It is not enough to be aware of different learning strategies; one must also be motivated to use them. This adds a third layer of complexity to the layers just described.

3. How students construe themselves as learners, in particular the goals they set for themselves, is important in order to understand why students are inclined to do what they do, which is sometimes not what is expected of them. Students may feel that using control strategies takes too much time and effort. Perhaps one or more psychological needs are thwarted in a learning context, so they will not identify with the goals of that context. This layer also includes their ability to define ongoing and upcoming activity in the light of their own wishes, needs, and expectations, and to protect their own goals from conflicting alternatives. Some students are able to pursue multiple goals simultaneously, whereas other students seem to pursue goals in a serial manner. CCC has several self-related constructs. Self-efficacy beliefs are central to SRL, and this concept is discussed further below.

8.2.2 Self-efficacy

The construct 'Self-efficacy' is an important one in CCC. Bandura (1986) defines 'Self-efficacy' as

> *People's judgements of their capabilities to organize and execute courses of action required to attain designated types of performances (Bandura 1986, p 391).*

Students develop self-efficacy beliefs in several ways. Experiences of success and failure are of course important to their self-efficacy beliefs, but it depends on how the experiences are *explained* through self-referential processes. A student with a high self-efficacy is more likely to attribute a failure to chance, whereas a student with a low self-efficacy belief may see the failure as confirmation that he or she cannot achieve. When a strong self-efficacy belief is developed, a few defeats may not matter. Furthermore, a student with high self-efficacy will intensify his or her efforts when problems are encountered, whereas a student with low efficacy may easily give up. Therefore this construct should correlate with 'Effort and perseverance', which is related to achievement standards. A high self-efficacy belief is important to setting high, but manageable standards for achievement. If the standards are too high, failure will almost certainly be the result. On the other hand, if the standards are so low that the goals are easily achieved then that is also a problem. By reducing the standards, success may be obtained with little effort. On the other hand, if one rewards oneself for mediocre results, one's self-esteem may suffer as a result.

Individuals also develop self-efficacy by comparing themselves to others, especially people who are important to them and have a similar background. In this realm of developing identity as a learner, two other CCC constructs in the 'Learning style' aspect focusing on whether the student prefers to compete and/or cooperate in learning situations are important. Bandura (1997) holds that learning environments that furnish cooperation are much healthier for developing self-efficacy beliefs in weak students than are learning environments based on competition, for instance by focusing on grades, or when the teacher compares students' achievements in front of the class.

The discussions that follow will draw on the theoretical framework just outlined, with a focus on learning strategies, self-efficacy and learning style. The various constructs in CCC will thus be covered in various degrees of details in the discussions, and only data for constructs that are part of the discussion are presented below. For each construct results will be given in the form of mean values as well as correlations with reading literacy. Results will be given for all Nordic countries together with the OECD mean and the highest ("max") and lowest ("min") values internationally.

8.3 Results

8.3.1 Learning strategies

Table 8.2 presents results for the construct 'Memorisation'. The construct has been derived from the frequency with which students used the following strategies when studying:

> *I try to memorise everything that might be covered; I memorise as much as possible; I memorise all new material, so that I can recite it; I practice by saying the material to myself over and over.*

The results in table 8.2 show that there are only marginal differences between the Nordic countries when it comes to the correlation between 'Memorisation' and reading literacy. The differences between the mean values of the construct are more striking. The lowest mean value is found in Norway, a mean value that is among the lowest of all countries participating in PISA. The difference between Norway and Sweden is surprisingly high and amounts to as much as 77% of the standard deviation.

Table 8.2 Results for the construct 'Memorisation'

Country	Mean	Correlation with reading literacy
Denmark	0.05	0.05
Finland	-0.10	0.07
Iceland	-0.27	-0.01
Norway	-0.60	-0.03
Sweden	0.17	0.09
Int. min.	-0.68	-0.16
Int. max.	0.89	0.14
Mean OECD	0.00	0.01

Table 8.3 shows results for the construct 'Elaboration'. This construct has been derived from the frequency with which students used the following strategies when studying: *I try to relate new material to things I have learned in other subjects; I figure out how the information might be used in the real world; I try to understand the material better by relating it to things I already know;* and, *I figure out how the material fits in with what I have already learned.* The results in table 8.3 show that the correlation between the construct and score in reading literacy is near to the OECD mean in all the Nordic countries. When it comes to the mean construct values, all the Nordic countries except Sweden have mean values below the OECD mean. Iceland has the lowest mean of all the countries participating in PISA.

Table 8.3 Results for the construct 'Elaboration'

Country	Mean	Correlation with reading literacy
Denmark	-0.12	0.13
Finland	-0.15	0.16
Iceland	-0.24	0.15
Norway	-0.22	0.16
Sweden	0.01	0.14
Int. min.	-0.24	0.01
Int. max.	0.47	0.35
Mean OECD	0.00	0.13

Table 8.4 shows the results for the construct 'Control strategies'. This construct has been derived from the frequency with which students used the following strategies when studying: *I start by figuring out what exactly I need to learn; I force myself to check to see if I remember what I have learned; I try to figure out which concepts I still haven't really understood; I make sure that I remember the most important things*; and, *when I study and I don't understand something, I look for additional information to clarify this.* The results in table 8.4 show that for this construct also the relationships with reading literacy are

close to the OECD mean in the Nordic countries. The differences between the countries are small. Again, larger differences are found regarding the mean values for the construct. Norway has the lowest mean value of all the countries in PISA, while Sweden is the only Nordic country with a mean value above the OECD average.

Table 8.4 Results for the construct 'Control strategies'

Country	Mean	Correlation with reading literacy
Denmark	-0.23	0.13
Finland	-0.47	0.16
Iceland	-0.36	0.15
Norway	-0.58	0.12
Sweden	0.03	0.20
Int. min.	-0.58	0.09
Int. max.	0.40	0.34
Mean OECD	0.00	0.21

8.3.2 Motivation

Table 8.5 shows the results for the construct "Effort and perseverance". The construct is based on these single items: *When studying, I put forth my best effort; when studying, I try to do my best to acquire the knowledge and skills taught; when studying, I keep working even if the material is difficult*; and, *when studying, I work as hard as possible.* The results in table 8.5 show that the correlation between this construct and score in reading literacy is similar in all the Nordic countries. Also the variation between the mean construct values is relatively small among the Nordic countries.

Table 8.5 Results for the construct "Effort and perseverance."

Country	Mean	Correlation with reading literacy
Denmark	-0.05	0.20
Finland	-0.03	0.26
Iceland	-0.09	0.22
Norway	-0.16	0.25
Sweden	0.02	0.17
Int. min.	-0.39	0.03
Int. max.	0.40	0.26
Mean OECD	0.00	0.15

8.3.3 Self-concept

Table 8.6 shows the results for the construct 'Self-efficacy'. This is a construct in the category *Self-concept.* The construct is derived from these single items:

I'm certain I can understand the most difficult material presented in texts; I'm confident I can do an excellent job on assignments and tests; I'm confident I can understand the most complex material presented by the teacher. The results in table 8.6 show small variations between the Nordic countries regarding the correlation between the construct and reading literacy, while the magnitude of the correlations is relatively high. Again, Sweden has the highest mean value.

Table 8.6 Results for the construct 'Self-efficacy'

Country	Mean	Correlation with reading literacy
Denmark	-0.03	0.35
Finland	-0.16	0.26
Iceland	0.04	0.38
Norway	-0.04	0.33
Sweden	0.19	0.34
Int. min.	-0.42	0.12
Int. max.	0.35	0.38
Mean OECD	0.00	0.24

Table 8.7 shows the results for the construct 'Academic self-concept'. This is the final construct in the category *Self-concept*. The construct is based on the following single items: *I learn things quickly in most school subjects; I do well in tests in most school subjects;* and, *I'm good at most school subjects.* The results in table 8.7 shows that stronger relationships between the construct and reading literacy than the OECD mean are found in all the Nordic countries. The relationships in Norway and Iceland are the strongest among all the countries. Denmark has a particularly high mean value for the construct.

Table 8.7 Results for the construct 'Academic self-concept'

Country	Mean	Correlation with reading literacy
Denmark	0.41	0.42
Finland	-0.05	0.44
Iceland	-0.04	0.47
Norway	0.05	0.47
Sweden	0.09	0.39
Int. min.	-0.96	0.11
Int. max.	0.43	0.47
Mean OECD	0.00	0.30

8.3.4 Learning style

Table 8.8 shows the results for the construct 'Co-operative learning'. This construct has been derived from students' level of agreement with the

following statements: *I like to work with other students; I learn the most when I work with other students; I do my best work when I work with other students; I like to help other people do well in a group;* and, *it is helpful to put together everyone's ideas when working on a project.* The results in table 8.8 show that there are significant differences between the Nordic countries regarding the correlation between this construct and score in reading literacy. The largest correlation is found in Norway, which in fact has the largest correlation of the countries in PISA. The lowest correlation is found in Sweden, and is close to the lowest correlation among all countries. The mean values also vary a lot between the Nordic countries. The difference between Denmark (highest mean) and Iceland (lowest mean) is as large as 79% of a standard deviation.

Table 8.8 Results for the construct 'Co-operative learning'

Country	Mean	Correlation with reading literacy
Denmark	0.50	0.05
Finland	0.04	0.12
Iceland	-0.29	0.11
Norway	0.17	0.19
Sweden	-0.21	0.02
Int. min.	-0.85	0.01
Int. max.	0.59	0.19
Mean OECD	0.00	0.08

In table 8.9 we have shown the results for the construct 'Competitive learning'. This construct was derived from the students' level of agreement with the following statements: *I like to try to be better than other students; trying to be better than others make me work well; I would like to be the best at something;* and, *I learn things faster if I'm trying to do better than the others.* The results in table 8.9 show that Norway again has the strongest correlation with reading literacy of all the countries in PISA. The differences between the Nordic countries are, however, small. The mean construct values vary considerably, with Finland having the lowest value and Sweden the highest.

Table 8.9 Results for the construct 'Competitive learning'

Country	Mean	Correlation with reading literacy
Denmark	0.19	0.17
Finland	-0.25	0.19
Iceland	0.01	0.23
Norway	-0.03	0.27
Sweden	-0.01	0.13
Int. min.	-0.38	-0.06
Int. max.	0.54	0.27
Mean OECD	0.00	0.14

8.4 Discussion

8.4.1 Learning strategies

Learning strategies are the most specific and confined constructs in the CCC, but also the most dependent on context; the task, the learning situation, who is taking part and so on. It is this context dependency that makes it difficult to interpret 'Memorisation' and 'Elaboration'. It is not so much the *frequency* of use of these strategies that identifies a student who can self-regulate his or her learning, but the fact that the student can flexibly adapt strategies according to the situation. Therefore, the poor correlation between learning strategies and reading score in the Nordic countries may indicate that CCC "smears out" this contextual dependency, not that learning strategies are of minor importance for achievement.

We note that among the Nordic countries Norway has the lowest mean for the constructs 'Memorisation' and 'Elaboration' (together with Iceland for the latter), and Sweden the highest. This may of course be because of the cultural effects discussed earlier, but it may also be because students in the two countries have different levels of 'metacognitive awareness'; students who know that they use 'Memorisation' as a strategy also know that they use 'Elaboration'. Remember, the CCC format requires the students to not only understand the question, but also to make a synthesis over a range of learning situations and time spans. The 'Control strategies' construct, which is the most metacognitively oriented construct, fits into this picture for Norway and Sweden, with highest mean for Sweden and lowest for Norway. Given this interpretation of the results, it seems that the other Nordic countries should take a close look at the Swedish results.

Although it is accepted that it would be simplistic to label the use of 'Memorisation' as 'bad', and 'Elaboration' as 'good', 'Elaboration' is certainly necessary for a deeper understanding of the subject matter. The somewhat higher correlation with reading literacy score indicates this. Apart from Sweden, all the Nordic countries score below the OECD mean on this construct, with Iceland actually defining the international minimum, closely followed by Norway. As a sensible use of learning strategies is essential for SRL, these results alone indicate that this aspect needs more attention in the future if SRL is taken as an important outcome of education.

8.4.2 Self-efficacy as part of self-concept

According to the theory, the construct 'Self-efficacy' is expected to be important for self-regulation, and this is confirmed by the CCC data. In all Nordic countries, 'Self-efficacy' correlates at the 0.5–0.6 level with 'Elaboration', 'Control strategies' and 'Effort and perseverance'. These constructs are all essential to SRL. 'Self-efficacy' has a fairly large correlation

with reading score. The correlations are among the highest internationally, with Iceland highest of all with 0.4.

Sweden has markedly the highest level of 'Self-efficacy'. On the other hand, Denmark has markedly the highest level on 'Academic self-concept', with a value close to the international maximum. The other Nordic countries cluster around the international mean. It is puzzling why Sweden and Denmark differ so much in this respect, as one would expect these constructs to be highly related, and 'Self-efficacy' and 'Academic self-concept' correlate 0.7 and 0.6 in the two countries. The correlations with the reading score tend to be stronger than the international mean for the Nordic countries, i.e. the constructs matter more to the reading score, most of all in Norway and Iceland.

'Self-efficacy' correlates differently to learning style, that is, the extent to which students prefer to compete or to co-operate when learning. (Note that this is not a dichotomy, and indeed the two variables are only weakly correlated for the Nordic countries). Self-efficacy correlates much more strongly with 'Competitive Learning' than with 'Co-operative learning'. In other words, students with high self-efficacy tend to prefer to compete, whereas students with low self-efficacy tend to prefer not to. Needless to say, it does not follow that students in the latter group would increase their self-efficacy if they become more competitive. Bandura (1997) argues that the causal relationship is rather the other way around. In learning environments with a strong focus on competition, for instance with a strong focus on grades, or a sequential whole-class teaching style where the teachers make frequent comparisons between students, it is the high-achieving students that benefit at the expense of the low achievers. Co-operation is more likely to strengthen the self-efficacy beliefs of both high and low achievers, and all students suffer when co-operation does not function well.

Denmark is close to the international maximum with a mean of 0.5 for 'Co-operative learning', and also has the highest score for 'Competitive learning'. As noted earlier, these constructs are weakly correlated, so that it is difficult to predict the score for one from the score for the other. Iceland has the lowest mean for 'Co-operative learning', whereas Finland has markedly the lowest score among the Nordic countries for 'Competitive learning'.

As might be expected from the discussion above, 'Competitive learning' correlates markedly more highly with reading score than does 'Co-operative learning'. For both constructs, Norway has the strongest correlation, while Denmark and Sweden have the weakest. To take up Bandura's point above, one should be cautious about placing too much emphasis on correlations with achievement, in this case the reading score. This point can be made for all constructs in CCC. Self-regulated learning is indeed important for students' achievement, but it does not follow that achievement should be the primary focus in the process of developing SRL among students. SRL includes complex processes of cognitive and metacognitive strategies, motivation and self-concepts, where achievements should be considered as one of the important variables rather than the ultimate goal. As noted by Bandura,

> *Educational practices should be gauged not only by the skills and knowledge they impart for present use but also by what they do to children's belief about their capabilities, which affect how they approach the future (Bandura 1997, p 176).*

With the self-constructs, the correlations are consistently higher than was the case with the cognitive and metacognitive learning strategies constructs. We see that among all constructs the correlations are highest for 'Academic self-concept'. This concept is the most general notion of self in CCC, and can be thought of as a synthesis of more specific self-concepts. To extend the discussion on learning strategies further, the results can partly be explained by noting that these constructs are less context-dependent than the learning strategies. The differences in correlations should therefore not be seen as a rank order of how important they are to SRL. As previously noted, it is the dynamics between the constructs in specific settings that makes the difference between the 'novice' and 'expert' learner.

In this chapter, some intriguing differences in cross-curricular competencies have been found between the Nordic countries. More sophisticated analytical techniques than we have used may give more clues about what they represent, but the extent to which the effect of cultural diversity on how students respond to CCC causes the differences is not known. A qualitatively oriented research project that focuses on students' reflections on their answers to the questionnaire could provide information on this point, and would be helpful in interpreting the answers as well. However, the existence of these variations alone should motivate comparative studies in a Nordic context on differences in the conceptualisation of cross-curricular competencies, their implementation and how they are evaluated.

References

Bandura, A. (1986). *Social Foundations of Thought and Action. A Social Cognitive Theory*. New Jersey: Prentice Hall.

Bandura, A. (1997). *Self-efficacy. The exercise of control*. New York: W. H. Freeman and Company.

Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research*, 31, p 445–457.

Knain, E. (2002). *Elevenes læringsvaner. Selvregulert læring som en viktig kompetanse på tvers av fag: Perspektiver og resultater*. Acta Didactica, nr 5, Department of Teacher Education and School Development, University of Oslo.

Pintrich, P. R. (2000). The Role of Goal Orientation In Self-Regulated Learning. In Boekaerts, M., Pintrich, P. R. & Zeidner, M. (eds.) *Handbook of Self-Regulation*. San Diego & San Francisco: Academic Press.

9 STUDENT AND TEACHER BEHAVIOUR

Rolf V. Olsen

9.1 Introduction

The school is an important unit of analysis in PISA. This is evident from both the sampling design, where the school is the primary sampling unit, and the instruments used, with a separate school questionnaire, and also from the published international report, where differences between schools are highlighted (OECD 2001). However, information from other international surveys shows that the differences between schools are small in the Nordic countries. Central to our school systems has been equality of opportunity to learn for all, independent of geographical, economic, ethnical, social or cultural background. As described in chapter 1, the Nordic population tested in PISA have not, for instance, been streamed into different schools or study tracks .

Overall, the same trend is apparent in the PISA data. The variance components in reading score *between* the schools relative to the total variance for the Nordic countries range from 8% (Iceland) to 19% (Denmark). These ratios are very low compared to the rest of the countries participating in PISA (OECD 2001; see also chapter 10).

In this chapter some constructs derived from the student questionnaire and some individual items from the school questionnaires are presented, mainly through descriptive statistics. The cluster of questions used for the analyses presented in this chapter is mainly related to how teachers' and students' behaviour in the classroom effects learning as perceived by the students and the principals respectively.

9.2 Measurement of the constructs

The constructs presented in this chapter are aggregates from the student questionnaire scaled such that 0 is the OECD mean and 1 is the overall standard deviation. The constructs presented are measures of the students' own individual perceptions of what happens in their classroom.

- **Teacher support:** This construct was derived from students' reports on the frequency with which, in their class of the language of assessment: *the teacher shows an interest in every student's learning: the teacher gives students an opportunity to express opinions; the teacher helps students with their work; the teacher continues teaching until the students understand; the teacher does a lot to help students*; and, *the*

teacher helps students with their learning (using a four-point scale with the response categories 'never', 'some lessons', 'most lessons' and 'every lesson').

- **Disciplinary climate:** This construct summarises students' reports on the frequency with which, in their class of the language of assessment: *the teacher has to wait for a long time for students to quieten down; students cannot work well; students don't listen to what the teacher says; students don't start working for a long time after the lessons begin; there is noise and disorder;* and, *at the start of class, more than five minutes are spent doing nothing* (using the same scale as above, but inverted so that low values indicate a poor disciplinary climate).
- **Teacher-student relations:** This construct was derived from students' reports on their level of agreement with the following statements: *students get along well with most teachers; most teachers are interested in students' well-being; most of my teachers really listen to what I have to say; if I need extra help, I will receive it from my teachers*; and, *most of my teachers treat me fairly* (using a four-point scale with the response categories 'strongly disagree', 'disagree', 'agree' and 'strongly agree').
- **Pressure to achieve:** This constructs summarises students' reports on the frequency with which, in their class of the language of assessment: *the teacher wants students to work hard; the teacher tells students that they can do better; the teacher does not like it when students deliver careless work;* and, *students have to learn a lot* (using the same scale as for disciplinary climate).

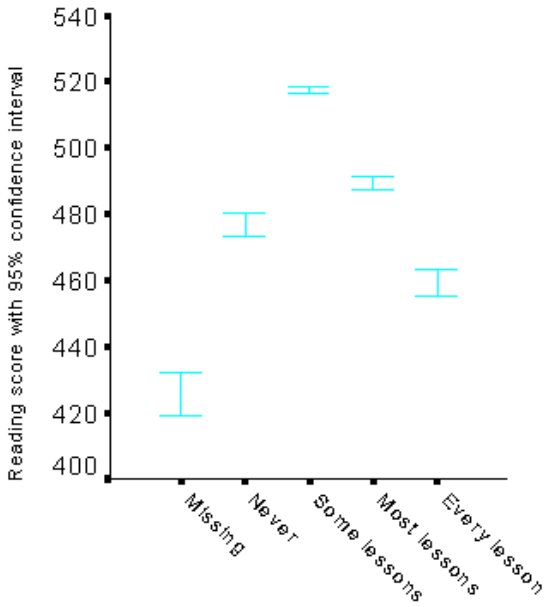
For obvious reasons performing analyses with these constructs is problematic. Firstly, the constructs based on the student questionnaire are students' perceptions and therefore not necessarily a good measure of some objective reality. Very often one can see that measures collected by self-report, intended to map what happens in the classrooms or schools, are biased such that high-achieving students seem to perceive their surroundings differently from low-achieving students. This tends to produce contra-intuitive effects or to mask effects. One example is given in figure 9.1. This figure shows the average reading literacy score (with confidence intervals) for students giving different responses to one of the questions included in the construct *Disciplinary climate*. This is a good example of how the relationship between school-related variables and achievement very often is curvilinear. One important implication is that the ordinary Pearson correlation coefficient between these two variables is close to 0.

Secondly, the students are answering questions relating to phenomena in their classroom. In PISA the school, and not the class, was the primary sampling unit. It is not obvious that, for instance, the disciplinary climate in

one classroom is typical of the school. This means that aggregating these data to school level[1] is not necessarily meaningful.

A third problematic aspect is related to the questionnaire answered by the principals. These items also do not measure the objective phenomena referred to in the question, but rather the principal's perception of these phenomena in his/her school. It is very likely that the degree to which the principals are familiar with the phenomena addressed in the items varies both within and between countries. We will return to all these problems later in this chapter.

Figure 9.1 Average reading scores for the different responses to the student questionnaire item "Students don't listen to what the teacher says" in Norway



9.3 Descriptive results

In the following section the mean values for the Nordic countries for the above-mentioned variables will be presented by line-diagrams (profiles) for each country, supplemented with international means (0), maxima and minima. In these diagrams the lines themselves have no meaning. The direction and steepness of the lines have no substantial meaning by themselves. Their function is to provide a visual framework so that the main characteristics can be identified at a glance.

[1] "Aggregating to school level" in this case implies that school averages are computed and analysed.

9.3.1 Students' views

Figure 9.2 The mean scores for the Nordic countries (coloured lines) and international ("Int.") extremes (black lines) on the four school-related constructs

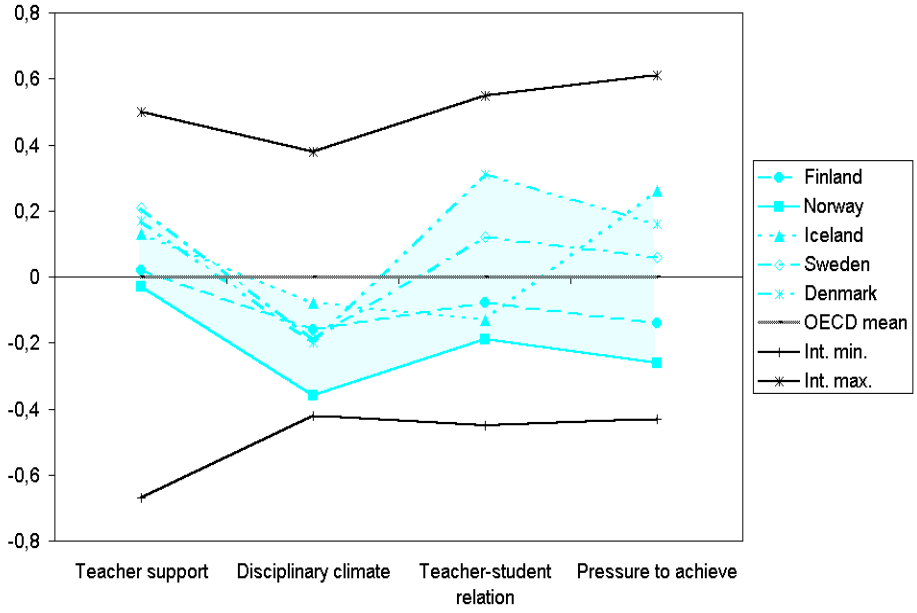
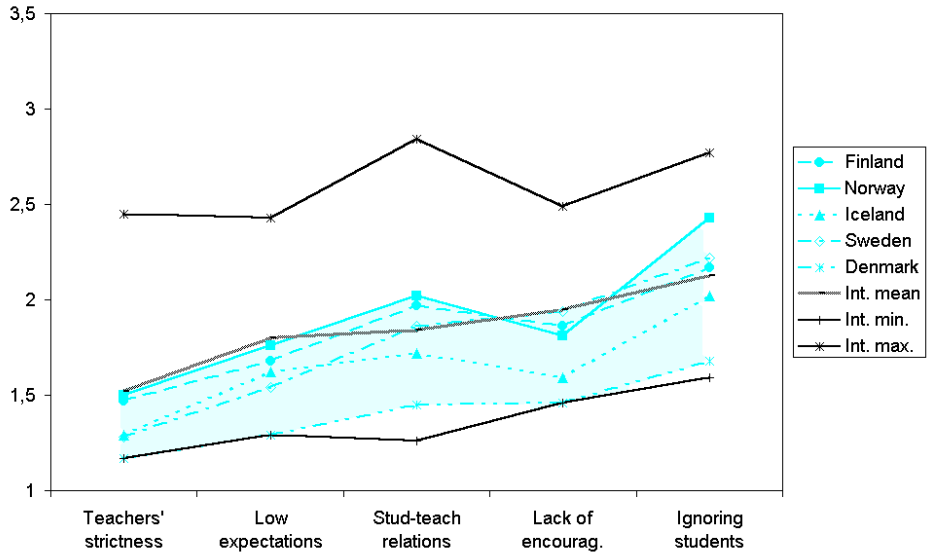


Figure 9.2 gives a summary of the four constructs from the student questionnaire. The figure shows that in the Nordic countries the students report to a larger degree than in other OECD countries that their teachers are supportive. They also perceive their classrooms as less disciplined learning environments than their OECD counterparts. However, there are surprisingly large differences between the Nordic countries for the constructs 'Teacher-student relation' and 'Pressure to achieve'. The Norwegian profile is the lowest of the Nordic profiles. The typical student in Norway reports that there is a relatively low pressure to achieve combined with a worse than average teacher-student relationship, and that the classroom is not particularly quiet. The Danish profile is very much the opposite of the Norwegian. In Denmark, students report a relatively high pressure to achieve and good teacher-student relationships, and the teacher is perceived to be supportive. The most dominant feature in the Swedish profile is the fact that the students perceive their teachers to be quite supportive. In Iceland the students report that there is a relatively high pressure to achieve. For Finland the most pronounced feature is that the profile is quite close to the OECD mean for all the constructs.

9.3.2 Principles' views

Figures 9.3 and 9.4 give similar profiles for some of the individual questions in the school questionnaire answered by the principal or another school-leader. Figure 9.3 presents the student-related factors and figure 9.4 presents the teacher-related factors. For all variables, high values mean that the principals identify this as a problem for learning. The questions reported in these figures all began with "In your school, is the learning of 15-year old students hindered by:…". A four-point scale was used ('not at all'=1, 'very little'=2, 'to some extend'=3, 'a lot'=4). The scale used in figures 9.3 and 9.4 has not been standardised. The absolute mean of the scale, the point halfway between 1 and 4, is 2,5.

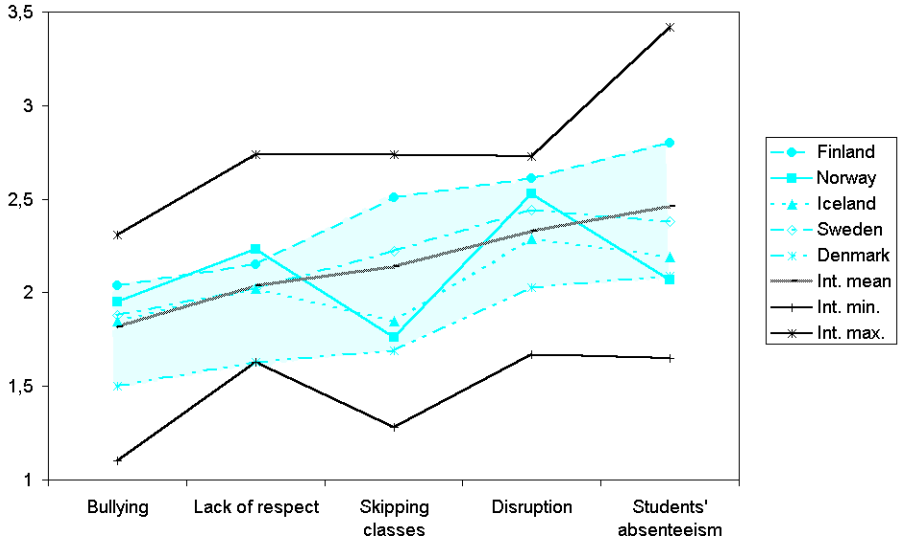
Figure 9.3 Nordic and international mean values for teacher-related factors in the school-questionnaire. Factors are sorted by increasing mean value in the OECD countries . High values indicate problems for learning.



One main characteristic of the factors associated with teacher behaviour (figure 9.3) is that the OECD mean is well below the absolute mean for all the questions. In other words, within the principals' own understanding of the scale, they do in general not perceive any major problems related to teacher factors. For the Nordic countries this feature is even stronger. Mostly the means are close to or below the international mean. As was the case with the data from the students' report (figure 9.2), the Danish and Norwegian profiles are the lowest and highest respectively of the Nordic profiles; however, this time in the opposite direction since the scales are reversed. The Danish profile is exceptional in the sense that from an international perspective their

headmasters think very highly of their teachers. Iceland has the same tendency, although not as pronounced.

Figure 9.4 Nordic means for student-related factors in the school questionnaire, sorted by increasing OECD mean. High values indicate problems for learning



The picture painted by the school-leaders is somewhat less positive for the factors associated with student behaviour (figure 9.4) compared to their assessment of teacher factors (figure 9.3). Even though the international mean is still below the absolute mean for the scale used, the overall profile is shifted upwards towards higher values, meaning that school-leaders identify these factors as causing more problems for the learning environment in their schools. This shift is even more pronounced for the overall Nordic profile. The Danish profile is still well below the other Nordic countries' profiles, and in Danish schools no particular learning problem is associated with student behaviour. In Finland the principals identify several of the student-related factors as being problematic. It is worth noticing that, compared to their Nordic counterparts, the Finnish principals report that students skipping classes or being absent from school is a substantial hindrance to learning in their schools. The principals in Sweden, Iceland and Norway all identify disruption as the single most important factor hindering learning in their schools.

As always with data given by self-report, one should keep in mind that the differences between countries may partly reflect the fact that, even if the phenomena addressed by the items were of equal kind and size in all countries, principals and students in different countries may have culturally conditioned

differences in their *perceptions of the phenomena*. Also, the variation in the ratings for the variables between countries may indicate that there are cultural features in the different countries that result in different *perceptions of the scale used*. Of course, this is impossible to control or correct for. However, it is reasonable to assume that comparisons between the Nordic countries are comparisons between more or less identical cultures.

9.4 Comparing students' and principals' reports

It is difficult to make a direct comparison between the students' reports and the principals' reports based only on the descriptive statistics presented in figures 9.2-9.4. Indeed, such a comparison is at the outset impossible because the students and principals were not asked identical questions, and they did not report their answers on the same scales. However, there are several indications that they seem to partly agree about which are the most important characteristics of teacher and student behaviour.

- The principals report that their teachers are not too strict and that lack of encouragement is not a problem. At the same time the Nordic students evaluate their teachers as being supportive.
- Overall, both the students and the headmasters identify the disciplinary climate in the classrooms as a main problematic characteristic compared to the other countries participating in PISA.
- Also, in both reports the same countries appear at each extreme of the Nordic range; Danish students and principals report no major problems associated with student or teacher behaviour, while in Norway to some extent the scores for these data are higher.

The first and third of these points are also supported when correlating the variables from the school questionnaire with the "corresponding" variable in the student questionnaire. These correlations were made with the student data aggregated to school level. The fact that the average student and principal to some degree describe their school in the same manner, combined with the fact that the variables relating to the same phenomena are correlated in the expected direction, is a strong indication that both students and principals at least partly relate to the same objective reality when answering these questions. By triangulating the data in this way the validity of our interpretations is strengthened. This also makes analysis of possible school effects interesting.

9.5 Possible school effects

When analysing these data, attention must be drawn to their hierarchical nature. Students are nested within classes, which in turn are nested within schools. In this section the relationship between achievement and the variables presented above is studied both at the individual student level and at the school level. For the school level a simple procedure whereby data at student level are aggregated to produce whole-school data has been used. This means that each

school is represented by their students' mean. In order to reduce outlier problems, schools with less than 10 students have been excluded.

Therefore, the results are not estimates of the population defined by PISA as being all schools with 15-year-old students. In the Nordic countries quite a few students live in rural areas with small schools. This is particularly the case for Iceland. Almost 50% of the Icelandic schools are not included in the school level analysis below. As a consequence of this the calculated correlation coefficients are non-significant for Iceland, and more seriously, the data are biased. Also, almost 30% of the Danish schools are excluded. Some of these schools are probably, as is the case in Iceland, schools in rural areas, but it is also important to note that in the Danish school system they have so-called continuation schools, many of which are small and therefore excluded in the analysis. Without going into detail, there are many characteristics of these schools and their students, which will almost certainly result in biased estimates for Denmark.

We must keep in mind that most of the overall variance in achievement is due to differences between students. It is likely that much of this variance cannot be explained by factors that can be manipulated through school policy actions. Policy decisions are more directly related to school factors than individual factors. Of course, these are in turn to some degree indirectly related to student factors. Therefore, even if the between-school variance is small it is important to analyse it in order to identify possible improvements that could be made by policy-makers. It is important to describe any noticeable effects of these variables at the school level simply because they largely address issues of concern to policy-makers.

In this section attention will only be given to the constructs from the student questionnaire describing the students' perception of what happens in their classroom. It should be remembered that for all constructs high values correspond to a positive situation. When correlating the student variables with reading achievement for the data at student level, the effects are in the same direction[2] for nearly all countries for all four variables[3], but they are quite small (see table 9.1). This first fact is reassuring in the sense that if the data had been faulty, these relationships would most probably have gone in different directions on pure chance. When producing the equivalent coefficients for the data aggregated to school level, the same direction occurs in the relationship, but now the coefficients are much higher, as can be seen from table 9.1. The most plausible interpretation of this fact is that these variables are related to genuine school phenomena, and they are not just measures of individual perceptions. In other words, even if these measures are directly related to students' perceptions of what happens in their classroom, the fact that these

[2] Direction is here not any statement of causal direction. It is only related to whether the correlations are positive or negative.

[3] The only exception is the correlation for 'Pressure to achieve' in Denmark at student-level. For all countries except Denmark this coefficient is negative

variables when aggregated to school level correlate significantly with achievement suggests that there are social mechanisms within schools that mediate student-teacher behaviour across classes.

Table 9.1 Correlation coefficients between classroom factors and reading literacy, given for data at student level (Ind) as well as at school level (Agg). n.s. = not statistically significant at the 0,05 level

	Denmark		Finland		Iceland		Norway		Sweden	
Construct	**Ind**	**Agg**	**Ind**	**Agg**	**Ind**	**Agg**	**Ind**	**Agg**	**Ind**	**Agg**
Teacher support	0,09	n.s.	0,07	0,15	0,09	n.s.	0,11	0,10	0,07	0,15
Disciplinary climate	0,08	0,30	0,1	0,25	0,07	n.s.	0,08	0,22	0,12	0,35
Teacher-student relation	0,16	0,26	0,15	0,21	0,18	n.s.	0,17	0,37	0,12	0,15
Pressure to achieve	0,03	n.s.	-0,18	-0,20	-0,15	-0,33	-0,12	-0,25	-0,14	-0,24

To sum up the results in table 9.1:

A notable result is that across all countries the aggregate for the students' perception of the disciplinary climate is positively correlated with achievement on the reading test, ranging from about 0,2 (Finland and Norway) to 0,35 (Sweden). These are significant results, not merely in the statistical sense. In fact, for Sweden this construct alone accounts for more than 12% of the between-school variance.

The teacher-student relationship also correlates quite strongly with reading score at both school and at student levels. At the student level this could very well be due to the fact that high achieving students get along well with their teachers. However, one could argue that the aggregated results are more likely to indicate that schools where teachers and students get along well, to a larger extent succeed in fostering the abilities measured by the reading test in PISA. This variable seems to be particularly important for Norway, explaining almost 14% of the between-school variance.

The construct "Teacher support" has a positive but weaker relationship with reading score than have the above two constructs.

The construct "Pressure to achieve" also has pronounced effects at the school level for all countries, except Denmark. In schools were students feel that the pressure to achieve is weak, the scores on the PISA reading test are higher than in schools were students perceive this pressure to be strong. Some might say that this is contrary to what might be expected. However, by looking through the statements to which students should agree or disagree it is obvious that this construct to some degree give a description of the tendency to which teachers give negative feedback to the students.

9.6 Concluding remarks

By combining the descriptive measures and the correlations with reading score, an interesting picture of school factors in the Nordic countries appears. In general the relationships with achievement seem to be similar across countries. Schools with high average achievement on the PISA reading literacy test are characterised by having supportive teachers with good relations to their students, working in classrooms with a good disciplinary climate where students do not feel that the pressure to achieve is too high. However, the school principals' and students' reports of these phenomena show interesting differences between the Nordic countries, which have been addressed in this chapter.

The descriptions given in the student questionnaire data have been triangulated with data from the school questionnaire, leading to the conclusion that the measures are working as intended. Also, the fact that the documented effects are larger for schools than for individuals alleviates our initial concern that aggregating these data to school level is not necessarily meaningful. This increases our confidence that analyses aiming at modelling between-school variance are not only warranted, but also technically possible, at least for Norway, Sweden and Finland. In this chapter the data have been studied across the hierarchical levels using very simple, some might even say simplistic, methods. The results presented in this chapter show that modelling the structural relationship of these variables, at both levels, but also across the levels (interactions) with more appropriate and complex methods (e.g. HLM[4]) is worthwhile.

In this chapter the most obviously school-related variables in the PISA database have been used in the analyses. Interesting school effects may also be found for other variables. In particular, the variables related to learning strategies (see chapter 8) are potential candidates for such analyses. It is reasonable to assume that students' use of learning strategies is structurally related to some of the variables analysed in this chapter.

References

OECD (2001). *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD Publications.

[4] Hierarchical Linear Modelling is a regression analytical tool that simultaneously analyses variables at both individual and aggregated levels and the interactions between variables at different hierarchical levels (see chapter 10)

10 THE TWO-LEVEL EFFECT OF SOCIO-ECONOMIC BACKGROUND

Jouni Välijärvi and Antero Malin

10.1 Introduction

In this chapter we will explore the data from the PISA literacy study to investigate how between-school differences appear in different Nordic countries., We will also look at the variation in the schools' socio-economic status (MEANISEI), which is the mean of the students' socio-economic background (ISEI; see OECD 2001, p.221), and especially the relationship between this status and the school's performance level. The PISA International Socio-Economic Index of Occupational Status (ISEI) captures the attributes of occupations that convert parents' education into income (Ganzeboom et al. 1992). The values of the index range from 0 (low) to 90 (high). In this connection, a school's performance level is represented by the mean of the students' scores on the combined reading literacy scale within the school.

Between-school variances are very small in the Nordic countries compared with the other OECD countries. This is the case for all three domains. In Denmark the variation is greater than in Finland, Iceland, Norway and Sweden, but even there differences between schools are remarkably smaller than in the OECD countries on average (OECD 2001, p.61).

When comparing the results between the Nordic countries we should bear in mind that the school and student samples differed slightly between the countries. The effects of these differences on between-school variation are hard to estimate. In Denmark, Iceland and Norway the share of the small student samples was clearly higher than in Finland and Sweden. This is due to differences in the organisation of lower secondary education, e.g. in terms of school size, and also differences in the sample designs. In Iceland, 45 % and in Norway, 22 % of the sampled schools had less than 15 students at each grade level on average. In Denmark the proportion of these small schools is 5 %, in Sweden 2 and in Finland only one %. In Finland 70 % of the sampled schools had more than 75 students at each grade level, in Sweden 62 %, in Norway 49 %, but in Denmark only 8 % and in Iceland 5 %.

There are also differences in the sample design between the countries (see (table 10.1). The number of schools assessed and the distribution of the student samples varied considerably. In Finland and Sweden the maximum number of students sampled in each school was 35, in Norway 30, and in Denmark 25. In

Iceland almost the whole target population was covered. The total number of assessed schools across all the five countries was 840.

Table 10.1 The number of schools assessed in Nordic countries

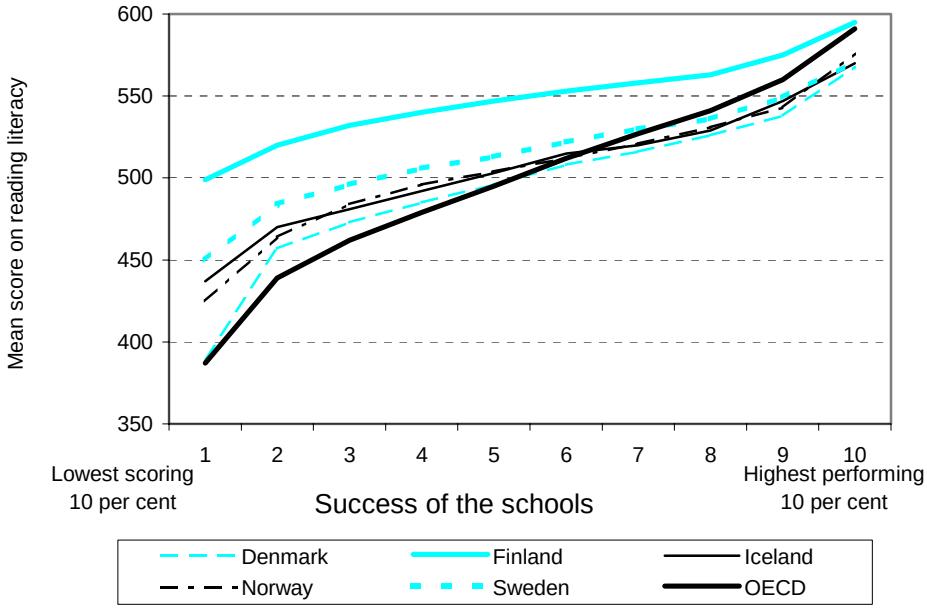
N of students assessed	Denmark	Finland	Iceland	Norway	Sweden
Less than 10	23	2	44	12	4
10 to 19	69	2	22	17	2
20 to 29	133	15	19	141	61
30 to 35		136	9	6	87
Over 35			36		
N of schools in total	225	155	130	176	154

10.2 Between-school variance

Although between-school variances in the Nordic countries were relatively small in comparison with other countries in general, there was still a considerable difference between the average scores of the highest and lowest ranking schools. Figure 10.1 describes differences among schools in five Nordic countries by comparing the average achievement levels of the best and worst performing groups of schools. The figure divides schools into ten percentage groups (10 % of schools in each group) representing different performance levels as indicated by their average on the combined reading literacy scale. Figure 10.1 shows that even the lowest ranking Nordic schools, except for Denmark, reach a clearly higher performance level than the lowest ranking schools in the OECD on average. The trend is clear: the lower the ranks concerned, the greater the difference in favour of the Nordic schools.

When the PISA schools were investigated according to their average performance level on the reading literacy scale, the respective means of the lowest quarter of schools were higher than the OECD on average by 29 points in Norway, 35 in Iceland, 50 in Sweden, and 90 in Finland. In contrast, Denmark's mean score for the lowest quarter of schools was only slightly (10 points) above the OECD average. In Finland, even the lowest performing schools reached almost the OECD average (500 points) of reading literacy, and only 4 % of the schools sampled failed to do this. In Sweden 29 %, in Iceland 41, in Norway 42, and in Denmark 48 % of the schools performed below the OECD average.

Figure 10.1 *Mean scores of the schools on the combined reading literacy scale in the Nordic countries*



Examination of the most successful group of schools evens out the differences between the Nordic and the other OECD countries. When comparing the best performing schools, the OECD average is clearly higher than the corresponding averages in Norway, Sweden, Iceland and Denmark, and approximately equal to that of Finland (figure 10.1). This also means that in many OECD countries the top schools showed a distinctly higher average performance than their counterparts in Finland, in particular, or in the other Nordic countries.

Equality of opportunity to learn is an aim highlighted in education policies across the Nordic countries. The task of the comprehensive school is to provide all children with equal opportunities for learning regardless of their particular school, background or circumstances. In comparisons across the OECD countries, the Nordic comprehensive schools seem to function quite well. However, it may be important to investigate the variation within each educational system because there seem to be some interesting differences between the Nordic countries.

When comparing the top schools with the lowest ranking group within each country, the differences in these groups' average performances on the combined reading literacy scale were also obvious in the Nordic countries (figure 10.1). For instance, the average performance level of the best 10 % of the schools is about one and a half proficiency levels (96 points; see OECD 2001, p.44-48)

higher than in the poorest performing 10 %. In Denmark the difference is two and a half proficiency levels (178 points), and the other Nordic countries have differences between these two values.

Even when the bottom and top 25 % of the schools were examined the variances remained considerable in each country. Again, in Finland the difference between the highest and lowest ranking quarters is the smallest of all Nordic countries (67 points), but still almost one proficiency level. In Denmark the corresponding difference was more than one and a half proficiency levels (115 points). As the proficiency levels are defined on a five-step scale, from the equality point of view the variation between schools cannot be considered insignificant in any country. Thus, ensuring equal educational opportunities for all children still remains a central challenge for education policies in all Nordic countries. In most other countries, though, this challenge is still much greater: in the OECD countries on average the difference between the best and poorest performing 25 % of the schools was 146 points on the combined reading literacy scale, i.e. no less than two proficiency levels.

10.3 Socio-economic status of schools and reading literacy

10.3.1 Constructing the model

This section deals with socio-economic background (the PISA International Socio-Economic Index of Occupational Status, ISEI; see OECD 2001, p.221) as a characteristic of schools. A school's socio-economic background is determined by the homes of its students. Variation in the schools' socio-economic backgrounds derives usually from two factors. First, the school's geographical location often determines the area where the students come from, which means that the particular student population represents the social structure of that particular area. Depending on the degree of regional differentiation, the schools' socio-economic status may vary more or less. Second, if students can choose schools according to their preferences and schools can select their students freely according to their own criteria, this often leads to differentiation between schools according to their social status.

The distribution of schools' on the scale of social status according to students' background varied among the Nordic countries to some extent. The values detected for the index concerned ranged from 27.2 to 79. In Iceland the schools' average social status was the lowest, 48.5, and the variation between schools the greatest (standard deviation 8.0). The social status was of schools' was highest in Norway with an average index value of 53.9, while the variation between schools was the lowest, with standard deviation of 5.9 points. Other Nordic countries were close to each other in this respect. Finland and Sweden had the same average (50.3) and also equal standard deviations (6.6). The

average index value for the social status of Danish schools was 49.6 points, with a standard deviation of 7.3 points.

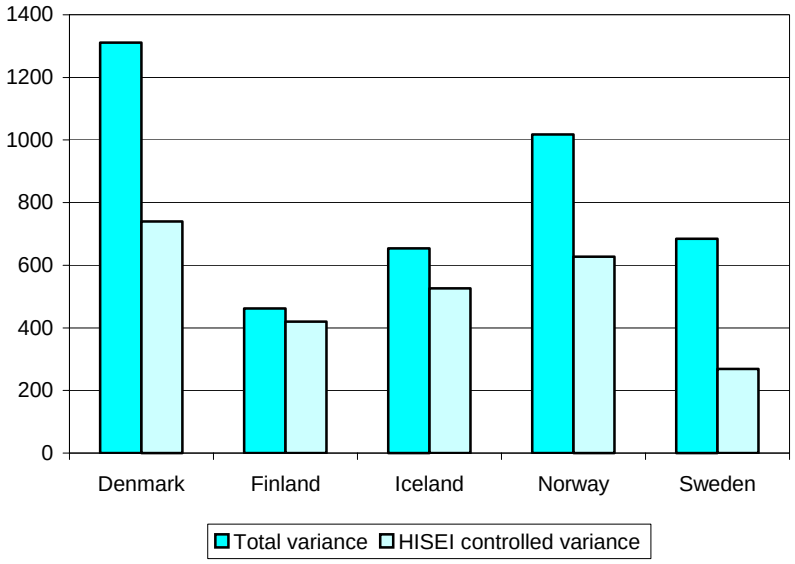
Next we wanted to find out to what extent the variation in schools' literacy performance could be explained by their social status. In the following analyses, the statistical method used is the two-level regression model (Bryk & Raudenbush 1992; Goldstein 1987, 1995), with students as level 1 units and schools as level 2 units. The response variable is the combined reading literacy score. The PISA International Socio-Economic Index of Occupational Status (ISEI) is used to describe the students' socio-economic background. The label HISEI indicates that the highest ISEI values of the two parents (or adult guardian) are used as the home characteristic. The school level socio-economic index is the mean of the students' index values in the school. The regression coefficients of these variables describe the changes in reading literacy score associated with moving one point on the socio-economic index scale. Students' gender was coded as 1 for girls and 0 for boys, in which case the coefficient connected with gender gives an estimate of how much better the girls are in reading proficiency compared with boys.

10.3.2 The results

In all Nordic countries the variation in schools' performance levels is clearly smaller than in the OECD on average (OECD 2001, p.61). On the other hand, there was considerable variation among the Nordic countries as far as the relationship between a school's social status and its students' average reading proficiency level is concerned. The results of our analysis are summarized in the appendix table where "MEANISEI" at level 2 describes the effect of schools' social status on students' achievement. At level 1 students' gender (FEMALE) and the direct effect of their social background (HISEI) were considered. In our analysis the main interest was in level 1 effects, and the level 2 effects can be determined as "controlled effects" in the model.

When the effects produced by the schools' social status (and students' gender) on the variation in literacy performance are standardised in the data, between-school variances diminish most considerably in Sweden and Denmark, whereas in Finland and Iceland the effect of the school's social status on students' average literacy performance is fairly small, as displayed in figure 10.2. In Sweden as much as 61% of the between-school variance in reading proficiency can be explained by differences in the social background of their student populations, which is even more than in the OECD countries on average (55%). In Finland this portion is only 9% and in Iceland 20%. It should be noted, however, that in all Nordic countries the overall variation between schools is small in comparison with the OECD average, and also the differences between the Nordic countries are fairly small. Therefore it is especially interesting that the detectable variation can be explained in very different ways in different Nordic countries.

Figure 10.2 *Total between-school variation and between-school variation when schools' social status (HISEI) is controlled*



In the construct model depicted in the appendix table the effect of students' social background on their proficiency level is divided into two components. On the one hand, there is the effect deriving from the whole school's social status, which affects various features of the learning environment and is reflected in students' performance (Level 2: MEANISEI). This effect can be interpreted as a "bonus" the school brings to each student's performance level. On the other hand, the social status of an individual student's family (Level 1: HISEI) has a direct effect on the student's performance level. The effects of between-school differences on the relative reading proficiency of boys and girls (FEMALE) were also considered in the model.

The appendix table shows that in Norway and Sweden and especially in Denmark the "bonus" derived from schools' social status and reflected in students' average reading proficiency is considerable and also statistically significant in comparison with Finland and Iceland. This result can be interpreted as showing that when the index value for school's socio-economic background (MEANISEI) increases by one point, the students' performance is raised on average by two points in Denmark and by 1.7 points in Norway and Sweden. This school-based effect added only 0.1 points in Iceland and only 0.3 points in Finland, values which were not statistically significant.

In relation to all OECD countries, however, the effect of the school's social status in the Nordic countries remains clearly below average: within the OECD countries as a whole, an increase of one point in school's social status produces on average a rise of 4.2 points in students' performance. As explained above,

this effect can be interpreted as a school-based average bonus deriving from the social status of a school and benefiting all its students regardless of their individual home background. For example, if the difference between two schools in terms of the index of social status equals two standard deviations (standard deviations are given in part 10.3.1), a student selecting the higher status school gets a "direct benefit" in terms of the PISA combined reading literacy scale as follows: in Denmark 29.2 points, in Sweden 22.4, and in Norway 20.1 points. In Finland and Iceland, in contrast, the corresponding effect was negligible compared with the other Nordic countries; such computational "benefit" was 3.9 points in Finland and 1.6 points in Iceland. The school-based effect is added to the other component, namely the effect deriving from the social status of the student's family.

The effect of the socio-economic background of the student's family on the proficiency level remained statistically significant for all Nordic countries, even after controlling for the school-based effect. In this respect there was little variation among the countries (see the appendix table). In Iceland the direct effect of the family background on student's proficiency was the smallest; in other words, on average when the index value for home socio-economic background (HISEI) increased by one point, the student's proficiency improved by one point as well. In Finland the increase in students' proficiency was 1.2 points, in Denmark and Sweden 1.5, and in Norway 1.6 points. In comparison with the OECD average, the direct effect of family background on proficiency is stronger in the Nordic countries. Above all, this is due to the fact that in the other OECD countries the effects of socio-economic factors tend to be manifest, more strongly than in the Nordic countries, as differentiation between schools along with different student selection patterns steered by the family's social status.

10.4 Conclusions

Equality of performance is apparent in Nordic schools when the data from the PISA literacy study are examined in the context of all OECD countries. However, remarkable differences also exist between schools in the Nordic countries if the five proficiency levels are considered. The correlation between schools' social status and their performance on combined reading literacy varies considerably. This result is open to various interpretations. One interpretation could be that in contrast to Finland and Iceland, schools in Sweden, Denmark and Norway are more clearly divided into those of low and high social status, according to the average social status of the parents. This might result from a more long-standing, at least from the Finnish perspective, policy of allowing students and their parents to choose the school they wish. Also many social scientists have argued that liberalised school selection practices will show first as increased variation between schools in terms of the social structure of their student population, and then as increasing differences in the schools' average performance levels. However, this doesn't seem to be

true when the distributions of school ratings in the Nordic countries are compared. The variance of schools' social status is quite similar in all five countries.

However, there also exist real differences in schools' social status in the Nordic comprehensive school systems, and these differences have apparent effects on students' individual achievements. The higher the social status of a child's learning environment the better will be the results in reading literacy. Differences between countries may be explained by slightly different structural and historical features of the systems. In Denmark and Sweden parents' freedom of choice between schools appropriate to their children has a much longer tradition than, for example, in Finland where legislation to this effect was not changed until the end of the 90's. An interesting pedagogical conclusion could also be reached that it takes a long time to effect a change in a school's social structures. Following the introduction of freedom of choice for parents, the first change in a school is in its social status, but changes in, for example, the motivational structures of a school or the students' commitment to educational achievements take much more time. If this explains the differences in results between Sweden and Denmark on the one hand and Finland and Iceland on the other, serious attention should be paid both in research and in school development planning to strengthening positive social processes, particularly in those social and pedagogical environments where a high social status does not support a school's aspiration for high academic achievements.

References

Bryk, A., & Raudenbush, S. (1992). *Hierarchical Linear Models: Application and Data Analysis Methods*. Sage.

Ganzeboom, H., De Graaf, P. & Treiman, D. (1992). A standard international socio-economic index of occupational status. *Social Science Research,* 21, p.1-56.

Ganzeboom, H., Treiman, D. & Donald, J. (1996). International comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research,* 25, p.201-239.

Goldstein, H. (1987). *Multilevel Models in Educational and Social Research*. Griffin.

Goldstein, H. (1995). *Multilevel Statistical Models*. Edward Arnold.

Husén, T. (1975). *Jämlikhet genom utbildning?* Natur och kultur.

OECD. (1999). *Measuring student knowledge and skills. A new framework for assessment.* Paris: OECD Publications.

OECD. (2001). *Knowledge and for life. First results from the PISA 2000.* Paris: OECD Publications.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2000) *HLM 5. Hierarchical Linear and Nonlinear Modeling.* Scientific Software International.

Summers, A.A. & Wolfe, B.L. (1977). Do schools make difference? *American Economic Review*, 67, p.639-652.

Välijärvi, J. & Linnakylä, P. (eds.) (2002). *Tulevaisuuden osaajat. PISA 2000 Suomessa.* Koulutuksen tutkimuslaitos and Opetushallitus.

Willms, J.D. (1986). Social class segregation and its relationship to pupils' examination results in Scotland. *American Sociological Review*, 51, p.224-241.

Appendix table *Results of the two-level regression model (see text)*

	Denmark	Finland	Iceland	Norway	Sweden	Nordic Average	OECD Average
Constant	489 (2.9)	522 (2.6)	490 (3.4)	488 (3.3)	500 (2.2)	496 (1.3)	491 (0.7)
LEVEL 2 (school):							
MEANISEI	2.0 (0.4)	[0.3 (0.4)]	[0.1 (0.4)]	1.7 (0.4)	1.7 (0.3)	1.2 (0.2)	4.2 (0.1)
LEVEL 1 (student):							
FEMALE	25,0 (3.1)	52.4 (2.4)	38.6 (3.7)	42.3 (3.8)	37.3 (2.7)	39.2 (1.4)	25.9 (0.6)
HISEI	1.5 (0.1)	1.2 (0.1)	1.0 (0.1)	1.6 (0.1)	1.5 (0.1)	1.4 (0.04)	1.0 (0.02)
School level variance	740	420	526	627	269	540	1436
Student level variance	7115	5971	7014	8378	6756	7078	5534
Explained school level variance (%)	43.6	8.9	19.6	38.4	60.7	34.6	53.5
Explained student level Variance (%)	8.3	14.3	7.8	10.2	10.4	9.8	5.5

Note: Coefficients in ordinary brackets are standard errors.

Coefficients in square brackets are not statistically significant.

11 READING ACHIEVEMENT IN 1991 AND 2000

Peter Allerup and Jan Mejding

11.1 Comparing reading in IEA 1991 with PISA 2000

As described in chapter 2, reading literacy in PISA is defined as more than just decoding written material or literal comprehension. It incorporates understanding and reflecting on texts and using written information in order to be able to function effectively in a knowledge-based society. The following definition of reading literacy was used in PISA:

> *"Reading literacy is defined in PISA as understanding, using and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential and to participate in society." (OECD 2000, p.18).*

But nine years prior to PISA, in 1991, the International Association for the Evaluation of Educational Achievement (IEA) conducted another large international reading literacy survey. At that time the following short definition of reading literacy was formulated:

> *"Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual." (Elley 1992).*

As can be seen from the two definitions, the overall concepts of reading literacy share mutual properties: both focus on a broad description where the actual use of reading by the individual is a central issue. In the IEA Technical Report (Wolf 1995) Warwick Elley describes the background for the IEA reading literacy test in the following way:

> *"The notion of functional literacy, with its connotations of being able to use one's literacy skill to function effectively within one's own society was popular in the early discussions, but some NRCs wanted to extend the notion beyond the basic levels needed for survival, to include higher-level thinking and the reading of good literature, for example. (...) It was also proposed at the first NRC meeting, that a cross-section of topic themes should be included, representing tasks that are likely to be encountered at Home, at School, in Society at large, and at Work."*

This conception of reading literacy is very much in line with the description we find in the framework for reading in PISA (OECD 2000):

> *"Literacy for effective participation in modern society requires mastery of a body of basic knowledge and skills. For example, reading literacy depends on the ability to decode text, to interpret meanings of words and grammatical structures, and to*

construct meaning at least at a superficial level. But reading literacy for effective participation in modern society requires much more than this: It also depends on the ability to read between the lines and to reflect on the purposes and intended audiences of texts, to recognise deviances used by writers to convey messages and influence readers, and the ability to interpret meaning from the structures and features of texts. Reading literacy depends on an ability to understand and interpret a wide variety of text types, and to make sense of texts by relating them to the contexts in which they appear."

There is little doubt that it is – to a large extent – the same underlying reading competencies the two studies want to measure. It is therefore also of interest to investigate how the results from these two studies of reading compare. Seen from a pedagogical point of view, however, PISA has the advantage of describing thoroughly in a reading framework the concept of reading literacy, and it is therefore better suited as a guideline for future educational planning and research.

11.2 Two different tests of reading

There are differences between the two studies on how the test results were collected. The IEA study was solely reading study, and the test information was collected from two test booklets containing only reading texts and items. All students used the same booklets, and this simplified the process of calculating a comparable score between students and countries. PISA, on the other hand, gathered information on reading, mathematics and science, and the reading texts and items were distributed across nine booklets. Each student only answered one booklet (in two parts), and a comparable score had to be calculated on the basis of different subsets of the available reading test items in different mixtures of reading, mathematics and science items. This is, however, a technical problem only (more on this later), and does not invalidate the total score of a country, but it does make it more difficult to compare the data between the two studies.

In the IEA study the reading texts were classified according to the type of text: was it a narrative text, an expository text or could it be classified as a document? PISA - as is also described in chapter 2 - reports on the aspect of the task according to what the text and the questions ask the reader to react to: Do they need to retrieve information? Do they need to interpret the text? Or do they need to reflect on and evaluate the text content or the text format in order to get to and react to the information at hand?

Whereas the IEA study relied on the multiple-choice format in the calculation of a reading score, the PISA study has a more complicated process of reaching a score. Only about half of the questions in the texts can be answered in a multiple-choice or in short answer format. The rest of the questions are asked in an open answer format that requires the student to formulate and write his or her own answer. A team of trained markers then evaluate and categorize the answers as either right or wrong. Some of the

answers can be given more points if the content is more complex and if they show a highly qualified understanding of the issue in question. But eventually we end up with the same structure in this test as in the IEA reading literacy study: an array of questions that can be ordered according to their difficulty. The more difficult it is to get the right answer or to get to a certain level of answer to a question, the more points it will give you in the final score. And this last step of converting right and wrong answers to a comparable score is accomplished through item analysis and Rasch scaling, which will be presented later in this chapter.

Both studies conducted item analysis and reported their results on a Rasch scale with 500 points as the international mean and 100 points as one standard deviation. But even though this was done it is not the same as saying that getting a score of 500 is equally difficult in both studies. The international mean is always dependent on which countries participate in the study, as is the rank each country gets in the study.

Even though scores were calculated on different aspects in the two studies, a combined reading score was also reported. In the Nordic countries there was generally not much difference in the way the students scored in the different aspect scales compared to most countries in OECD (OECD 2001) and it is thus justifiable to look only at the combined reading score when comparing the Nordic countries as a whole.

How well are you doing then, if you get a score of 500? This is not easy to determine based on the score itself. The score will always be at an arbitrary level, and will not tell you actually what you are capable of doing. This is why in PISA 2000 the results are also reported for 5 reading levels. And the five levels are described in terms of the complexity of texts and questions associated with the level. At each level the percentage of the 15-year-old population in the country capable of solving tasks at that level of complexity is reported (OECD 2001). So how is it then possible to compare the results of the two studies? Let us look a little closer at that question.

11.3 Test equating

The need for statistical methods supporting the art of translating test scores from one test to another, test equating, has increased over the past decennium. Nordic schools have from time to time come out badly by international ranking studies and the public reaction to those studies has been substantial in some of the countries. Both IEA 1991 and PISA 2000 deal with students' reading skills, and we evidently need methods that enable us to equate the results of the two tests. Rasch item analysis, as developed by the Danish statistician Georg Rasch, is one of the ways to proceed.

11.3.1 The Rasch model

Georg Rasch repeatedly told a story to his students about how he was once confronted with the task of equating two different spelling tests. He had available only the students' test scores, no actual data from the tests (items) themselves, and from these scores he drew a simple X-Y scatter plot. This is the starting point for all kinds of analysis of co-variation between two test scores, and probably the means for even non-statisticians to 'translate' test scores from one test to the other. Instead of concentrating on the degree of correlation between the two tests, which is characteristic of classical statistical analysis of test equating, Rasch asked for more information about the responses to the single items on which the test scores were built. Because single item response data were not available to Rasch, he proceeded theoretically by setting up a set of assumptions for these unobserved responses. He wanted the X-Y plot to reflect something test-specific, i.e. something that was independent of the population of students (who in fact, by definition, influence the correlation).

These considerations led Rasch to a mathematical formalisation, in terms of a comprehensive statistical model, for the probability of responding 'correct' to each of the items (spelling the words correctly) in the test. From theoretical requirements concerning the interpretation of this X-Y plot, Rasch deduced a statistical model for responding to each of the items of the test. Problems concerning equating the two spelling tests were consequently transferred from problems displayed in the X-Y plot to problems concerning the structure of single item responses.

The simple Rasch model (Rasch 1960) for two response categories assigns a probability for student No. v to answer 'correct' for item No. i. In this model the *individual* student's 'ability' emerges through a parameter σ_v, which is *specific for student No* v, together with a measure *specific for item No i*, a 'difficulty' parameter θ_i. These measures are combined to determine the probability for a correct answer. The Rasch model is a statistical model for *single* item responses. It has to be emphasized that based on this model for the *distinct item response*, the statistical properties of the student test scores (i.e. the summed item responses) are *derived as mathematical consequences* from the model itself. Consequently, the statistical distribution of the test score is a distribution, depending on the individual parameter σ_v and the item parameters θ_I, which cannot be evaluated independently, irrespective of the model. (Readers interested in an elaborated statistical background for the Rasch model in relation to test equating, are referred to the literature e.g. Allerup 1994, 2002).

11.3.2 Test equating in the Rasch model

Under the Rasch model, test equating is defined as the process of transferring 'true scale' information regarding the σ 'abilities' from one test (test 1) to another (test 2) in such a way, that it takes care of the fact that item difficulties

θ_i may vary between the two tests, considering both 'content' related matters and item difficulties. The practical steps to ensure that this can be achieved are the following:

- First, the very existence of a σ-scale specific for test 1 is tested. This is done by exercising test statistics (Allerup 1994, 1995, 1997) for the fit of the Rasch model to the item level data for test 1. Notice that it is necessary to have access to the data at single item level.
- The same test procedure is repeated for test 2, testing the existence of a σ-scale specific for test 2. The two σ-scales need not be identical at this stage.
- On acceptance of σ-scales for each of the two scales, it is finally tested, if the two σ-scales are identical.

Under the Rasch model, test equating reflects a property of the two tests: It is hypothesized that items from test 1 can be merged with items from test 2 so that σ-abilities measured by the combined set of items remains the same as measured by the two tests.

One of the useful mathematical consequences of fit by the Rasch model to item level data is that the σ-scale can be estimated using *any subset* of the original items (Rasch 1960).

11.3.3 Equating IEA and PISA

Both the IEA and PISA studies conducted Rasch model analyses for the field trial data before main study data was collected. It is therefore assumed that the simple Rasch model can adequately describe the students' responses to the main study items.

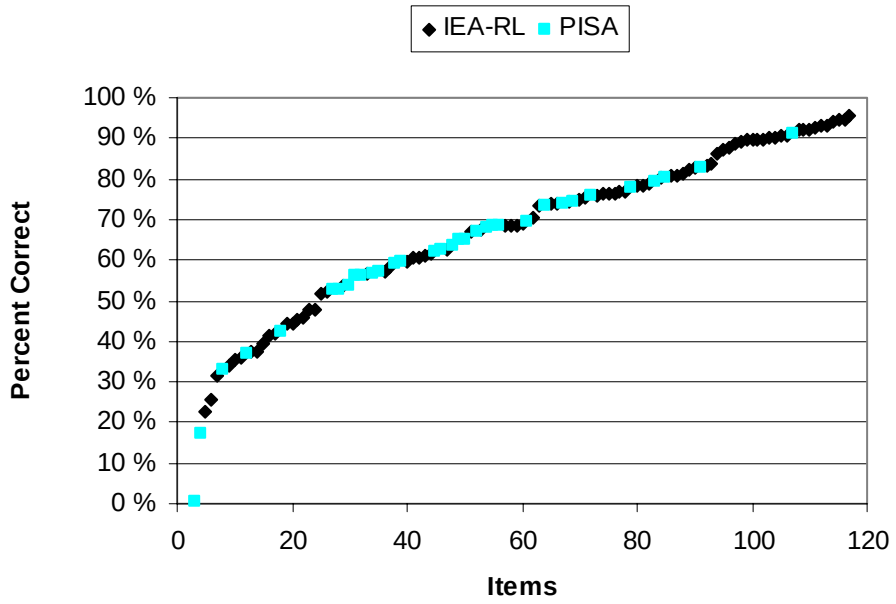
Some test equating procedures have, of course, already been undertaken prior to this attempt, when information from the nine different PISA booklets were combined into one reading scale and, likewise, when the results from the two booklets of IEA were combined. Conducting test equating in practice means, usually, that items enjoying a certain overlap cover all booklets. This was done most rigorously in the IEA mathematics and science study, TIMSS (Beaton et al. 1996). Here each booklet contained a common 'core' set of items in excess of the items specific for that booklet).

For the Danish PISA study a 10^{th} booklet was constructed containing a subset of items from both the IEA and the PISA tests and the Danish sample size was enlarged to accommodate for this extra booklet (keeping the 10^{th} booklet students out of the national PISA sample).

Although any selection of items from IEA and PISA can be used for the 10^{th} booklet, it is recommended that items representing a broad range of difficulties are selected. Figure 11.1 shows the item difficulty among the selected items in booklet 10, where items are arranged from the hardest items to the easier ones. (Each score point for a 2-point item is here regarded as a separate 1-point item.) The figure shows that there are sufficient items from both tests at both

ends of the difficulty scale to make a test equation between the two tests. It is therefore possible to measure the items from both studies on the same Rasch scale; thereby making the two scales for the two different studies comparable.

Figure 11.1 Distribution of difficulties (percent correct) for IEA and PISA items in Booklet 10



Having done this, it remains to be seen if the combined sets of test 1 (IEA) and test 2 (PISA) items in booklet 10 constitute one (so-called homogeneous) scale with one latent σ-ability scale. If this can be accepted by adequate statistical tests, the PISA items are thereby equated with the IEA items. This is, of course, a valid result for items in the supplementary booklet only. However, the result will also be valid for the PISA items in the other nine booklets as well, and for all items in the two IEA booklets. This is proved by the fact that both the PISA and IEA items have separately been 'approved' using the Rasch model.

A first point of analysis is, however, to explore whether the IEA reading items in the special 10th booklet are on the *same* scale of difficulty as those in the IEA 1991 study, in other words, whether the difficulty if the items, i.e. the θ-difficulty parameters in the Rasch model from IEA 1991, are the same as the θ-difficulty parameters found for the special 10th booklet in PISA 2000. It would, of course, be a problem if the reading scale had changed since 1991, i.e. if the relative difficulties had changed. In order to control for this, it is necessary to have access to old IEA 1991 data at the individual student and item level. Approximately 80 items were re-used in the PISA 2000 study.

The details of the statistical tests for consistency between the IEA reading scales will not be given here (Allerup 1995, 1997) but analysis shows, that the

IEA items have indeed kept their (relative) difficulties across the period. The so-called 'pop-B' IEA reading scale has stayed the same across the years 1991 to 2000. It is also worth mentioning that the national study 'F2000' (Allerup et al. 2001) made use of a subset of IEA items and found the same result. It is satisfying that the calibration (viz. the fit to the Rasch model) of a reading scale in 1991 seems to work years later.

The 1991 calibration of an international reading scale involved Denmark as only one part of the chain, but the international calibration required all countries to accept successful 'approval' of the Rasch model. It is therefore to be expected that the calibration can be extended to other countries as well.

There are two parts to booklet 10, presented to the students with a break in between. More insight into the details of the tests are required in order to see if this situation adequately satisfies the criteria set by the Rasch model. However, numerical statistical tests based on the conditional distribution of test score 1, given that total score = test score 1 + test score 2, adequately and 'objectively' (i.e. depending on the item difficulties θ only) confirm the impression gained from the figures:

The IEA items and the PISA items constitute one scale of difficulty, from which a common student ability σ-scale can be estimated.

This also appears to be valid when the two parts of booklet 10 are compared. Student abilities can therefore be estimated using any subset of the items, whether they originate from the IEA or the PISA study.

11.3.4 Danish results

From the successful fit of IEA 1991 and PISA 2000 reading items, an immediate question can be answered: Is the level of reading ability among 1991 14 year olds the same as for 15 year olds in 2000? The answer can be given either in terms of values on the 'true' latent σ-scale for student abilities or in terms of student scores. Without losing information these scores could be transformed to 'percentages correct' for each student. It is a matter of taste whether one prefers one or the other set of values. Both studies conducted item analysis and reported their results on a Rasch scale. The choice is independent of the Rasch model analysis, but will set certain limits for the reader concerning the interpretation of the results. In the case of the σ-scale scores, a difference between student achievements of, say '10.5 Rasch score (σ-values) points' will be the same, whatever the actual levels of the achievements of the students are. This will not be so if scores are reported as 'percentage correct', because a difference of, say, 5.5% between two students will not be the same if 5.5% is the result of 53.5% - 48.0% or 97.5% - 92.0%. However, in nearly all international studies the σ-values have been used with the values centred and re-scaled to different points on the real axis, e.g.: international mean = 500, international standard deviation = 100.

Table 11.1 Average levels of student performance in reading in Denmark measured using IEA 1991 reading items and three groups of students (rounded figures).

Study and population	Mean σ-values	St.dev. σ-values	Mean %	No. of obs. (N)
IEA 1991 DK (14 years)	525	78	74%	3743
PISA 2000 DK 10th booklet (15 years)	504	75	71%	492
F2000 (14 years)	505	71	72%	2556
PISA 2000 OECD mean	500	100	61%	13500

The answer to the question above is given in table 11.1, where both means of reporting are used. It is clear, that the average reading level of PISA 2000 15-year-old students in Denmark falls below the average level of the 1991 14-year-old students. These results are identical to the results from the national 'F2000' study (Allerup et al. 2001), which also demonstrated a significant decrease in reading levels from 1991 to year 2000 for 14-year-old students. The data in F2000 and in PISA 2000 were actually collected in the same spring. Although the PISA 2000 study was an age-based sample and F2000 (and IEA91) were grade-based samples and this comparison is based on cross-sectional and not longitudinal data, it seems as though in Denmark there is no increase in reading competence between the ages of 14 and 15.

11.3.5 Nordic results

In the 1991 IEA study 31 countries participated in the study. This is the same as the number of countries participating in PISA 2000, but only 18 countries participated in both studies. When student abilities (the σ values) are estimated from the data, the numerical values, the so-called Rasch scores derived from the Rasch analysis programs, are distributed approximately from –3.0 to 3.0. In order to create an international reference scale, 1000 students are selected from each country and the resulting σ values for all students are re-scaled, using a linear transformation, to the international mean of 500 and the international standard deviation of 100.

Any scaling procedure using a *fixed r*eference point (e.g.500) is, of course very dependent on exactly *which* countries are included in the calculations. Furthermore, fixing the mean to a certain value, 500, prevents immediate comparisons between repeated measurements taken using the same test. It has been shown that although PISA and IEA tests are different in content, they are psychometrically 'the same'. In that particular sense PISA 2000 is a repetition of the IEA 1991 study. It is, therefore, not possible immediately to judge from the two sets of results whether the countries evaluated 'as a whole' have done worse or better. Only the *relative ranking* of a country, seen in the context of participating countries, can be read from the international ranking tables.

In IEA 1991 Denmark's σ value was around 525, well above the international mean, while in the PISA study its score of 497 indicated a ranking a little under the international mean. According to these two measurements Denmark has dropped down the relative ranking list of participating countries.

There are at least two kinds of analysis to hand for evaluating the differences between IEA 1991 and PISA 2000. Restricting the analysis to only those 18 countries that participated both times is one way of calculating the differences. The other takes as its point of departure the fact that an 'absolute' judgement is available for Denmark, because of the PISA 2000 test design included IEA 1991 items in booklet No. 10.

Taking the latter method first, approximately 3500 IEA 1991 students and approximately 500 PISA 2000 students are available for analysis. Based on the information from booklet 10 it is estimated that the responses of the15-year-old PISA 2000 students *to the IEA 1991 items* were equivalent to 504 points. This value indicates a reduction equal to 525 minus504 which equals approximately 20 Rasch points.

The other analysis based on the 18 countries participating in both tests leads to the following calculations. First, it is noticed that the IEA 1991 average score for the 18 countries is close to 525 (incidentally the same as Denmark's score), while the PISA average for the same 18 countries is close to 503. The IEA partners are consequently 'weaker' competitors compared to the partners in PISA 2000. As such, the 18 countries as a whole have 'moved down' in the international ranking by approximately (25 – 3 =) 22 Rasch points. The first three columns of table 11.2 give the original scores for each country for the two tests and the difference ("Diff") between the two.

Subtracting 25 Rasch points from the IEA 1991 scores and 3 Rasch points from the PISA 2000 scores in table 11.2 (two first columns) will lead to revised values (revIEA and revPISA), which are both subject to the running standard: Mean = 500. It cannot, unfortunately, be guaranteed that the standard deviations for all students contributing to the scores in the two columns are equal to 100. In fact, it is not possible to ensure that this is the case unless the original 1991 analyses can be repeated for the 18 countries only. The last column in table 11.2, 'revDiff', is the result of measuring the change between the years 1991 and 2000 in the context of these 18 countries.

Table 11.2 International (IEA/PISA) - and re-scaled Rasch scores (revIEA) from IEA 1991 and (revPISA) from PISA 2000. Diff –measures are simple differences.

Country	IEA	PISA	Diff	revIEA	revPISA	revDiff
Belgium	481	507	26	456	504	48
Ireland	511	527	16	486	524	38
Canada	522	534	12	497	531	34
Spain	490	493	3	465	490	25
Norway	**516**	**505**	***-11***	**491**	**502**	**11**
Finland	**560**	**546**	***-14***	**535**	**543**	**8**
New Zealand	545	529	*-16*	520	526	6
Denmark	**525**	**497**	***-28***	**500**	**494**	*-6*
Italy	515	487	*-28*	490	484	*-6*
Iceland	**536**	**507**	***-29***	**511**	**504**	*-7*
Sweden	**546**	**516**	***-30***	**521**	**513**	*-8*
USA	535	504	*-31*	510	501	*-9*
Greece	509	474	*-35*	484	471	*-13*
Switzerland	536	494	*-42*	511	491	*-20*
Germany	526	484	*-42*	501	481	*-20*
France	549	505	*-44*	524	502	*-22*
Portugal	523	470	*-53*	498	467	*-31*
Hungary	536	480	*-56*	511	477	*-34*
Mean	*525*	*503*	*-22*	*500*	*500*	*0*

In table 11.2 there are now two educated guesses for assessing the change between IEA 1991 and PISA 2000: “Diff” and “revDiff”. One immediate reaction is, of course, that these two columns cannot both reflect the truth, since they are contradictory. However, notice from the PISA levels listed in column No. 2 that the international level for Denmark is 497, which is close to the booklet No. 10-based estimate of 504 reported earlier. The value 504 is obtained from a procedure based on approximately 500 students responding to *one* booklet, while 497 is based on many more students, adjusted across *nine* booklets. This may well account for the difference between 504 and 497. Accidentally, the 18 countries enjoy a common average close to the running standard of 500. Accepting the value of 503 as the running standard, it can therefore be concluded that the “Diff” IEA-PISA measures listed in column No. 3 reflect the change between IEA 1991 and PISA 2000 *in the context of all participating IEA 1991 countries.*

As mentioned above, the ‘revDiff” column measures the change *in the context of the 18 countries*. The two difference columns are therefore measuring two different things! No doubt, another third ‘revDiff’ column would emerge if another set of countries were to be selected as ‘partners’ for the international comparisons. With these comparisons it is like the rankings in international sport games: the change of position from one year to another depends on the competitors. And only on the competitors!

In the case of Denmark the results for IEA 1991, 525, and PISA 2000, 504, (497) emerge from the fact that we have access to student responses from the same students to items from both sets of test material. In the case of the other Nordic countries we do not have access to responses to both types of test material from the same students. In the absence of this data the only way of evaluating the change between 1991 and 2000 is by assessing the *relative* position of the country on both occasions. As can be seen from table 11.2 ("Diff"), Iceland and Sweden dropped significantly while Norway and Finland dropped less during the same period, *in the context of the IEA 1991-countries!*

Because of a fortuitous shift in competing countries between 1991 and 2000, both Finland and Norway made *relative progress* while Sweden and Iceland dropped slightly in the rankings ("revDiff" column in table 11.2), *in the context of the 18 participating countries.*

A few cautionary remarks should be added concerning sampling issues that have an impact on comparability. In IEA 1991 different populations in Canada (only British Columbia) and Belgium (only French), took part than in PISA 2000. For example, in Belgium in PISA 2000 the difference between the two language groups was 56 points in favour of Flanders. Therefore, their improved ranking does not necessarily reflect a real improvement. In the case of the Nordic countries, in 1991 only Finnish-speaking students in Finland were internationally assessed, whereas in 2000 the Swedish speakers were also represented in the sample. In 1991 as a national option, the Swedish-speaking schools were also assessed and they scored significantly lower in the narrative and expository domains. Since some "competitors" were different, detailed interpretations of relative standings should not be based only on the information in this chapter.

11.4 Summary

Two international studies, IEA and PISA, have investigated students' reading ability using data collected in 1991 and 2000 respectively. Even though both studies report their data on a proficiency scale with 500 as the international mean and 100 as one standard deviation this does not imply that the same score in both studies necessarily signifies the same proficiency level. The two studies focus on different aspects and types of reading ability, but both compute a combined score as an overall measure of reading proficiency. All five Nordic countries participated in both studies, and the reported subscales for reading did not differ much from the computed overall scores. It can thus be argued that use the combined reading score is justifiable as a means of comparison for investigating the development of reading proficiency over the time span of the two studies.

In order for a comparison to take place the two tests must be equated, and we explain how it is possible – in the light of the properties of the Rasch model – to adjust the scores to the same scale. To do this it is necessary to have access

to responses from the same students to items from both tests. It was therefore decided in Denmark to add an extra 10th booklet to the rotation of booklets in the PISA 2000 study enlarging the Danish sample accordingly. The 10th booklet contained a sample of IEA 1991 and PISA 2000 reading texts and questions (items), and as it was added to the rotation of booklets it was distributed to a random, representative sample of Danish 15 –year olds.

Because of the link between the two tests established from the Danish data it is now possible to expand this comparison of results to other countries participating in both studies. When this is done for Nordic countries it appears that Finnish and Norwegian 15-year-olds read with almost the same proficiency as did the 14-year-old students in these countries nine years earlier. In Denmark, Iceland, and Sweden, however, the 15-years-old students in 2000 read at a lower level than did their 14-year-old counterparts nine years before.

Unfortunately, it was decided in PISA 2000 not to include items linking to the IEA 1991 Reading Literacy study. The above analysis would have been even more accurate had we had access to student responses to items from both studies from all countries. It also shows the value of – from time to time – conducting studies with test material with well known properties. Had a link not been established for the Danish data there would have been no way of estimating the relative decrease in reading ability for the older students in Denmark, Iceland and Sweden. Fortunately we will have this type of data in the next cycle of PISA in 2003 and again in 2006. We will therefore in the future be better able to monitor changes in reading ability over time.

References

Allerup, P. (1994): Theory of Rasch Measurement; *The International Encyclopaedia of Education*, vol. 8, Pergamon Press.

Allerup, P. (1995): The IEA Study of Reading Literacy; in *Curriculum Assessment Issues: Messages for Teachers*, Falmer Press.

Allerup, P. (1997): Statistical Analysis of Data from the IEA Reading Literacy Study; in: *Applications of Latent Trait and Latent Class Models in the Social Sciences*; Waxmann.

Allerup, P. (2002): *Test Equating*, paper presented at: 7th Roundtable on Assessment. Conference on “Test Equating”, Canberra.

Allerup, P., Mejding, J. & Zeuner, L. (2001): *Færdigheder i læsning og matematik – udviklingstræk omkring årtusindskiftet*, Evaluering af Folkeskolen år 2000. Copenhagen: Undervisningsministeriet, Denmark.

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzales, E.J., Kelly, D.L., & Smith T.A. (1996): *Mathematics Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.

Elley, W.B. (1992): *How in the world do students read?* The Hague: International Association for the Evaluation of Educational Achievement.

OECD PISA (2000): *Measuring Student Knowledge and Skills. The PISA 2000 assessment of reading, mathematical and scientific literacy*. Paris: OECD Publications.

OECD PISA (2001): *Knowledge and skills. First results from PISA*, Paris: OECD Publications.

Rasch, G. (1960): *Probabilistic Models for some intelligence and Attainment tests*, Copenhagen: The National Danish Institute for Educational Research.

Wolf, R.M. (ed.) (1995): *The IEA Reading Literacy Study: Technical Report*. The Hague: International Association for the Evaluation of Educational Achievement.

12 UNITY AND DIVERSITY OF READING LITERACY PROFILES

Svein Lie and Astrid Roe

12.1 Focus on item-by-country interactions

The present chapter deals with similarities and differences between countries concerning patterns of student responses to reading items in PISA. Based on an analysis of percentages of correct responses for each item we will focus on the extent to which Nordic students have similar patterns of strengths and weaknesses. In particular we will display and try to understand the differences between the Scandinavian and Finnish reading profiles.

The PISA project represents a typical quantitative and comparative study, and one of the main goals is to establish valid and reliable estimates of student achievement. Within each of the subject domains, student scores and weights are calculated to optimise the way the student samples represent the populations and how the student data can be aggregated to the country level. Regardless of the measurement model (the Rasch model is being used in PISA, see chapter 11) the main idea behind a test score is that responses to individual items are not interesting beyond their contributions to the overall score. From a psychometrical point of view, the details about how students from different countries respond to individual items, often called "item-by-country interactions", should be regarded as a sort of random noise, or "error variance" (Olsen et al. 1991). Large item-by-country interactions for single items are often avoided, mainly because of the fear that they may reflect a particular cultural bias or even translation issues that represent challenges to the reliability of the test.

Seen from a different perspective, the details of these interactions represent something very interesting, namely a guide to the strengths and weaknesses of each country. Thus a matrix of percent correct responses for all items and for all countries provides a valuable starting point for looking for similarities and differences between countries. The item-by-item sets of percent correct responses (hereafter called "p-values") establish highly interesting country-specific educational "fingerprints". It is the aim of the present chapter to use such sets of "fingerprints" in an analysis of similarities and differences between Nordic countries in reading literacy. Our main questions are: To what extent can the Nordic countries be established as a robust group? And further, which countries are central in challenging our unity? Finally, we will also consider

which other stable country groups can reasonably be formed, and what are the possible grouping mechanisms.

Based on data from the IEA TIMSS study in 1995 (Beaton et al. 1996a & b) some analyses of similarities between countries in science and mathematics have been carried out (Zabulionis 1997, Lie et al. 1997, Kjærnsli et al. 2002, Kjærnsli & Lie 2002). These analyses clearly show that groups of countries can be established based on similar patterns of responses. For science and mathematics it can clearly be seen that the country grouping has to do with both language and other cultural - and even geographical - factors. For instance, the close similarities between English-speaking countries are clearly partly due to precisely the fact that: they are English-speaking countries. But there are also strong cultural traditions in these countries that develop students' self-esteem and the courage to express their opinions. On the other hand, the "East Europe" group consists of countries with very different languages, but with some obvious common cultural traits (Kjærnsli et al. 2002).

12.2 Methods and results

12.2.1 Correlations between countries

As mentioned above, our starting point was a matrix of p-values by item by country. A few countries had data missing for some items, due to a mismatch with the Rasch model caused by translation or printing errors, or for some particular cultural reason. Each blank cell was replaced by an expected value based on the international p-value for the item and on the particular country's score on *all* items. Given that we wanted to pursue the fine structure, we then calculated the cell residuals by subtracting from each cell value the average over countries for the actual item and the average over items for the actual country. Thus we were left with a residual matrix, where each cell indicates how much better or worse than expected that particular country scores on that particular item. The fact that some countries score higher than others and that some items are harder than others no longer shows up in the data.

The first set of results is simply the correlations between residuals in each Nordic country and all other countries. The correlation coefficients are shown in table 12.1. To simplify reading of the table, significant correlations are shown in bold (correlations 0.15 or higher) or italic (-0.15 or lower). It can be seen from the table that there is a clear linkage between the Nordic countries, particularly between the three Scandinavian countries (Denmark, Norway and Sweden). Iceland correlates significantly with Finland, Norway and Sweden, but not with Denmark in spite of their close historical and cultural links. Furthermore, Finnish students seem to be only weakly linked to their Scandinavian peers, but more strongly linked to Iceland, Germany and Switzerland. It is worth mentioning that there exist strong historical connections between Finnish and German pedagogy.

Table 12.1 Correlations between Nordic and all countries. Significant positive correlations are bold, significant negative are italicised.

	DENMARK	FINLAND	ICELAND	NORWAY	SWEDEN
AUSTRALIA	**0,20**	0,00	0,01	-0,01	0,13
AUSTRIA	0,07	0,10	-0,06	-0,09	0,09
BELGIUM	-0,04	0,14	*-0,17*	-0,03	0,04
BRAZIL	-0,10	*-0,27*	-0,07	-0,13	*-0,24*
CANADA	0,08	*-0,18*	-0,04	-0,04	-0,12
CZECH REP.	-0,15	0,04	-0,05	-0,11	*-0,24*
DENMARK	**1**	-0,11	-0,05	**0,21**	**0,20**
FINLAND	-0,11	**1**	**0,19**	0,00	0,08
FRANCE	-0,08	-0,06	*-0,17*	-0,09	-0,02
GERMANY	**0,22**	**0,15**	-0,01	-0,03	**0,17**
GREECE	-0,06	*-0,15*	0,07	0,06	-0,13
HUNGARY	*-0,17*	0,13	*-0,21*	0,05	-0,09
ICELAND	-0,05	**0,19**	**1**	**0,15**	**0,19**
IRELAND	0,14	-0,05	-0,09	0,08	0,02
ITALY	*-0,29*	-0,07	-0,11	*-0,31*	-0,08
JAPAN	*-0,15*	**0,15**	**0,15**	-0,01	0,07
KOREA	*-0,15*	-0,01	0,06	-0,04	-0,08
LATVIA	*-0,18*	0,05	0,07	*-0,21*	*-0,30*
LIECHTENSTEIN	-0,02	0,11	-0,11	-0,02	0,00
LUXEMBOURG	0,10	0,08	**0,17**	0,09	**0,24**
MEXICO	*-0,21*	*-0,33*	-0,11	*-0,24*	*-0,36*
NEW ZEALAND	**0,22**	-0,06	-0,10	0,03	0,08
NORWAY	**0,21**	0,00	**0,15**	**1**	**0,43**
POLAND	-0,01	-0,03	-0,04	*-0,17*	*-0,26*
PORTUGAL	*-0,16*	*-0,23*	-0,03	*-0,19*	*-0,18*
RUSSIA	*-0,18*	0,04	0,06	*-0,19*	*-0,17*
SPAIN	-0,02	*-0,22*	*-0,21*	0,02	-0,03
SWEDEN	**0,20**	0,08	**0,19**	**0,43**	**1**
SWITZERLAND	0,10	**0,16**	-0,11	0,08	**0,16**
UK	**0,22**	-0,05	-0,11	0,03	0,12
USA	0,09	*-0,20*	0,04	-0,03	0,03

Some other features are worth mentioning. Firstly, the Danish students seem to share some common characteristics with their English-speaking peers (as well as with Germany), particularly in the UK, Australia and New Zealand, whereas this is not the case for the other Nordic students. Secondly, all three Scandinavian countries have in common some particularly large negative correlations with Latvia, Russia, Portugal and Mexico. It should also be pointed

out that a relatively strong negative correlation with Mexico is a common feature for all the Nordic countries.

12.2.2 Clustering of countries

Another way of looking at similarities and differences between countries is to try to establish country clusters based on the residual matrix. Cluster analysis is a useful tool for this purpose. Instead of using correlations as a measure of similarity, cluster analysis allows us to calculate "distances" between countries in a number of possible ways. First one must decide on how distances between countries are to be calculated. The most usual measure is the (straight or squared) Euclidian distance, in which the distance between two countries is calculated from the sum of squared differences between residuals. Since contributions from each item are squared, possible outliers (e.g. translation weaknesses) will have a large influence, so this method is avoided here. Instead, we have applied what is often called the "Block" distance, which simply consists of adding the absolute differences between residuals.

Next we need to establish a rule to decide which countries or country clusters should be combined, according to their "nearness", at each step. As the first step, the two "nearest" countries form a group. As the next step, two other countries group together, or one country links to the group already made. It will make a difference to this process how distances from one country to a country group are measured. Here we have used the average distance between all members of the group. Similarly, the distance between two clusters is defined as the average of the distances between all pairs of cases in which one member of the pair is from each of the clusters. Using the alternative "nearest neighbour" method to measure the distance would have been less robust, since we would rely more on a small number of distances for each cluster.

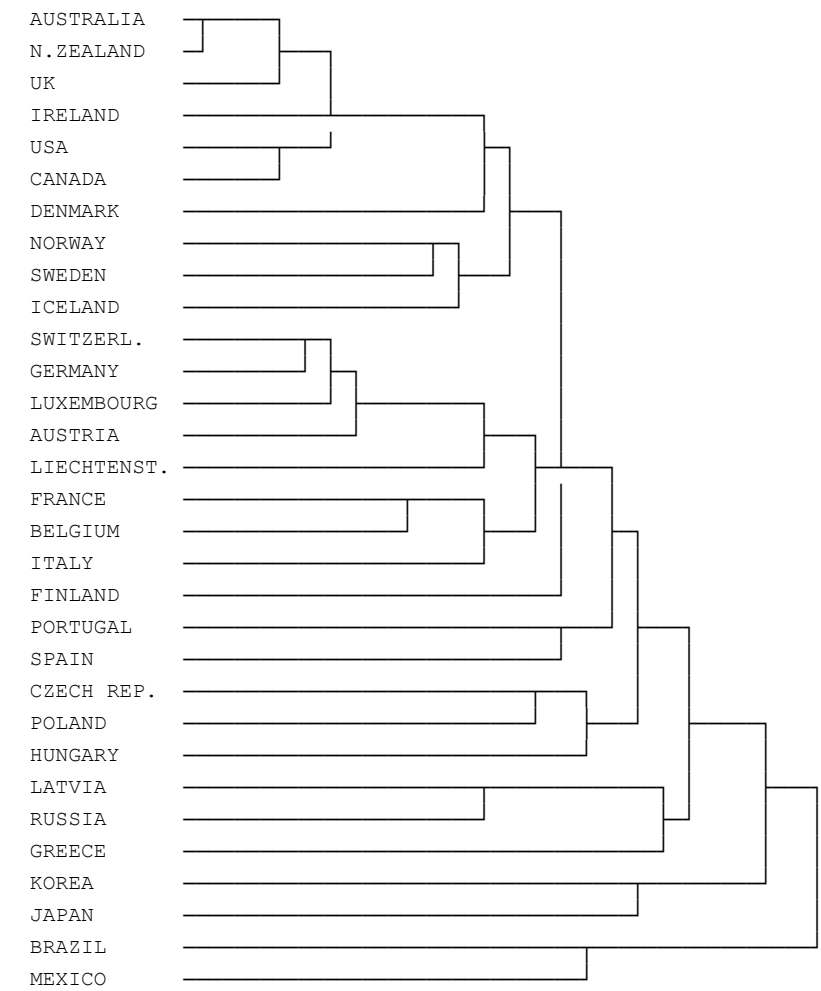
Figure 12.1 is a so-called dendrogram that displays the resulting clustering process. This figure shows how and at what "distance" countries link together into clusters. The following comments refer to what happens when we move to the right, i.e. to larger distances. The two most similar countries are New Zealand and Australia, and the UK joins these two somewhat "later". At about the same "time" the USA and Canada come together and eventually they join with Ireland to form an English-speaking cluster. At about the same "time" Switzerland and Germany combine, followed by Luxembourg and Austria, and eventually also by Liectenstein to create what one is tempted to label a German-speaking cluster, even though Switzerland and Luxembourg are multi-phonic countries. Similarly, France, Belgium (Flemish and French parts combined) and Italy form a cluster, but that cluster is less easy to label (it could possibly be called a romance language group?).

What mainly concerns us here is how the Nordic countries behave in this clustering process. As can be seen from the figure, none of the Nordic countries are particularly close to each other, but Norway and Sweden do link together,

and a little later Iceland joins the two in a Nordic cluster. Denmark appears to be "torn" between two nearby clusters, the Nordic and the English-speaking clusters, and ends up joining the English-speaking cluster, largely due to the attraction of the most "English" members, the UK, New Zealand and Australia (see table 12.1). The fact that Denmark is drawn to the English-speaking cluster is not a robust finding, and could well be changed by changes in the clustering criteria discussed above.

Finland is displayed as a rather different and atypical Nordic country in this analysis. Only rather late in the clustering process does Finland join the mega-cluster formed by the combination of all the above-mentioned clusters.

Figure 12.1 Dendrogram for country clustering (see text for explanation)



12.3 Differences between the Nordic countries

If we look at the Nordic countries in particular, we find from table 12.1 and figure 12.1 that there is a relatively strong linkage between the Scandinavian countries (Denmark, Norway and Sweden). On the other hand, Finland seems to be more closely related to Iceland, Germany, Switzerland and Japan than to any of the Scandinavian countries. How can these results be explained? Are the differences and similarities mainly language-based or could there be more important cultural explanations? The language theory is only relevant for some of the groups of countries, namely the Scandinavian and the English-speaking ones. Finnish, German and Japanese are perhaps as different as languages can be, so there are good reasons to look for additional explanations. In the following section we will focus on how Finnish student responses in reading differ from the responses from the Scandinavian countries, taken as one group. Furthermore, we will look for possible cultural explanations for this difference.

We will explore the characteristics of the 25 items where the difference between Finland and the average for the Scandinavian countries is greatest in either direction. For each item we will compare the residuals obtained as explained earlier, namely the measures of the countries' performances on individual items relative to what is expected based on the overall results. The 25 items where the Finnish residuals are furthest ahead of the average Scandinavian residuals will be termed "Finnish items". Likewise, the 25 items where the residuals for Scandinavia exceed those for Finland the most will be called "Scandinavian items". These items will be compared to all the items in the whole test, here named "All items".

12.3.1 Item categories

There are many ways of categorising the PISA items. We will start by looking at the item categories presented in chapter 2, table 2.1, which are identical to the categories used in the PISA framework. Here every item is classified by text structure, text type, reading context, item aspect, and item format. In the category *text type* there are five different non-continuous text types (charts/graphs, forms, maps, schematics, tables). These are collapsed into one text type, *non-continuous,* because each of the five categories represents very few items.

Table 12.2 shows the percentage distribution of items within each text category. As far as text structure is concerned, the table shows that there is no noticeable difference between the distributions of the sets of items. If we look at the text types, the differences are more striking for two of the categories. Firstly, there is a much higher percentage of items connected to *argumentative* texts among the Scandinavian items than among the Finnish items. Items connected to argumentative texts are also over-represented among

Scandinavian items compared to All items. Secondly, Finnish items are more dominated by items connected to expository texts than are Scandinavian items. Among the Scandinavian items there are in fact very few connected to expository texts. Finally, when it comes to reading context there is no striking difference between the student groups.

Table 12.2 Percentage distribution of items connected to different text categories within Scandinavian items, Finnish items and All items

		Scandinavian items	Finnish items	All items
Text Structure	Continuous	76	68	69
	Non-contiuous	24	32	31
Text type	Argumentative	36	8	14
	Descriptive	8	8	9
	Expository	8	28	24
	Injunctive	4	12	7
	Narrative	20	12	14
	Charts/graphs, forms, maps, schematics and tables	24	32	31
Reading context	Educational	40	24	28
	Occupational	12	20	15
	Personal	24	28	21
	Public	24	28	36

Table 12.3 shows the distribution of items classified by the two item-specific categories, item format and item aspect. Looking at the first category, compared to Scandinavian items and All items there is a slightly lower percentage of *multiple choice items* and a higher percentage of *short response items* within the group of Finnish items. The differences between Scandinavian and Finnish students are more striking (as well as statistically significant, Chi-square test, $p<0.01$) when the reading aspect is considered. There is a remarkably high percentage of *reflect items* and a very low percentage of *retrieve items* within Scandinavian items, particularly compared with Finnish items, but also compared with All items.

The percentage distributions shown in tables 12.2 and 12.3 give some indication of characteristic differences between Scandinavian and Finnish students concerning reading literacy. One preliminary conclusion could be that Scandinavian students seem to be relatively good at reading and reflecting on argumentative texts, and correspondingly poor at retrieving information from expository texts. However, this would be an over-generalisation based on the findings above, which do not provide sufficient evidence to draw any conclusion about what the differences really mean. Looking at differences related to text content and task difficulty may take us a step further.

Table 12.3 Percentage distribution of item categories within Scandinavian items, Finnish items and All items

		Scandinavian items	Finnish items	All items
Item format	Multiple choice	44	28	47
	Short response	20	40	22
	Open constructed response	36	32	31
Item aspect	Retrieve	12	40	29
	Interpret	48	48	49
	Reflect	40	12	22

12.3.2 Content, difficulty and 'booklet effect'

The texts in PISA are not classified by content. There are, however, several ways of creating text categories based on content. One way of classifying the texts that would make sense in PISA is by audience. We find that there is a clear distinction between texts that are written for a young audience and texts written for all age groups. Furthermore, there are some texts that were not originally meant for the PISA study, namely the texts from the International Adult Literacy Study (OECD 2000).

Nearly one third of the texts in PISA are explicitly related to young people's lives or interests in one way or another. These texts are either written by teenagers, about teenagers or directly addressed to teenagers. Most of the texts in PISA are not written for a specific age-group, they are texts that could be of interest for any reader, regardless of age. All the narrative texts and most of the expository texts in PISA are in this category. The last "audience category" of texts contains the items that are not originally designed for the PISA study but for a study meant for readers from the age 16 to 65 (IALS).

Table 12.4 Distribution of items connected to texts based on the "audience categories".

	Scandinavian items	Finnish items	All items
Young readers	56	16	29
IALS	0	28	12
All readers	44	56	59

Table 12.4 shows the percentage distribution of items by the audience categories described above. The table shows that items connected to texts written for a young audience are strongly (Chi square, $p<0.01$) over-represented among Scandinavian items, whereas Finnish students seem to be particularly good at reading the texts from the IALS study for adult readers.

Next we will focus on item difficulty. In PISA, the best measure of item difficulty is the so-called threshold value. This value can simply be explained as the ability level of students with 50% probability of getting the item right. The measure of ability level used here is the PISA ability scales (see chapter 1). The average threshold value (489 points) for the 25 Scandinavian items is much lower than the average threshold value for the Finnish items (540 points).

A final way of categorising the items is by test booklet. Each of the nine test booklets consisted of four clusters of items. In booklets 1 – 6 the reading items represented the first three clusters while the last cluster was either mathematics or science. In booklet 7 there were reading clusters only, and in booklet 8 and 9 the two first clusters were mathematics and science while the last two were reading clusters. Two clusters were only located in the last part of booklets 7, 8 or 9. The remaining seven clusters were only located among the three first clusters of booklets 1 – 7. It was reported after the analyses of the PISA 2000 results that scores for reading items that were located in the last part of the booklet were significantly lower than the average of all countries. This was referred to as 'the booklet effect'. By dividing the reading items into 'booklet effect' or 'no booklet effect', it will be possible to see if Scandinavian and Finnish students are equally affected by the effect.

The analysis showed that 24 percent of all items are located in the last part of the booklets. Among the 25 Finnish items there are 50 percent from the last part of the booklets. Among the Scandinavian items only 16 percent are from the last part of the booklets.

12.3.1 Discussion

Earlier in this chapter we suggested that Finnish students are better at retrieving information from expository texts. After examining differences between the two groups based on text content and task difficulty we also found that Finnish students are very good at interpreting and retrieving information from texts that are not written for a young audience; furthermore, they perform very well on the most difficult tasks. Among the tasks where they outperform their Scandinavian peers most, there are very few tasks that demand reflection and evaluation.

It seems as if Finnish students do not give up on tasks that they find difficult or boring to the extent that Scandinavian students do. They seem to have the energy and self-discipline to keep their concentration throughout the test. Finnish students also seem to be very good readers in general. To be able to answer the most difficult questions properly, as the Finnish students do so well, one needs to read thoroughly and understand the full meaning of what is written and be able to read between the lines. On the other hand, Finnish students do not seem to manage equally well the tasks where they are asked to state their own opinions, i.e. when the information is not to be found in or interpreted from the text.

Scandinavian students seem to have strengths in quite different fields. They are quite good at reflecting on and evaluating argumentative texts, especially when the tasks are connected to texts that are written for young people and are not among the most difficult ones. They seem to have problems with texts that they find uninteresting, boring or complicated. Finnish students, on the other hand, seem to be capable of mastering these texts surprisingly well.

These findings raise several questions: Does the ability to perform well on difficult tasks connected to demanding texts reflect the more hardworking and disciplined nature of students in Finland? Does the lack of ability to manage difficult and boring tasks indicate that Scandinavian students generally manage to do only the tasks that they enjoy and escape boring tasks?

If the answer to the questions above is yes, is it then possible to ascribe the reasons for the differences between Finnish and Scandinavian teenagers to cultural differences between Finland and the Scandinavian countries? Are Finns in general more serious and hard working than Scandinavians? Is what you achieve at school more important from the point of view of getting a good job in Finland? Are Scandinavian teenagers generally more spoiled and irresponsible, because they believe they will get a well paid job anyway? Have parents and teachers in Scandinavia lost their authority?

The Norwegian historian Eli Moen states that Finland has become a Great Power in advanced technology, while Norwegian industry is still mainly based on the production of raw material. In an interview, "The curse of luck" (2003), she discusses to what extent historic and cultural factors are decisive for economic development in a country. The luck of having had easy access to rich natural resources like wood and oil has lead to what she calls the laid-back attitude towards the economy and politics that one finds in Norway. "Why should we refine our oil when we earn money just the same?" she asks on behalf of the Norwegians. She holds that Norwegians do not seem to feel the same pressure to develop new methods for manufacturing and refining as happens in Finland. She sees the same kind of attitude in connection with the development of intellectual resources. A combination of luck and anti-intellectualism has led to a lack of emphasis on research and education in Norway. Finland, on the other hand, has over the last 200 years been struck by more wars and famines than Norway, and people have had to struggle to survive. The Fins have learnt to be prepared for the worst and to trust themselves. This mobilises resources in a completely different way, she claims (Løvhaug 2003, Moen 1999).

Data from PISA 2000 contains information from both the student and the school questionnaires about what happens in the classrooms. However, there is insufficient information to answer the questions we asked above. Nevertheless we suggest that there could be a motivation factor involved, and that it could have something to do with differences in the school culture. If future PISA assessments include more questions about school culture, this hypothesis can be tested more rigorously.

References

Kjærnsli, M., Angell, C. & Lie, S. (2002). Exploring Population 2 students' ideas about science. In: Robitaille, D.F. & Beaton, A.E. (eds.): *Secondary Analysis of the TIMSS Data* Kluwer Academic Publishers, Dordrecht/ Boston/ London. p. 127-144.

Kjærnsli, M. & Lie, S. (2002). TIMSS Science Results seen from a Nordic Perspective. In: Robitaille, D.F. & Beaton, A.E. (eds.). *Secondary Analysis of the TIMSS Data,* Kluwer Academic Publishers, Dordrecht/ Boston/ London. p.193-208.

Lie, S., Kjærnsli, M. & Brekke, G. (1997). *Hva i all verden skjer i realfagene?* (Norwegian national TIMSS report). Oslo: Department of Teacher Education and School Development, University of Oslo.

Løvhaug, J. W. (2003). Norsk utakt: Flaksens forbannelse. In *Apollon*. Research Magazine for Oslo University.

Moen, E. (1999). *Globalisering og industripolitiske strategier – En sammenligning av Finland og Norge*. På http://www.sv.uio.no/mutr/rapp1999/Rapport41.html

OECD (1999). *Measuring students knowledge and skills. A new framework for assessment.* Paris: OECD Publications.

OECD (2000). *Literacy in the Information Age. Final report of the International Adult Literacy Survey.* Organisation for Economic Co-operation and Development, Paris, and the Minister of Industry, Canada.

Olsen, R.V., Turmo, A.& Lie, S. (2001). Learning about students' knowledge and thinking in science through large-scale quantitative studies. *European Journal of Psychology of Education.* Lisboa. 16 (3), p. 401-418.

Taube, K. & Munck, I. (1996). Gender differences at the item level. In Wagemaker, H. (ed.): *Are girls better readers? Gender differences in Reading Literacy in 32 Countries*. Amsterdam: International Association for the Evaluation of Educational Achievement.

Zabulionis, A. (1997). *The formal similarity of the mathematical and science achievement of the countries.* Vilnius University pre-print No 97-19. Vilnius University.

13 FUTURE CHALLENGES TO NORDIC EDUCATION

Pirjo Linnakylä, Astrid Roe, and Svein Lie

International assessments like PISA reveal more clearly than national tests the special characteristics of a nation's educational culture as well as its relations to other school cultures. From close up it is often more difficult to see where the strengths lie and what potential for improvement there is, and also what is weak, stagnant, problematic or challenging in the culture. Likewise, international assessments provide an opportunity to learn about other education systems and their methods of solving problems in education and to assess the effectiveness of these solutions. For smaller nations, international studies also provide an opportunity to display their education system and school culture in an international context and to build links with other countries while developing their own system, curriculum and teaching style.

13.1 Reinforcing high quality and equality

The PISA 2000 results have shown that the Nordic education systems have proved reasonably successful in providing the majority of students with a solid foundation in the core subject areas of the comprehensive school. This means that young people in the Nordic countries are well prepared for further studies, for transition to working life and for full participation in the knowledge society. It is to be hoped that the outstanding success of Finnish students will give rise to educational discussions in all the Nordic countries as to how we can share our knowledge and visions, solve our common problems and develop our schools jointly for the benefit of all the Nordic nations and the young people responsible for their common future. The average levels of achievement for the other Nordic countries in some domains, and for Norway in all domains, has raised concerns and requests for various educational reforms. And there is no doubt that the PISA results have provided important insight into possible strategies in this process.

The process of standardising student scores internationally used in PISA does not allow easy comparison of student achievement with earlier assessments. As shown in chapter 11, however, there is ample evidence that the level of reading literacy has actually decreased internationally, as well as in the Nordic countries. Even if international assessments provide an opportunity to compare with international standards, it should be clear that no absolute "standards" can be established simply by averaging country scores. A lowering

of reading ability in the industrialised world is an educational challenge which has to be viewed in the light of other educational and social trends, like the ICT revolution and the increasing role of "infotainment". Therefore, in future phases of PISA measurement of trends will play an important role.

The PISA results also show that an education system can succeed in combining high quality performance with a high level of equality, and this was evident in most Nordic countries, particularly in Finland, but also in Iceland and in Sweden. However, the pursuit of equity will have to be a major objective in the future development of Nordic schooling, because at the same time social and cultural reproduction seems to be gaining ground across the Nordic countries. This is especially the case in Norway and Denmark, where the relationship between students' socio-economic background and school achievement proved surprisingly high.

The Nordic comprehensive school systems display certain differences as far as the social status of schools is concerned. Furthermore, these differences have apparent effects on students' individual achievements. In Denmark, Norway and Sweden, these school effects were stronger than in Finland and Iceland. Differences resulting form the school effects between the countries may be explained by the slightly different structural and historical features of the systems. In Denmark and Sweden, the freedom of parents to choose the school appropriate for their children has a longer tradition than, for example, in Finland, where it was not until the late 1990s that the legislation was changed. Following an increase in freedom of choice for parents, a change in the social status of a school takes place first whereas changes in, for instance, the school's motivational climate or the students' commitment to educational achievements seem to take more time. If this is what accounts for the differences found in the results between Sweden, Norway and Denmark, on the one hand, and Finland and Iceland, on the other, serious attention should be paid both in research and in school development to the strengthening of positive social processes, particularly in those social and pedagogical environments where the aspiration of schools for high academic achievement is not supported.

The depth of the Nordic principle of equity, however, has recently been put to a severe test owing to the increasing numbers of immigrant students and a growing cultural heterogeneity. In order to be able to tackle this aspect of equity, Finland and Iceland, in particular, will have a lot to learn from Sweden, which has had ample experience of immigration both in the past and in the present.

13.2 Pedagogy according to the individual student's needs

Even though the performance of Nordic students in PISA 2000 proved high quality, there is still room for improvement. Such improvements, however, presuppose an increased focus on the individual needs of students. In PISA, for

example, in every Nordic country, and particularly in Denmark, Norway and Iceland, more than 15 per cent of the students were found to have severe difficulties in coping with the reading literacy demands of today's knowledge society. By international standards, this proportion is not particularly high. If a principal goal is equal learning opportunities for all, however, it should be considered high. Moreover, it is definitely far too high in a society where every individual's mental growth and the nation's competitive edge are based on the ability to build up knowledge about and competence in life-long learning, domains in which literacy skills play a fundamental role. The sounder, more comprehensive and more equal this foundation, the better it promotes the individuals' quality of life and the economic growth of the nation, and the more effectively it prevents polarisation or marginalisation in terms of further studies, work, and social and cultural life. The Nordic comprehensive school must therefore continue struggling to minimise the proportion of students at risk. Joint efforts should be focused on understanding various types of learning difficulties, both neuro-psychological and socio-cultural, and their background factors in order to develop innovative support and rehabilitation programmes to conquer learning problems.

The PISA findings also suggest that it is imperative in the development of Nordic pedagogy to see the close relationship between the cognitive and affective elements of learning. Poor readers need both affective and cognitive support. Affective support can be increased through real-world interaction, by such means as authentic, interesting and exiting texts, personal choice of books and other reading materials, experiencing the joy that can be derived from reading and peer collaboration. These instruments can contribute to the affective elements of instruction and may also help to developing effective cognitive strategies. Interest and activity factors, if evenly spread, could also significantly reduce the gender gap and improve both the equality and the quality of reading literacy achievement. In reducing gender differences the Danish pedagogy could serve as a model for the Finnish school system, in particular, while the Finnish pedagogy could provide expertise in reinforcing more general interest and active engagement in reading. The PISA results also suggest that the Nordic countries could learn something, particularly with regard to boys' reflective and evaluative reading, from Anglo-American pedagogy, especially the pedagogies of Australia, Ireland, New Zealand and the United States. The PISA results for these countries showed smaller gender gaps than the results of most of the Nordic countries in reflective and evaluative reading, in particular. This was the domain in which all the Nordic countries face a serious challenge.

In developing reflective and critical reading, linguistic and rhetoric knowledge of texts should be systematically enhanced through instruction, so that students are able to see beyond the surface and realise how various texts are constructed. Students need to understand how authors use language, style and structure to produce a certain effect; how metaphors and implicit information are embedded in texts; and how inferences can be drawn by

reading beyond and between the lines with a reflective and critical mind. Depending on the type of text, and the purpose and situation of the reading, students have to learn to apply various reading strategies. Continuous texts demand a different strategy compared to non-continuous texts. In reading continuous texts, fiction requires a different strategy from that of expository prose. Furthermore, tasks calling for retrieval of information require different approaches from those involving developing an interpretation or reflecting on texts. To accomplish all this, teachers and test developers need to know a lot more about authentic texts as well as about tasks that are typical of boys' and girls' lives outside school. It is also possible that Nordic countries have something to learn form each other as far as different approaches to reading strategies are concerned. Finnish students significantly outperform their Scandinavian peers in reading literacy in PISA and when it comes to retrieving information from difficult expository texts the difference is strikingly large. The Scandinavian students, on the other hand seem to be better at reflecting on and evaluating argumentative text written for a young audience.

In the Nordic countries, interest and engagement in reading, especially as concerns reading fiction, are evidently considered part of a feminine culture., Nordic pedagogues responsible for curriculum development and instruction should invest heavily in breaking this cultural code, and should jointly develop a Nordic reading and literature pedagogy that looks for new ways of approaching literature, with a view to inspiring boys and helping them to realise that reading fiction can be enjoyable and interesting. Furthermore, parental involvement should be encouraged, and parents, particularly fathers, should be made conscious of the role model they provide for their sons. Young people's eyes need to be made aware that even 'a real man' reads, and not only newspapers, reports or Internet texts but also books, including fiction. On the other hand, electronic texts, which play a significant role in students' free time activities, should also be a part of the every-day learning environment. In today's learning environments, students need and should be able to develop skills to search, contrast, combine and critically evaluate information as well as to share, argue and reflect on text contents and forms. Using electronic texts in instruction may also be a way of arousing the interest of 'nerds' - usually boys - in reading and even in reading fiction, which provides endless possibilities of enriching the imagination and experiencing other worlds - both desirable and undesirable - with a critical eye.

There is also, however, the possibility that the PISA reading literacy tests have been biased in favour of girls. Compared to the IEA Reading Literacy Study conducted in 1991, the gender difference in favour of girls has increased. In PISA, the largest gender differences were found in items connected to continuous texts, especially long narratives and argumentative and injunctive texts, which are not typical of boys' reading material. Girls likewise seem to have an advantage in the open response items, where they can express their understanding and reflect in their own words, which again points to the fact that writing skills may play a certain role in the PISA assessment. Boys are not

at such a disadvantage when it comes to non-continuous documents such as charts, maps and figures. These texts have a relatively small amount of written information, yet on the other hand, they require the ability to understand, contrast and combine various types of information as well. In future assessments, texts and tasks may have to be redesigned to correspond to both girls' and boys' genuine reading interests and response preferences.

The gifted students have traditionally received meagre attention in the Nordic countries, and this lack of attention has often been mistaken for the opposite, namely the concern shown for the least successful students. Yet, it seems that even the development of students with widely differing knowledge and skills can be appropriately enhanced in heterogeneous groups as long as the teacher is willing and capable enough and has sufficient resources for within-group differentiation. Again, the pedagogies of England and New Zealand could be worth studying as part of the development of a pedagogy that takes full advantage of the whole potential of gifted students in various subject areas. Japanese and Korean instruction methods could also stimulate new ideas for enriching both conceptual and process approaches in mathematics and science instruction. This, of course, should be done without jeopardising special support for students with learning difficulties. It is reasonable to suggest that instruction aimed at enhancing the scientific literacy of both low and high achieving students should emphasize science as a specific culture and a specific way of reasoning and thinking. This includes the idea that time must be spent on interpretation, reflection and discussion on how science relates to the larger context of community and society.

13.3 Enhancing self-regulated and lifelong learning

In PISA 2000, self-regulated learning was chosen as a core construct of cross-curricular competencies, which included self-concept, motivation, learning strategies and learning styles. Self-concept, in turn, contained self-efficacy, which was understood as the students' own judgements of their capabilities for attaining standards. This was clearly strongest among Swedish students. Academic self-concept, on the other hand, was highest in Denmark which had a score close to the international maximum. Finnish students, however, scored lowest in both self-efficacy and academic self-concept, which is interesting considering the fact that their scores were above the international means. This might be connected to the high expectations of Finnish teachers and parents concerning academic achievement. Even though high expectations may have a positive effect on performance, educational achievement should be judged not only by the students' present skills and knowledge but also on how they affect the students' beliefs about their capabilities for future learning. These beliefs are known to affect the students' motivation for future studies. Nationally, self-efficacy and particularly academic self-esteem correlated strongly with literacy performance. The correlations tended to be stronger in the Nordic countries than elsewhere.

Students' various learning strategies also showed a minor relationship with achievement in the Nordic countries. High performance was accordingly associated with Nordic students' above average awareness of their own learning habits and their ability to control the learning process. In today's world, the ability to apply effective learning strategies – identifying, consciously selecting and controlling the efficiency of these strategies in the various subject areas, tasks and contexts of the school – will continue to be one of the major pedagogical challenges for our schools. The PISA findings suggest that the flexibility to be able to adapt various strategies relevant to the situation and purpose of a learning activity is of great importance, more so than how frequently various learning strategies are used. Thus, in the assessments to come learning strategies should also be examined, together with various learning tasks and contexts.

Students' learning styles were assessed from a co-operative and competitive learning point of view. Competition seemed to support high achievement, while co-operation supported achievement more generally. Competition likewise correlated more strongly with self-efficacy than co-operation. That is, students with a high self-efficacy tended to prefer competition, whereas the opposite was true for students with a low self-efficacy. Thus in learning environments with a strong focus on competition, for instance where students are frequently being assessed, the high-achieving students benefit at the expense of low achievers. Co-operation has a greater chance of improving the self-efficacy of low achievers as well.

The Nordic results showed that Danish students achieved close to the international maximum in co-operation and were highest in competitive learning. Iceland had the lowest mean score in co-operation, whereas Finnish students showed the lowest use of competitive learning strategies among the Nordic countries. Nationally, competitive strategies correlated considerably more highly with the reading achievement than did co-operation. Norway had the strongest correlation, while Denmark and Sweden had the weakest. We should be cautious, however, about putting too much emphasis on these correlations. Here, again, the findings concerning the various learning styles should be examined as they relate to various learning tasks and situations and various types of students. In addition, we should also consider what kind of influence various learning styles have on students' social skills, which may be as important as academic achievement in their future.

13.4 Warming up the Nordic school climate

When we combine all the descriptive measures of school climate and correlate them with reading performance in PISA 2000, the correlations seem to be quite similar across the Nordic countries. Schools with high average score points in reading are characterised by supportive teachers with good relations to their students and by classrooms with a good disciplinary climate, where the students do not feel that the pressure to achieve is too high. However, the

school principals' and students' reports on the above phenomena showed some interesting differences between the Nordic countries.

The Nordic students' descriptions of the student-teacher relations and the learning environments are generally positive from an international point of view. There are, however, some remarkable differences between the Nordic countries. Norwegian (and to some extent Finnish) students express a less positive view than do their Nordic and most of their international peers. In Norway, the rather problematic picture that has been drawn of noisy classes with little teacher support and low teacher expectations has been noted as a particular concern in the national educational debate. Obviously, such a picture has to be viewed in the context of a Norwegian national curriculum that strongly enhances student-centered and even student-initiated activities.

The Nordic principals' perceptions of school climate proved, on average, more negative than those of the Nordic students. There were also, however, some remarkable differences between the Nordic countries in the principals' views. In Denmark and Iceland, the principals were quite satisfied with the students' behaviour; whereas in Finland, Norway and Sweden dissatisfaction was more widespread. Expectations may vary because the school climate is positively correlated with reading achievement at national levels. In many previous studies the school climate has been found to influence the students' motivation for and attitude to future learning. School principals, in particular, should therefore be conscious of the possibilities for stimulating teachers and students in order to improve the well-being of their own schools and to work towards a warm and innovative learning environment.

The results presented in this report have shed light on different aspects of the education systems in the Nordic countries. We have seen how the Nordic countries are similar in some respects and different in other. Some countries seem to be more successful than others in obtaining equity together with a high quality of learning outcome for their students. However, each Nordic country has particular qualities that other countries can learn from, and even a country that is generally successful in most fields may have something to learn from other countries.