

AKTUELLA ANALYSER 2009

With a focus on Reading Comprehension

An analysis of the differences and similarities between
international comparative studies and national syllabuses



Skolverket

With a focus on Reading Comprehension

An analysis of the differences and similarities
between international comparative studies
and national syllabuses

www.skolverket.se
ISBN: 978-91-85545-60-5
Form: Ordförrådet AB

Preface

The Swedish National Agency for Education (Skolverket) has, in a previous summary (*Internationella studier under 40 år* [‘International studies over a period of 40 years’, trans.]), described Sweden’s participation in international comparative studies of knowledge. From this it is evident that the scope of Sweden’s participation has expanded significantly during the past decade. During this time, the aim of these studies has shifted from describing and comparing the results of different countries – and possibly also from explaining and understanding results in light of background factors on different levels – to focusing increasingly on monitoring change in knowledge over time, i.e. trend measurements. Interest in trend studies has also increased nationally in recent years. The major national evaluation which was done in 2003 (NU-03) can be viewed as an example of this.

It is costly, however, to perform assessment studies and in many cases a complex methodology and plan is required in order to secure good quality. For this reason it is advantageous to be able to share the cost of method development and analysis with many participating countries. At the same time there is, of course, a risk that internationally developed instruments are not applicable to the Swedish situation.

Is it possible to use the international studies to measure trends from a national perspective? How well do the frameworks for the international studies agree? Do we risk losing our distinctive national character? Are our pupils’ results judged on the basis of misleading principles? Do our national objectives differ from the discernible objectives of the international studies?

In the present study four assessment studies have been analysed: PIRLS 2001, RL 2001, PISA 2000 and NU-03. The aim of this study is to examine the similarities and differences among the frameworks for PISA, PIRLS, and RL with regard to reading literacy on the one hand, and the syllabuses for Swedish with a focus on reading comprehension on the other. Exercises in the international studies and in NU-03 have also been studied and related to the syllabuses.

Similar comparisons in mathematics and the natural sciences will be accounted for in a later report. The results reported are intended to support the Swedish National Agency for Education’s assessment of how the evaluation of knowledge is to be formulated.

The report, which has been developed within the framework for the project *Analys och syntes* [‘Analysis and synthesis’, trans.], was written by Gustav Eriksson, Department of Scandinavian Languages, Uppsala University. Views have been expressed by Karin Hector-Stahre, Kristian Ramstedt and Anita Wester.

Stockholm, 12 June 2006

Ann Carlson Ericsson
Department Manager

Anita Wester
Project Manager

Content

1	Introduction and objectives	6
1.1	About the studies.....	7
1.2	Outline.....	7
2	About reading and reading comprehension	8
3	Syllabus for Swedish	11
4	Measuring reading comprehension	14
5	Studies – framework, concretion and results	19
5.1	RL and PIRLS	19
5.2	PISA 2000.....	25
5.3	National evaluation	31
6	The Agency's assessment	37
7	Literature	41

1 Introduction and objectives

The ability to read and understand is highly valued in our society. All public administration places major demands upon the ability of its citizens to assimilate the written word. The written language is a communication medium which dominates to a large extent the daily acquisition of all forms of knowledge. Regardless if we are to assimilate advanced theoretical models, or the date information on a milk carton, we must be well acquainted with the form and function of the written word; an acceptable and meaningful existence is dependent largely upon our literacy. Entry into the community of the written language can be the beginning of a lifetime of learning where time and space may cease to be determining factors. With the help of the written language readers can assimilate texts that were written in another time, in another space.

The significance of the written language is described in detail in the steering document¹ for the Swedish compulsory school. As the chapter below will show, there is a great deal of focus on literacy. These skills are some of the keys to opening the doors to life as an active member of a society. Given the significance of the written language it is of considerable value to learn how Swedish schools manage to educate their pupils to become active citizens in accordance with said arguments.

Four studies have been analysed in the present study: PIRLS 2001, RL 2001, PISA 2000 and NU-03. The objective of this study is to examine the similarities and differences between the framework for PISA, PIRLS and RL with regard to literacy on the one hand, and the syllabuses for Swedish with a focus on reading comprehension on the other. Exercises in the international studies and in NU-03 have also been studied and related to the syllabuses.

The studies have different authors. The OECD is behind PISA (Program for International Student Assessment), the IEA organisation is behind PIRLS (Progress in International Reading Literacy Study) and RL (Reading Literacy), and finally the Swedish National Agency for Education has produced NU-03 (the Swedish National Evaluation of the Compulsory School in 2003). Both PIRLS and PISA report that Swedish pupils hold their own from an international perspective. PIRLS shows that pupils in their third school year are among the best participants and above average. PISA reports that pupils in their ninth school year perform above the OECD average. Only three countries show better results. That Sweden performs well from an international perspective is thus evident. The RL study reports, however, that the level of reading literacy has declined and that Swedish pupils' performance was significantly poorer in 2001 than in 1991. Even NU-03 reports a decline.

Each study has advantages and unfortunately also disadvantages. The variation potential in the design of the studies is one explanation; it is possible to define the population in different ways (for example by age or grade) and also to use samples of different size, with different kinds of texts, and with different kinds of instruments. The objective of the present report is to present a highly concrete account of the content of these studies with regard to their texts and

¹ This report deals especially with compulsory schooling.

instruments. The studies will be compared with one another from two perspectives. The theoretical basis for each study will be related to the syllabus, and the actual design of the study will be related to the syllabus as well as to national tests. By discussing the aforementioned studies in relationship to, in the first place, the national test in Swedish for grade 9, but also the national test in grade 5, we will be able to form an opinion about how to make use of their results.

1.1 About the studies

Two of the studies presented below were done in the early 2000s. The other two are trend studies, which were done just after the millennium shift, as well as at the beginning and middle of the 1990s respectively. The purpose of these studies differs somewhat. PISA includes references to how pupils are being equipped for a forthcoming life as active citizens. This is also the case in *PIRLS*. The primary objective of the *RL* and *NU-tests*, however, is to measure literacy over time.

The studies are all well-reasoned, but offer partly different designs. We should study what follows with an awareness that none of the studies are 'incorrectly' constructed. On the other hand, they can correspond more or less well to the national syllabus for Swedish and the design of the national test in Swedish.

1.2 Outline

Chapter 2 accounts for theories about the assimilation of written language and reading comprehension. In order to partake of the execution of these studies it is necessary to have certain basic facts, which illustrates the complexity of the concept of *reading*. Chapter 3 introduces the syllabus for Swedish. The studies cited are more or less relevant depending upon the extent to which there is a clear correlation with the syllabus. This is to say that relevance is relative; the studies can uphold scientific validity and still fall short of the requirements presented in the syllabus. In order for the studies to be relevant in Sweden they must test what is referred to in the 'goals to attain' in the syllabus. In the latter case, it is necessary to question whether or not the studies test what it is said that pupils in the Swedish schools should be capable of. Chapter 4 gives a picture of the way reading comprehension is tested today. The national test of Swedish serves as a base. The studies themselves are accounted for with concrete examples in Chapter 5, where their theoretical framework, results, and any correlation to the syllabus are presented. Chapter 6 concludes the report with a summary and discussion of the goals of the syllabus versus the content of the theoretical framework of the studies.

2 About reading and reading comprehension

Extensive research is being done in the field of reading comprehension. There are different theories about how children assimilate the written language and how they then best learn to understand text. It is clear that reading and writing development are important parts of the school system. This has been pointed out in political contexts, for example by setting up a national centre for language, reading, and writing development. The Minister of Schools and Adult Education, Ibrahim Baylan, (2005*) has explained the aim of the centre: "The objectives of the centre are to contribute to the elevated competence of teachers in pre-schools and compulsory schools, to promote contacts between research and the school system, and to encourage an interest in and motivation to read and write among children and young adults in pre-schools and compulsory schools."

There are various ways to grasp the term *learning* which refers to so many different things. It is important to make a distinction between decoding and understanding. The best way of assimilating the written language is a controversial issue, and as with so much other research there is no complete agreement in the area. A clear dividing line is, however, difficult to distinguish; views have converged in practice and there is a prevailing consensus that a combination of various procedures yields the best results. (Liberg, 1993:13)

For the sake of simplicity, it can be said that there are basically two main positions with regard to the reading process. These were represented most explicitly by Ingvar Lundberg and Åke Edfeldt. These two positions have come closer to one another in recent years. (ibidem) While Lundberg (1984) is of the opinion that reading itself, i.e. the decoding, is the fruit of an analytical process, Edfeldt (1982) has a synthetic bias. In this context, the process is often referred to as *bottom-up* or *top-down*. A word is apprehended either by analysing first the units, i.e. sounding out the parts in order to apprehend the whole (*bottom-up*), or by apprehending the word as a whole, which is then analysed (*top-down*). When Liberg accounts for how these views have been compromised, it is in terms of *combination*, *interaction* and *flexibility*. (1993, p. 13)

When the decoding itself is automatic, i.e. when the reader does not have to think about how he or she goes about reading, it is a question of understanding how the words interact with one another and how to comprehend what is written. Åke Edfeldt (1982) states that: "The reader does not comprehend at the moment of decoding." (p. 88) An automatic reading process is one thing, but understanding the significance of what one reads is another.

Understanding a text is to some extent dependent upon one's familiarity with the medium itself. To *speak like a book* is a figure of speech, hardly indicative of reality. The sender and the receiver interact in a conversation. The speaker can make use of gestures, tone of voice, and use other words or expressions upon request to explain and clarify the meaning of what is said. The writer, on the other hand, must rely on his or her ability to anticipate questions and convey the intended meaning as clearly as possible. The text itself is the only as-

sistance the writer can give. The naturally large differences between the written and spoken language take time to assimilate. In addition, the written word must vary more than speech in order for the reader to have the energy and interest to continue reading.

Lundberg (1984) has described the meaning of *reading comprehension*. He elucidates the significance of *syntactical competence* in the interpretation of a text. He mentions, for example, forms which seldom occur in spoken language: "In the written word, for example, children may soon encounter the appositive construction 'Jan, the boy in a brown shirt, has a new ball'." (1984, p. 93) The words between the commas constitute an apposition, i.e. in this case it is something that has to do with *Jan*. Given greater familiarity with the medium itself, i.e. as syntactical competence grows, it becomes a matter of perceiving what is hidden in the text or, in other words, "reading between the lines". (1984, p.94) Even if one is well acquainted with the form of the written language, it is necessary to make use of one's knowledge of the surrounding world in order to understand a text. Once again we borrow an example from Lundberg, who emphasizes that understanding a text is not a matter of interpreting sentence by sentence. Sentences are connected and Lundberg illustrates with: " 'Lisa came flying down the stairs. Kalle rushed off to get a doctor.' Most readers would interpret this to mean that Lisa fell down the stairs, and that Kalle ran to get help because she was injured." If the first sentence had been followed by another ('Karl rushed up and kissed her') then the significance would have been dramatically different. The reader must make *inferences*, i.e. add more information in order for the sentences to connect correctly. The reader constructs a scenario from the text and makes use of inferences in order to understand. Taking another example from Lundberg:

'Per went to the cinema on Wednesday. He ran into an old friend.' The reader automatically infers that Per met his friend on Wednesday at the cinema. The first sentence provides the spatial and time framework for the interpretation of the second sentence.

According to Lundberg, another important component of reading comprehension is text association. When encountering the sentences: " 'A bus came rushing down the hill. The bus nearly hit a truck.' " the reader must understand that it was the same bus that came rushing that nearly hit a truck. The definite form, *the bus*, is the syntactical signal that is conveyed, a so called reference. Sometimes these must be combined with inferences: " 'Anna unpacked the lunch basket. The juice was warm' ". The definite form of *juice* indicates that the information refers to the given circumstances. But since the word *juice* is not mentioned, an inference must be made, by which the reader associates *juice* with the probable contents of a lunch basket. (1984, p. 95f)

Much more could be said about reading comprehension. This review is not, however, intended to comprehensively account for the complexity of the topic. Before leaving the subject, a few comments will be made regarding Lundberg's theories on the diagnosis of reading comprehension.

According to Lundberg there are two main underlying reasons for poor reading comprehension: (1) poor decoding skills and (2) poor general knowledge of the world, required understanding, and poorly developed classification in conjunction with reading texts in a particular field. It can be difficult to construct a

diagnostic test to determine which skills need to be reviewed. In tests designed with short-answer, multiple-choice, and fill-in-the-blank questions, for example, it is difficult to discern the reader's problems. And diagnostic tests with many inference questions risk unfair treatment of pupils from "socially and culturally impoverished homes". (1984, p. 109) It is evidently a question of diagnostics and not an exam. However, opinions can be worth taking into consideration.

3 Syllabus for Swedish

The theoretical bases for school work are the curriculum and the syllabuses. These steering documents reflect views on knowledge and indicate the goals and focus of the education. Below is an outline of the set requirements for a pupil to receive the grade 'Pass' in Swedish. The emphasis is on reading comprehension.

The concepts *goals to strive towards* and *goals to attain* are established in the steering document *Lpo94*. These can be seen as interdependent. The goals to strive for indicate the focus of school work – the education – and the goals to attain account for the minimum acceptable outcome of the education.

A great deal of importance is associated with literacy in *Lpo94*. The school system is charged with the responsibility for "promoting the pupils' learning and thus preparing them to live and work in the society." This is where language plays an important role. In work to develop the ability to be an active participant in society *Lpo94* states that: "Language, learning, and the development of a personal identity are all closely related. By providing a wealth of opportunities for discussion, reading and writing, all pupils should be able to develop their ability to communicate and thus enhance confidence in their own language abilities. What is mentioned in the curriculum applies of course not only to the subject Swedish. Other subjects can, of course, also offer opportunities for communication with the kind of results mentioned above. It is, however, in the subject Swedish that these elements are the centre of focus and emphasis. In reading the curriculum for Swedish one should accordingly pay notice to these introductory paragraphs..

The 'goals to aim for' and 'goals to be attained' are stated in the *Syllabus for Swedish*. As in *Lpo94*, reference is made, of course, to the significance of language: "One of the school system's most important tasks is to create good opportunities for pupils' language development." It is further asserted that "The significance of the written word is great and society places demands on one's ability to cope with, evaluate, and assimilate texts." If we pause to consider this description of the subject, it can be interesting to comment briefly on its significance. The school is expected to prepare the pupil to for an active life in society. One of the more important tasks of the school system is to create the opportunity for language development, which then leads to the ability to live and work in the society. To this end, the actual conditions in the school as well as the way in which the knowledge is measured, are interesting. The national tests of subjects ordered by the Swedish National Agency for Education are expected to measure "useful" knowledge, as will be discussed later.

As a point of departure the goals 'to aim for' described in the syllabus provide a sense of what teaching should concentrate on. Teaching should be such that the pupil develops a desire *to read to learn* and a desire to read on his or her own. It should also develop a desire to use language to create. The importance of developing the imagination of the pupil is also emphasised here. Among the 'goals to aim for' it is stated that pupils with adequate linguistic competence. "have the courage, desire and ability to express themselves in many different contexts, and by means of writing and speech acquire an instrument for thinking, learning, communicating and exercising influence."

It is important that the school strives to ensure that pupils "develop their ability to read, understand, interpret and experience texts of different kinds and adapt their reading and work on texts to its purpose and character." Their work should also be directed towards deepening insight into the basic patterns and grammatical structures of the language, its development and origins.

Teaching should also mean that pupils "develop the ability to use different opportunities to obtain information, acquire knowledge of the language and functions of the media, as well as develop their ability to interpret, critically examine, and evaluate different sources and their content."

The goals that pupils should have attained by the end of the fifth school year include: "to be able to read with fluency, both aloud and to themselves, and understand events and meaning in books and non-fiction written for young children and young persons, and be able to discuss their reading experiences, as well as to reflect over texts." They should also be able to apply the most common rules of the written language and to produce texts for different purposes.

By the end of the ninth school year pupils should also "be able to read literature appropriate to their age from Sweden, the Nordic area, and other countries, and also read non-fiction and newspaper articles on general subjects, as well as be able to reproduce the contents coherently and also reflect over this." By the end of the ninth school year, pupils should also be able to produce texts for different purposes.

It is apparent that the syllabus does not present particularly specific goals to attain. The general character of the text does not mean, however, that demands are unclear. Nor does it mean that it is difficult to measure whether or not pupils have attained the goals. The syllabus places major emphasis, as previously stated, on preparing a pupil for a life as an active member of the society. For those who wish to absorb the substance of social debate, for example, it is necessary to know how to interpret information and to adapt reading to the situation; the latter is a 'goal to aim for' that is emphasised in the syllabus.

In addition, the text around the goals to be attained is worth noting. It points, as mentioned previously, to what teaching should include. It states, for example, that pupils should be able to read different kinds of texts, understand the Swedish language and its structure, deepen their insight into basic patterns and grammatical structures, and apply the standards of the written language. Furthermore, it states very concretely that pupils should be able to write legibly and use computers as an aid. Finally, demands for insight into learning and the development of one's ability to critically examine texts are also worth mentioning. Bear in mind, however, that these demands are *not* the same as the goals to be attained. This is what the pupil should strive to attain and what teaching should include. For one who wishes to read and interpret the 'goals to be attained', the 'goals to aim for' is a good tool.

Goals that pupils should have attained by the end of the fifth year in school

Pupils should

- be able to read with fluency, both aloud and to themselves, and understand events and meaning in books and non-fiction written for young children and young persons, and be able to discuss their reading experiences, as well as reflect over texts.

- be able to produce texts for different purposes as a tool for learning and communication,
- be able to orally relate and present something so that the contents are understandable and brought to life,
- be able to apply the most common rules of the written language and the most common rules of spelling, as well as be able to use dictionaries.

Goals that pupils should have attained by the end of the ninth year in school

Pupils should

- be able to actively take part in conversations and discussions, as well as be able to share the thoughts of others, and be able to present work orally so that the contents are clear and understandable,
- be able to read literature appropriate to their age from Sweden, the Nordic area, and other countries, and also read non-fiction and newspaper articles on general subjects, as well as be able to reproduce the contents coherently and also reflect over this.
- be able to read, reflect over, and put into a context some literary works and authors that have been influential in affecting how people live and think,
- be able to appreciate and reflect over and evaluate the contents and means of expression used in pictures, films and the theatre,
- be able to write different kinds of texts so that the contents are clear, and in doing this apply the standards of the written language, when writing by hand and using computers,
- have a knowledge of the language which makes it possible to carry out observations of both one's own and others' use of language.

The syllabus and the studies

When partaking of the PISA, PIRLS and NU-03 studies, it is important to carefully read certain parts of the syllabus. Perhaps we should focus on the following quote, which states that pupils should: "be able to read literature from Sweden, the Nordic area, and other countries, appropriate for their age, and also read non-fiction and newspaper articles on general subjects, as well as be able to reproduce the contents coherently and also reflect over this." In order for any great importance to be attached to these studies, it is reasonable to expect exercises to be constructed in such a way that the latter demands are measured. In the next section, the theoretical framework for these studies and how this is reflected in the exercises is presented. The exercises in the studies are then compared with those of the national test of the subject.

4 Measuring reading comprehension

It should be possible to compare the exercises in PIRLS, PISA and NU-03 with the content of the syllabus. Since this report is focused on reading literacy, some of the goals listed in the syllabus will be excluded from the subsequent discussion.

It would be reasonable to expect the national test of Swedish to be a standard formula for testing the attainment of the goals of the syllabus. In this context it is interesting to closely examine the exercises in the test, which is the basis of the argument below. Firstly, it should be pointed out, however, that it is ill-advised to believe that the current form of the national test will always be applicable. Without attempting to predict the future we can rest assured for the time being that today's test is a reasonably good instrument.

Measuring the demands in the syllabus places, of course, demands upon the test designer. What exercise is capable of revealing the extent to which pupils can "read literature appropriate to their age from Sweden, the Nordic area, and other countries, and also read non-fiction and newspaper articles on general subjects, as well as be able to reproduce the contents coherently and also reflect over this."? None is probably the right answer. A good test of reading comprehension must contain several different *kinds of texts*. The questions must be well-balanced; information can be coherently rendered if one has read a text in order to report information and thus concentrated on what is central in the text. The pupil must be given several different ways to show that he or she has understood. How well the pupil is able to coherently report a course of events, reflect, interpret and draw conclusions about a text is tested also in the pupil's own formulation of the answer. It is also important that the test is constructed in such a way that the pupil does *not* have to make extensive use of knowledge about the surrounding world in order to answer the questions. It is primarily the *ability to read* and not an ability to read based upon a broad general knowledge that is the focus of attention (see the section on Reading and reading comprehension). A conventional sort of exercise: *Explain what happened and the consequences. Motivate how you would have reacted in a similar situation* would fulfil the demands cited above. As previously stated, the demands of the syllabus are not at all unclear, and should be possible to test.

Example: National test

The Advanced Studies in Modern Swedish group (FUMS) in the Scandinavian Languages Department at Uppsala University is commissioned by the Swedish National Agency for Education to construct the national tests. The national test for the ninth school year is the only mandatory test in the compulsory schools and can be described as an accepted way of testing proficiency in reading comprehension, as well as oral and written presentation. Since the purpose of PIRLS and PISA is to measure reading comprehension, only the reading comprehension part of the national test is accounted for here. As far as this report is concerned, the national test serves thus as the basis for the comparison between PIRLS and PISA as well as the trend study in NU-03. After the section on the ninth school year test, there are comments on the fifth school year test (which are relevant to the comparison with PIRLS).

There are three parts of the test: A reading, B oral work, and C written work. Part A is to be completed in approximately 80 minutes and uses a booklet with a new theme each year. The booklet that was presented at the time data was gathered for NU-03 contained nine different kinds of texts on the theme of *change and reflection*, with specific questions on five of the texts. The two final questions on this part of the test were open questions, in the sense that the answers could relate to all of the texts. Since the purpose of the national test is to measure the ability to read different kinds of texts, both fiction and non-fiction are represented. The various sorts of texts and related exercises are illustrated below.

The text booklet contains excerpts from *Populärmusik från Vittula*. [‘Popular Music from Vittula’, trans com.] This particular excerpt talks about the hardships associated with forming a music duet between the main character and his friend. Under peer pressure, for example, they perform in school during a lesson reserved for entertainment by the pupils, the so-called “fun hour”.

There are five exercises associated with the text. The first is the kind where the pupil is asked to tick a ‘true’ or ‘false’ box indicating whether or not the information is correctly reported. The purpose of the third exercise is for pupils to use their own words to formulate something which is not expressly stated in the text. The exercise asks how the narrator sang when he was alone and in the text it says: “I ... opened my mouth. and began to sing. ... But it was a song without sound, just my lips were moving ... I was moving on the inside to the music ...” In the fourth exercise, the pupil is asked to motivate the reaction anticipated by one of the test authors to an incident that occurs with one of the secondary characters in the text. Here the pupil is to draw some conclusions about the incident and reflect on its significance, despite the fact that it is not expressly stated in the text. The fifth exercise deals with a metaphor, the significance of which the pupil is asked to and explain.

The Equal Opportunities Ombudsman Claes Borgström has written the next text, *Skärp dig, sa Elin* [‘Wise up, said Elin’, trans.], which is an informal article in *Pedagogiska magasinet* [‘Pedagogical Magazine’, trans.]. Borgström reasons about institutions and people who have affected his choice of profession. He reflects on what has made him the person he is.

This text, which is considerably shorter than the previous one, is followed by three exercises. The first two ask the pupil to draw conclusions from the text and reflect on the meaning and significance of various passages. The third exercise is to find the meaning of the word “*JämO*” [abbrev. for Swedish Equal Opportunities Ombudsman, trans.] in the text. This involves a (relatively) simple ability to make inferences.

A section consisting of 5 exercises has been constructed using three text passages, which in one way or another have to do with physical health: (1) “Extra physical education on the schedule improves grades” is an article published in the magazine *Svensk idrott* [‘Swedish sports’, trans.], which describes a school model where there are opportunities for extra physical activities. (2) “Facts about physical inactivity” accounts for the National Institute of Public Health’s definitions of physical activity and gives figures (percentages) regarding the physical activity of adults and teenagers. Finally (3), there is also a table that describes in percentages the physical training habits of youths in 1993/1994 and 1997/1998.

In the first exercise, the pupils are to use their own words to interpret what the authors of the article meant were the positive effects of the school model. Exercises 2 and 3 ask pupils to draw conclusions from surrounding text in order to account for the significance of two words. Exercises 4–6 are based upon the brief factual text (2) and the adjacent tables (3). The purpose of exercises 4 and 5 is to locate explicitly reported information and the purpose of exercise 6 is to interpret the figures in the various tables.

The aforementioned open exercises were formulated as follows: "24. Which *two* texts (from the entire booklet) do you think have something in common? *Explain* the connections you see." The next exercise is understood by relating to the images in the booklet. Andy Warhol's *Tripple Elvis*: is found in this booklet, for example. "The booklet is called 'Förändringar och funderingar' ['Change(s) and reflection(s)', trans. com.]. Which image do you think best matches the concept 'change'? *Describe* the image and *motivate* the way you think it relates to change."

As described above, various kinds of text are presented in the booklet. In order to assess a pupil's solutions, the answers to the exercises are rated on a scale. There are four levels, 0-3, where a "0" means that the pupil does not fulfil the most basic requirements. Not all the exercises enable a rating on all levels. For example, a relatively simple exercise can only be rated right or wrong, which means a value of 0 or 1. Below are the exercises associated with text from *Populärmusik från Vittula* [Popular Music from Vittula] and *Skärp dig, sa Elin* ['Wise up, said Elin', trans. com].

Populärmusik från Vittula (p. 3–5)

1. Tick true or false for the following statements based upon the entire account!
 - a) The narrator's father liked Elvis Presley very much.
True ☐ False ☐
 - b) Niila made fun of the narrator when he was practicing in the garage.
True ☐ False ☐
 - c) The gramophone at school played the music too fast.
True ☐ False ☐
 - d) The roar of the audience could be heard all the way to the church.
True ☐ False ☐
 - e) The guys grumbled enviously because the girls appreciated the performance.
True ☐ False ☐
2. During which period of time does the story take place? Tick one box.
☐ from the 60s on ☐ during a school term ☐ during three weeks
3. How did the narrator sing?
 - a) when he practiced alone.
 - b) when he practiced with Niila.
4. You can assume that the narrator's sister is going to be angry after the "concert".
What two reasons does she have for being angry?
5. "We hade /.../ blown up each other's fantasies to huge, colourful balloons", it says about the boys in one of the stories. Explain the quote. What are the consequences of the boys' inflated dreams?

It may seem as though a quick reading of the text is enough to arrive at a correct answer to exercises 1 and 2. The answers are expressed in other words and wording in the text. The pupil has to interpret and reflect on the meaning of the text. A pupil who has just skimmed through the text will have problems; more

careful reading is required. In exercise 3 the pupil is to answer in his or her own words. The situation is similar here; the information is embedded in the text and one who has assimilated the content can easily formulate a correct answer. As previously mentioned, the text stated: "But it was a song without sound, just my lips were moving ... I was moving on the inside to the music ...". This was, however, preceded by: "I ... opened my mouth. and began to sing." A more careful reader understands immediately, while the careless reader will, as suggested before, have problems. In their text on the assessment of pupils' answers, the national test group states: "Answers that demand more *detailed explanations, examples, or several aspects of the exercise* (e.g. both explanation and motivation) can be regarded as *more or less complete rather than right or wrong...*" (FUMS etc., 2003.) What is important, quite simply, is thus that the pupil is capable of reflecting and understanding the content, and then answers as completely as possible.

In exercise 4 the pupil is once again asked to retrieve information, which is more embedded in this case. Here the task is to evaluate information that is given and interpret its meaning. The author writes: "and I had borrowed – without permission – my sister's *Jailhouse Rock*" and "the heavy needle bounced around like a woodpecker's beak, pecking at the unarmed record." Later passages let us know that the "borrowed" record has probably not survived. An alert reader can sense the sister's presumed anger, but one who is looking for explicit evidence will not find it.

Finally, in exercise 5 the pupil is asked to interpret a quote, a metaphor. The reader must be skilful in drawing conclusions and understand what is not expressed. The pupil must know how to read between the lines, reflect upon and give a coherent account of the content.

A pupil's ability to find information, interpret and reflect are put to the test. Pupils who answer the questions satisfactorily have thereby fulfilled the goals that are stated in the syllabus for Swedish. Having seen this concrete example of the way in which reading comprehension is measured we have at least a sense of the complexity of the procedure. In order to determine the extent to which a pupil has not only read but also understood a text requires many questions on a variety of levels.

Texts and exercises in PIRLS, PISA and NU are described further along in this report. It can be of value to make a note of the structural disposition of the above-mentioned material.

National test in the fifth school year

The national test of Swedish in the fifth school year is different from the test in the ninth school year. We will discuss here the test that was given in 2006. The collection of texts is less comprehensive and the design of the test is different. There are two longer texts, both fiction and a spread with the headline "smått och gott" ['bits and pieces', trans. com], containing a poem, a quote, a letter to the editor, and some sayings. All of the test exercises are of an open character (which is contrasted in the design of the trend study, see the forthcoming section). We will acquaint ourselves cursorily with two texts and some of their associated exercises.

De tre grottornas berg ['Three cave mountain', trans. com.] by P-O Enquist tells of some children and their stay with their grandfather in Värmland. The

children visit a cave, their grandfather tries to teach them rhetoric, and in the end they run into a bear.

A careful reading is required in order to answer the questions correctly. To do the exercises the pupil must read between the lines and make inferences. None can be solved by simply copying text from the booklet. Let us take a look at question 1: *Who is Mischa?* This may seem easy to answer, but nowhere in the text does it state explicitly that Mischa is a dog. To answer the question the reader must make use of the illustration alongside the text and understand the significance of "And [sitting] farthest away Mischa, who was holding Marcus by the paw". The paw here belongs to Mischa, which appears in the text. We are also going to read questions 8 and 9.

Question 8: *Why are some words written with capital letters?* The capital letters in the middle of certain words are clearly a breach of written language conventions for one who is familiar with the latter. Their purpose is to convey that the text expresses something that is said with emphasis. For one who has not read the text carefully, or is not well acquainted with the most common orthographic rules, it is difficult to interpret the significance of the upper-case letters.

Question 9: *Ia and Marcus are siblings. Marcus is five years old. How old can Ia be? Explain what you think about her age.* Ia's age is not mentioned in the text, but she complains about Marcus, who still doesn't get certain things despite the fact that he is five. Furthermore she speaks to him as though he were an idiot. It is necessary for the reader to make an inference like "of two siblings, one of whom understands more than the other, the former is usually the elder". A correct answer is based largely on reflection and interpretation.

To sum up, the test in the fifth school year can be good practice for future exercises. The tests are similarly constructed, but the one for the ninth school year is longer and has different kinds of questions. Interpretation and reflection are central to both of the tests.

5 Studies – framework, concretion and results

5.1 RL and PIRLS

5.1.1 Framework

PIRLS is an acronym for *Progress in International Reading Literacy Study* and *RL* stands for *Reading Literacy*. The International Association for the Evaluation of Educational Achievement (IEA) has authored these two studies. RL (henceforth referred to as *the trend study*) was first done in 1991 and then again in 2001. The purpose was to measure whether or not a change in reading literacy had taken place. The exercises and framework associated with the trend study in 1991 were the same in 2001. We should bear in mind that the framework for the trend study was not, for obvious reasons, associated with the current syllabus, which came into use in 1994. It is, however, still interesting to study its theoretical basis in light of the syllabus for the purpose of identifying possible similarities and differences.

The first part of the trend study was implemented in 1991. The target group was pupils in their third school year. Rosén and Gustafsson (2004), who authored the Swedish study relate, by quoting Taube, the following definition of reading proficiency which served as the basis for the work with instrument design:

Reading proficiency is the ability to understand and use forms of the written language that are required by the society and/or are of value to the individual.
(Taube, 1995, p. 13 in Rosén and Gustafsson, 2004, p. 16)

The definition emphasises "functional reading ability" as well as "significance with reference to the value of an individual's ability to read ... for his or her own pleasure". (Rosén and Gustafsson, 2004, p. 16) It is fairly comprehensive and general. Nonetheless it is possible to study it in light of the syllabus and discover a similar ambition in the latter text (see section 3). What is perhaps missing from the trend study and emphasized in the syllabus is the question of reflection. On the other hand, it is possible to argue that it is of value for the individual to be able to reflect over the meaning of a text. To reflect over a text can be to understand the text on a deeper level and associate it with one's self.

The trend study was implemented in some 30 countries in 1991 and in 9 countries in 2001. Based upon the definition, exercises were constructed for three kinds of text: *fiction*, *non-fiction*, and *information materials*. The second part of the trend study was implemented at the same time as PIRLS 2001, but is a separate study. The most important differences between the trend study and PIRLS are summarized at the end of this section.

The aim of *PIRLS* (Progress in International Reading Literacy Study) is to measure the reading proficiency of pupils in "the upper of the two adjacent grades with the most nine year olds. In most countries this is the fourth grade". (Hasselrot, 2003, p. 8). That Sweden was to test pupils in the fourth school year was thus clear. Despite this, it was decided to test pupils in the third school year too. This was motivated by the fact that the trend study tested pupils in

their third school year. The idea was that the trend study would be compared to PIRLS, and thus maximum comparability was wanted.

The study is intended to recur every fifth year, and since the PIRLS tests were conducted in 2001 and 2006 the next years to collect data would be 2011 and 2016 etc.

PIRLS measures knowledge in children who, in relationship to their age, should read in order to learn. The theoretical point of departure in PIRLS is *reading literacy*, which is translated into *läskompetens* in Swedish. This is used to refer to:

the ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers, and for enjoyment. (ibidem, p. 25)

Several aspects of the definition deserve attention. Reading literacy means that the pupil has enough knowledge to understand a text not just superficially, but on a deeper level: "In accordance with this definition it is assumed that the reader actively constructs meaning, is aware of effective reading strategies and able to reflect on a text." (ibidem, p. 26)

The first part of the definition deals with the knowledge and skills needed to participate in a society. A person who lives in an organised society must be able to assimilate information in public contexts, receive directions and notices. The second part is about recreational reading (ibidem).

In the theoretical framework for PIRLS emphasis is made on two primary objectives of reading: (1) "Reading to obtain a literary experience" and (2) "reading to acquire and make use of knowledge." (ibidem, p. 28) In other words, reading literacy is to be measured in cases where the purpose is (1) pleasure (2) learning.

In PIRLS, as opposed to the trend study, both multiple-choice questions and exercises requiring that a pupil to produce written text are used. A total of 98 exercises were constructed. The purpose of these was to measure the pupil's ability (a) *to find and explicitly report stated information*, (b) *to draw simple conclusions*, (c) *to interpret and integrate thoughts and information* and (d) *to scrutinize and evaluate various aspects of a text*. Each pupil had two texts with associated exercises.

The intention of PIRLS and its exercises appear to correspond well with the syllabus and the goals it has set for fifth year pupils. The goals of the syllabus are that pupils in their fifth school year will "be able to read with fluency, both aloud and to themselves, and understand events and meaning in books and non-fiction written for young children and young persons, and be able to discuss their experiences from reading, as well as reflect over texts." (Syllabus for Swedish) We can assume that when the syllabus refers to *books* it means fiction, since books in this context are distinguished from *non-fiction*. PIRLS is intended to measure reading literacy, both reading for enjoyment as well as reading to learn. In this respect the main purpose of PIRLS corresponds with an important paragraph in the syllabus. Other parts of this paragraph PIRLS can be found in the four aspects of reading literacy that PIRLS aims to measure. The syllabus says that pupils in the fifth school year should understand the developments of a text and be able to reflect on this. The concept of *reflection* is broad, and when

reading PIRLS in light of the PIRLS framework we find that (b), (c) as well as (d) can be contained in what the syllabus refers to as reflection.

The main differences between PIRLS and the trend study are elucidated in an *unpublished* article (Gustafsson & Rosén, 2006). PIRLS and the trend study are largely in agreement, but generally exhibit differences in focus. The trend study measures reading literacy with three different kinds of texts: *fiction*, *non-fiction*, and *information materials*. The emphasis in PIRLS, on the other hand, is on two text categories: *reading for experience* and *reading to acquire information*. (ibidem, p. 5) Information material does not have the same prominent place in PIRLS (though it is represented). Furthermore there is a notable difference in the construction of the exercises. The trend study consists almost solely of multi-choice questions along with a few exercises where the respondent is to fill in using one or two words. Half of the exercises in PIRLS are the kind where pupils answer using their own words. (ibidem, p. 3) In addition, there is a notable difference in the length of the texts, where the texts in PIRLS tend to be longer (but with more time allotted) and include a dimension of reading literacy that is not tested in the trend study. (ibidem, p. 7) Regarding the latter, pupils are to produce a metatext by formulating sentences that evaluate the style and content of the text. (ibidem, p. 6) It should be pointed out, however, that only 15 percent of the texts in PIRLS measure this dimension.

The purpose of PIRLS corresponds thus well with the aim of the syllabus for Swedish. It can also be said that the aim of the trend study corresponds too, since PIRLS's definition is more of an extension than a total revamp. Some exercises are described below, and one is reproduced in full.

5.1.2 Exercises²

Haren varnar för jordsbävning ['Hare heralds the earthquake', trans.] is a text intended to provide a literary experience. When a fruit falls, an anxious hare mistakes it for an earthquake and warns a total of 10 000 other hares. A lion runs into the hares and asks the anxious one to show him where he discovered the earthquake. There the lion can find out what caused the quake and thus demonstrate it and thereby calm the hare.

After the reading, pupils answer multiple-choice questions and give a written answer based upon their interpretation of the text. The first question has to do with what worried the hare most and there are four answers to choose from. The last of the eight questions in the exercise asks whether or not the pupil thinks that the lion likes the hare.

Cykla Nordälvsleden ['Follow the River Nord Trail', trans.] is a text in the category information material. It's a leaflet from Zippy, a bike-for-hire company that tells how much it costs to hire a bicycle, a stretch of the Nordälv trail and sights along it.

As with the exercises after the text about the hare, pupils are to begin by locating certain information. Pupils are to use the table with prices to draw conclusions about how a family constellation affects the price. Later they are asked to motivate their choice of sights based upon the brochure's description of a stretch of the trail and possible places to stop.

² Exercises in the trend study are not shown here. As previously mentioned, these all consist of multiple-choice questions, the form of which is comparable to examples (1) and (2) in the exercises associated with *Haren varnar för jordsbävning* below.

The exercises serve the aforementioned purpose of measuring various reading skills. To ask, as in the latter exercise, why a certain passage in the text is given requires an ability to draw conclusions. In order to answer the question about the trail head, requires finding explicitly stated information. When one is then asked to motivate where a particular family might stop, it is necessary to reflect on the sights and places that would be suitable for the various family members. This is based largely on previous knowledge.

Haren varnar för jordsbävning

Below are all the exercises that follow the text *Haren varnar för jordsbävning*.

The exercises that are not followed by a parenthesis are of the open type.

1. What was the hare's greatest worry? (A: A lion, B: A crash, C: An earthquake, D: A falling fruit)
2. What made the whole earth shake?
(A: An earthquake, B: An enormous fruit, C: The fleeing hares, D: A falling tree)
3. Things happened quickly after the hare shouted "Earthquake!"
Find and copy two words in the story that show this.
4. Where did the lion want the hare to take him?
5. Why did the lion drop the fruit onto the ground? (A: To make the hare run away, B: To help the hare get the fruit, C: To show the hare what had happened, D: To make the hare laugh)
6. How did the hare feel after the lion dropped the fruit on the ground?
(A: Angry, B: Disappointed, C: Foolish, D: Worried)
7. Write two ways in which the lion tried to make the hare feel better at the end of the story.
8. Do you think the lion liked the hare? What happens in the story that shows this?

Four of the exercises that follow *Haren varnar för jordsbävning* are multiple-choice questions. These can be quite intricate, which the examples in the test indicate. Generally speaking the hare in the story seems to be worried; when the lion confronts him he answers "timidly". This can make it somewhat difficult to accurately determine who or what the hare is most afraid of.

In the exercise the pupil is to determine what exactly has happened. Once again, the one who has carefully assimilated the content can conclude that it is a large piece of fruit that has caused the trembling. The pupil who has read the text superficially gives in perhaps to the hare who believes that there was an earthquake. The answers to both exercises 1 and 2 are explicit, but require a careful reading. Exercise 1 includes a measure of assessment, where the pupil is to use clues to determine what the hare is most afraid of.

A linguistic observation is requested in exercise 3. The pupil must reflect on the content and become clear as to what makes the reader understand the course of events. In exercise 4 the pupil is asked to recount the course of events. The lion says in the text: "... do you think you are brave enough to show me where this terrible catastrophe occurred?" Once again, the careful reader should not encounter any major problems in completing the exercise. In order to satisfactorily complete exercises 5 and 6, the pupil is required to interpret and reflect on the content of the text. There is no request for a specific answer here, but a simple inference is required for the pupil to show that he or she has understood the content. To a greater extent than the two previous exercises, exercises 7 and 8 are questions that request and demand that the pupil reflects upon and interprets the text. The pupil must be aware of what has happened and be able to evaluate it.

The pupil who has read the text carefully will be able to answer all the questions in the exercises. The text is relatively independent and it is not necessary for the pupil to know a lot about a particular subject in order to understand. The exercises measure how the pupil has understood the course of events and been able to reflect on the text. As stated previously, pupils in the fifth school year should understand the course of events and meaning of a text and be able to reflect on it. It is possible to demonstrate proficiency at this with the exercises shown above.

5.1.3 Results from PIRLS and the trend study

Third and fourth year pupils have been tested with PIRLS. Thus results of these groups will be presented separately.

The target group for *PIRLS* is defined as pupils in "the upper of the two adjacent grades with the most nine year olds. In most countries this is the fourth grade." Sweden participated, as mentioned above, in both third and fourth grade. Swedish pupils in the 4th school year perform very well, in fact better than those from all the other participating countries. They not only perform best, but the distribution of their results was the lowest. Swedish pupils had an average of 561 points, to be compared with the international average of 500 points. Accordingly, pupils in their fourth school year perform very well and significantly better than the international average.

Swedish pupils in their third school year, who are number 19 internationally, also had results that were significantly better than average. Remember that they are being compared with pupils who have attended school for one more year. As an explanation for the somewhat poorer results, seen from an international perspective, *PIRLS* talks about a negative skewness, a sort of "tail" of low values (Hasselrot, p. 10). The low performance can be understood with an awareness of the situation in the fourth school year, when the "tail" is gone because many of the poorest performers have improved their results. This may be a consequence of improved reading comprehension, i.e. a better ability to decode words.

When we then take a look at the trend study the situation is different. The results indicate that reading literacy was better in 1991. Of the nine countries that participated in the trend study only two countries, Sweden and the U.S.A., showed a decline. Sweden placed number six of nine, ahead of Slovenia, Singapore and Hungary. If Sweden had performed as well as in 1991, it would have shared the lead with Italy and Iceland.

In *PIRLS*, *boys and girls* have been separately accounted for and clear differences can be seen. Sweden is one of the countries with by far the greatest difference, as it places over the international average with regard to differences that are to the girls' advantage. The parents' education level is clearly correlated to the pupils' achievement in reading comprehension, in Sweden as well as in other countries (Hasselroth, 2003).

In summary, it can be said that pupils in the fourth school year do well, in fact best of all the participating countries. Pupils in their third school year do not do as well. They place 19, but still significantly above the average. Furthermore, the trend study shows that there has been a decline and that reading literacy was better in 1991. In addition, the observation is made, as in *PISA*, that girls read better than boys.

5.1.4 Framework and information relative to the syllabus and the national test

PIRLS/RL

With the aforementioned differences between the trend study and PIRLS as a point of departure, thoughts and conclusions about the framework of the studies in relationship to the syllabus are summarized below.

Initially it is interesting to study how the definition of reading literacy (*reading ability* in the case of the trend study) corresponds with the syllabus. PIRLS' definition is somewhat more comprehensive than that of the trend study, but both nevertheless strive to achieve the same goals. One who has a reading ability/literacy can use the written language to become an integral part of society and for the sake of his or her own pleasure. In PIRLS' there is the additional emphasis on being able to extract meaning from various types of text. Consequently, without a biased reading of the definitions, it is possible to assert that PIRLS' further development is actually an explanation of what is already inferred in the trend study. Society imposes demands on one's ability to assimilate various kinds of texts. Pupils in the fifth school year should, according to the goals to attain "be able to read with fluency, both aloud and to themselves, and understand events and meaning in books and non-fiction". This quote is equivalent to the substance of the definitions. In the same goals to attain it is stated that pupils should "be able to discuss their experiences from reading, as well as reflect over texts". Reflection and discussion are not mentioned in either the trend study or in PIRLS. The definitions are, however, so general that reflection and discussion could very well fall under the wording "valued by the individual", to the extent that *discussion* and *reflection over texts* are intended to mean something which the individual can benefit from.

A prerequisite for the relevance of these studies in Sweden is that the definitions are in agreement with the ambition of the Swedish syllabus. If we take a look at the actual content we see that while the above expressed definitions are fairly close to one another, they are in practice a bit different from one another. This is expressed primarily in the various dimensions of reading that are incorporated in the texts and exercises. The trend study focuses on three kinds of texts. In PIRLS there are two main categories of texts, each of which has two sub-categories. The difference is that information material is represented to a greater extent in the trend study and that a kind of text, or dimension, in PIRLS – the one which gives rise to the questions about the style and tone of a text – is not represented in the former. In PIRLS the pupil is asked to comment on style and tone as well as to reflect over and discuss (however limited in scope) the text. This is a goal specified in the Swedish syllabus.

When we take a look at the national test instead, we find that the way questions are posed in PIRLS corresponds more closely to that of the national test. The trend study consists almost solely of multiple-choice questions and only a few exercises where the pupil is asked to fill in using his or her own words. In the latter case, one or two words are requested. The pupil is forced in PIRLS to reflect over the entire text (exercise 8, p. 22) and the exercises are generally like those of the national test. In no way can we write off the method that makes use of multiple-choice questions, but we do note that PIRLS poses questions in a way that more closely resembles the national test than the trend study (see the section on the national text for students in the fifth school year).

That the trend study notes a decline in reading literacy is alarming and the question of the significance of the texts and exercises in the trend study versus PIRLS is critical. Gustafsson and Rosén (2006) are of the opinion that an important factor to consider in an analysis of differences between PIRLS and the trend study has to do with reading speed. They assert that the reader must be considerably quicker in order to satisfactorily answer the questions in the trend study. We must bear in mind that pupils in their third and fourth school year have yet to encounter the national test. But the question remains, whether or not the way of dealing with a text is also an important factor if we are to understand the results of the studies. To the extent that there is a demand that the point of departure and way of handling a text in the Swedish syllabus and the national test correspond as closely as possible, it appears that PIRLS (however marginally) corresponds more closely than the trend study.

5.2 PISA 2000

5.2.1 Framework

OECD's initiative PISA (The Programme for International Student Assessment) measures 15 year olds' reading literacy, mathematics literacy and science literacy. The purpose of PISA is to study how different educational systems help to equip 15 year olds for adult life. The first PISA study was done in 2000 and the other in 2003. This report refers only to the study done in 2000.

PISA recurs each third year with an emphasis on one of these three subjects. In 2000 the focus was on reading literacy, and thus science and mathematics were afforded less space. Mathematics was emphasised in the 2003 measurements and science in 2006. The reason why the two subjects which are not main subjects are included in every study is that this is necessary to implement trend studies in all three subjects.

The concept of *literacy* is used in the theoretical framework around PISA, which is understood as 'the ability to understand, use and reflect over texts in order to achieve personal objectives, develop one's own knowledge and potential, and to participate in society'. In PISA *läsförmåga* [literally 'reading ability', trans.] is proposed as the Swedish equivalent to the English term. (Taube, Petersson & Molander, 2001, p. 22)

PISAS's exercises are constructed to measure three dimensions of reading comprehension: (1) *processes*, (2) *knowledge and understanding* and (3) *context*. In simple terms these dimensions can be understood as: (1) strategy for reading (on the assumption that different texts require different strategies), (2) what one obtains from the reading, and (3) the context in which the texts are read.

PISA makes use of three sub-scales to present the results: *information retrieval*, *interpretation* and *reflection*.

Table 1 Subscales in PISA

Information retrieval	Interpretation	Reflection
[...] the ability to localise information in a text	[...] the ability to understand and draw conclusions from one or more parts of a text	[...] the ability to relate a text to one's own experiences, knowledge and views.

(ibidem, p. 29)

The student is given points for each correct answer. The number of points vary depending upon the level of difficulty. A simpler exercise gives less points etc. The results are incorporated in one of five levels, where the first is the lowest, based upon the number of points the pupil obtains.

Table 2 Score Points at each proficiency level in PISA.

	Level 1	Level 2	Level 3	Level 4	Level 5
Points	335–407**	408–480	481–552	553–625	> 625

(ibidem, p. 31)

PISA's framework is obviously complex and places major demands on the exercises. Once again we should read from the syllabus on Swedish, which states that pupils should: "be able to read literature appropriate to their age from Sweden, the Nordic area, and other countries, and also read non-fiction and newspaper articles on general subjects, as well as be able to reproduce the content coherently and also reflect over this." The syllabus states that different kinds of texts should be understood. A measurement which is made in relationship to the syllabus should therefore contain different kinds of texts, which of course require different strategies. This is also evident in the various dimensions of PISA (primarily (1)). To reproduce the contents and reflect, as it is stated in the syllabus, is reflected in the latter dimensions. If we compare the exercises in PISA with the exercises in the national test of Swedish, the picture becomes clearer as to how not only the framework but the exercises (i.e. that which the pupil actually confronts) are equivalent to the "usual" way of tackling the demands of the syllabus. The exercises are described briefly below.

5.2.2 PISA – questions

PISA 2000 consists of a total of 138 reading questions, which are based on fiction and non-fiction as well as discursive texts. Below are some of all the questions presented, in brief, in PISA. With suitable passages from the syllabus for Swedish in mind, it may be worthwhile to obtain a sense of the content of the texts and how their associated questions are constructed in PISA. After the summary, the text *Graffiti* with associated questions is presented in full.

Graffiti consists of two texts. Both are discursive and public, since they were published on the Internet. The writers *Helga* and *Sophie* look at both the prevalence of graffiti and the extent to which it is damaging. Four exercises are connected to the texts. The pupil is asked to choose among four statements to indicate what the text is about. The pupil is to use his or her own wording in the remaining questions. The pupil is to interpret the writers' intended meaning in certain passages, and reflect over the content by evaluating the argumentative discourse, tone and style.

Distribution of labour is presented in PISA as a representation of a schematic text in educational contexts. A tree diagram describes how an unnamed country's working population classifies itself into different categories, such as *in the labour force*, and *employed* or *unemployed*. The pupil is to use the diagram to answer five questions with different demands, and in order to obtain the maximum number of points the student must understand facts that are not explicitly expressed.

The fictional piece *Gåvan* ['The Gift', trans.] is about a woman whose house is no longer on solid ground due to a flood. The woman who is locked inside the house is visited by a puma that sits down, in a noisy and threatening manner, on the veranda. The woman doesn't manage to shoot the animal. After a while she decides to feed the puma with ham. The next day the puma is gone from the veranda.

Seven questions follow *Gåvan*. The purpose of these is to call upon the pupil to provide answers that require information retrieval, interpretation and reflection. On the basis of what has happened, the pupil is to use assertions to describe the woman's behaviour and find support in the text for opposite views. Statements claim that (1) the woman is "heartless and cruel" and (2) that the woman is "an enormously empathetic person".

Polisens vetenskapliga vapen ['The scientific weapons of the Police', trans.] is a non-fiction example of a magazine article. The author accounts for the concept of *DNA* and how it can be used in the police's preliminary investigations. A somewhat longer passage is followed by three tables of facts that provide details on parts of the article. Four exercises follow the text. The pupil is to answer questions about details in the text and interpret the underlying meaning of the text. The latter is done with the help of the multiple-choice question: "What is the author's main objective?"

The exercises meet the objective of PISA to the extent that they require that the pupil is able to make use of techniques to find information, interpret and then reflect over it, in order to obtain points. Given the fact that PISA contains different kinds of texts, it largely resembles the national test of Swedish. It has a broad scope and also enables pupils to convey their knowledge on various levels.

An example of how the questions are designed follows, using the aforementioned texts about graffiti. The choice of this text is motivated with reference to the opening paragraph of Lpo94 [Swedish Curriculum for the Compulsory School System, trans.]. An active citizen must, as has been mentioned so many time before, be able to assimilate and critically evaluate arguments.

Graffiti

I'm simmering with anger as the school wall is cleaned and repainted for the fourth time to get rid of the graffiti. Creativity is admirable but people should find ways to express themselves that do not inflict extra costs upon society.

Why do you spoil the reputation of young people by painting graffiti where it's forbidden? Professional artists do not hang their paintings in the streets, do they? Instead they seek funding and gain fame through legal exhibitions.

In my opinion, buildings, fences and park benches are works of art in themselves. It's really pathetic to spoil this architecture with graffiti and what's more, the method destroys the ozone layer. Really, I can't understand why these criminal artists bother as their "artistic work" are just removed from sight over and over again.

Helga

There is no accounting for taste. Society is full of communication and advertising Company logos, store names. Large intrusive posters on the streets. Are they acceptable? Yes, mostly. Is graffiti acceptable? Some people say yes, some no.

Who pays the price for graffiti? Who is ultimately paying the price for advertisements. Correct. The consumer.

Have the people who put up billboards asked your permission? No. Should graffiti painters do so then? Isn't it all just a question of communication – your own name, the names of gangs and large works of art in the street?

Think about the striped and chequered clothes that appeared in the stores a few years ago. And ski wear.

The patterns and colours were stolen directly from the flowery concrete walls. It's quite amusing that these patterns and colours are accepted and admired, but that graffiti in the same style is considered dreadful.

The times are hard for art.

Sophie

The two letters above come from the internet and are about Graffiti is illegal painting and writing on walls and elsewhere. Refer to the letters to answer the questions below.

Question 11: GRAFFITI

The purpose of each of these letters is to

- A Explain what graffiti is
- B Present an opinion about graffiti.
- C Demonstrate the popularity of graffiti.
- D Tell people how much is spent removing graffiti.

Question 12: GRAFFITI

Why does Sophie refer to advertising?

Question 13: GRAFFITI

Which of the two writers do you agree with? Explain your answer by using your own words to refer to what was in one or both of the letters.

Question 14: GRAFFITI

We can talk about what a letter says (its content).

We can talk about the way the letter is written (its style).

Regardless of which letter you agree with, in your opinion, which do you think is the better letter?

Explain your answer by referring to the way one or both letters are written.

Question 11: This question offers a possibility to reveal whether or not the pupil can search for superficial knowledge. It should be obvious that the correct answer is (B), but the distracters manage to get the pupil to decide what the main purpose of the text is.

Question 12: The pupil is to answer this question in his or her own words. In this case there are more or less correct ways to handle the exercise. The pupil cannot, of course, refer to some nonsense, but it is reasonable to expect correct answers to have different wording.

Question 13: The task here is to reflect over the meaning of the text. The pupil must prove that she or he can relate to the texts, and assess and evaluate the arguments. The pupil must have the proficiency to transform the text into her or his own thoughts and then into a factual written answer.

Question 14: The question requests that the pupil analyse the style of the texts. Once again the task is to reflect over the content, this time about how it is presented and what effect this has on the content. The pupil must draw a conclusion regarding words that are of fundamental significance and phrases that manage to convince or not.

In order to answer the above questions in a reasonably correct way it is important that the pupil masters techniques for retrieving as well as reporting information. The construction resembles that of the national test of Swedish and it would suffice to say that the exercises in PISA bear up in comparison with the goals to attain stated in the syllabus. This is similarly true for several of the exercises associated with each text, where pupils are given the opportunity to demonstrate their proficiency on several levels. We can also agree that the exercises in PISA mirror the framework of the investigation. The aforementioned dimensions are found in the exercises. The various reading scales thus also fit in. Information retrieval, interpretation and reflection are the focuses of these exercises.

5.2.3 PISA – results

PISA measures three aspects of reading comprehension: *information retrieval*, *interpretation* and *reflection*. In the study, proficiency is presented on various sub-scales, but here only the composite scale is taken into account. This is because the whole, at least as it viewed from the perspective of the syllabus, is what is important.

As previously mentioned, a special rating scale has been developed for PISA. The examinee is rated on a level depending upon the number of points that are obtained.

The mean value for Swedish pupils on the composite reading literacy scale is 516 PISA points. Within OECD Sweden places higher than the mean for all participating countries (which is 499 points). Swedish pupils perform significantly worse than students in only three other countries and on the same level as in nine countries. Sixty-seven (67) percent of the Swedish students obtained more than 481 PISA points, a score which corresponds to the grade Pass. How the number of points obtained correlates with compulsory school grades has been examined by analysing how pupils on each grading scale did in PISA. This means that pupils with the grade Pass, for example, had an average PISA score of 481 points. The mean for pupils who had Fail as an average grade was slightly greater than 400 points. Approximately 12 percent of the Swedish pupils had scores of 400 points or less.

It is interesting also to take a look at the distribution: "In summation, Sweden has comparatively little total variation. [...] The very weakest and the very best Swedish readers performed well, but far from best compared with corresponding groups in many other countries." (ibidem, p. 44)

In PISA there has been an attempt to analyse the results by studying how proficiency varies among and within schools and among students. This survey shows that the greatest differences are within schools and among students. We must remember that the total variation is not comparatively large, but it is still of interest to ascertain the variables that covary with the results. With the help of data about parents' profession, coupled with international data on the relationship between profession and income, a socioeconomic index (sei) has been calculated for the pupils. Subsequently, 2,000 randomly selected results of reading comprehension and all reading and sei averages from the participating schools are placed on a scale where the connection between sei and the PISA outcomes are shown. Here it can be seen that pupils' results in PISA on average covary with placement on the sei scale. Simply speaking, there is a connection between the pupils' socioeconomic background and their performance in school.

In PISA, girls outperform boys; girls are overrepresented on the three highest PISA levels and boys dominate on the two lowest. Of the boys, 40.1 percent are on level 2 or below. Of these, 4.6 percent are below level 1. Of the girls, 24.9 are on level 2 or lower. With regard to the top-performing pupils, 15.1 percent of all girls and 7.4 percent of all boys are in this group.

The results are also classified in order to analyse the extent to which migration patterns affect the results. It is shown that pupils with Swedish as their mother tongue do best. The group that has both Swedish as well as another native language place next. The group with a mother tongue other than Swedish has the poorest performance.

In summary, three of ten pupils fail to achieve what corresponds to a Pass in PISA. There is a correlation between socioeconomic background, gender, and mother tongue with regard to performance.

5.2.4 Framework and exercises relative to the syllabus and national test

The definition of reading literacy in PISA, which is basis of the entire survey, is as previously stated: "the ability to understand, use and reflect on written texts in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society". The definition is without a doubt comparable to that in PIRLS, where an active citizen of a society, as well as personal goals are referred to. PIRLS talks about the pleasure of reading, while PISA focuses on the development of one's potential. As previously mentioned, the definitions are of a general nature and not worded identically. This is to say that the definitions correlate since it is possible to associate the wording in one definition with that of another. We have already established that the definitions in PIRLS and the trend study are in agreement with the goals to attain in the syllabus. In the ninth school year these goals are somewhat different, however, which is why we should study PISA's point of departure in light of the syllabus, which says:

Pupils should be able to read literature appropriate to their age from Sweden, the Nordic area, and other countries, and also read non-fiction and newspaper articles on general subjects, as well as be able to reproduce the contents coherently and also reflect over this,

Does PISA's definition correspond to the goals to be attained? The goals of the syllabus are more precise than those in the PISA definition, at least in the sense that the syllabus indicates the kinds of texts which the pupil is expected to understand. It is, however, reasonable to assume that what is referred to as texts in PISA in reality refers to different kinds of texts and are at least suitable for the age group. In the Swedish syllabus it says that pupils should be able to read non-fiction and other texts and be able to use them for retrieving information and reflection. This corresponds to "understanding and using" in PISA. PISA also emphasises active participation in society, which is comprehensive and can be applied to the syllabus.

The actual wording of PISA's definition should be viewed in light of the syllabus as well as the national test. We begin with the former, as indicated above, noting that PISA makes use of various dimensions in the construction of the exercises, and that the exercises and their answers are incorporated into three scales: *information retrieval*, *interpretation* and *reflection*. PISA's texts and exercises require the pupil to retrieve information, interpret and reflect. This is also emphasised in the Swedish syllabus. Non-fiction and newspaper articles often require information retrieval, and fiction requires interpretation and reflection. All kinds of texts can, however, require the use of all the procedures. The kinds of texts and exercises applied in PISA correspond to the demands imposed by the syllabus. Once again we return to exercise 14, which was reported above.

The text about graffiti was of a discursive nature and gave the pupil exercises that definitely measure what PISA states that it measures. Exercise 14 asks pupils to reflect over their opinions and to constructively assess and evaluate a text:

Question 14: GRAFFITI

We can talk about what a letter says (its content).

We can talk about the way the letter is written (its style).

Regardless of which letter you agree with, in your opinion, which do you think is the better letter?

Explain your answer by referring to the way one or both letters are written

In previous questions the pupil has been asked to retrieve information and interpret a text. In comparing the way exercises are constructed in PISA, like the earlier PIRLS, with that of the national test, it must be said that they correspond. The way of dealing with the texts is central since the theoretical bases are so close to one another. In a comparison of sections on PISA and the national test above, the similarities in the construction of the exercises and the theoretical bases appear greater than the differences. Of course the national test is more comprehensive and takes place over a longer period of time, but the underlying thought and the actual execution have similarities. PISA appears, from the perspective of the framework and the questions in the test, to be an adequate instrument for measuring the goals to be attained according to the national steering documents.

5.3 National evaluation

5.3.1 Theoretical basis

The report *National Evaluation of the Compulsory School in 2003 (NU-03)* describes the situation in a number of subjects in the Swedish compulsory school. The purpose of the evaluation was "to provide a national picture of goal attainment in the compulsory school for Swedish over a ten year period. The goal attainment that is studied concerns the overall goals of the curriculum as well as the subject specific goals in the syllabus". (Elmeroth, 2005, p. 22)

In NU-03 there is no theoretical framework equivalent to those in PIRLS and PISA. On the other hand, the results are based largely on reports about the results of the subject-specific tests. The tests reported in NU-03 are described below.

One of the goals of NU-03 was thus to note changes over a ten year period. Since the subject-specific tests in their current form were first used in 1998, two texts were available from previous evaluations. In the study as well as here, they are referred to as *NU tests*. *Första skoldagen* ['The first day of school', trans.] was published in 1992 and *Våra arbeten* [Our work, trans. com] in 1995 (as a part of the compulsory school evaluation, *Utvärdering av grundskolan*). The texts are adapted to pupils in their fifth school year, but were also given to pupils in their ninth school year. One reason for this was to assess progress in the intervening school years. Unlike RL 1991/2001, NU-03 reports problems with the texts. One is, of course, that the texts are too easy for ninth grade pupils, but this was probably also true earlier in 1992 and 1995. Another asserts that the way of dealing with the texts in the NU tests is old-fashioned compared to today's. It is, however, probably not so completely different that the pupils had trouble understanding the questions (especially considering the low level). What should be attached with extra importance is the fact that many of the respondents seem uninterested in participating. The authors of the report imply also that poor motivation may have affected the outcome: "In correcting the tests it is sometimes evident that the pupils have not exerted themselves to a maximum." (ibi-

dem, p. 29) Elmeroth writes about the significance of motivation for school results. She refers to Giota who has shown that motivation is an important factor and can report that motivation to study Swedish and attitudes toward what is interesting in Swedish indicates a decline from 1992. Hypothetically, had there been greater interest in Swedish results, then results might have been better.

The design of the NU tests

We will acquaint ourselves with the two NU texts and their associated questions. By way of conclusion, half of the exercises in the anthology *Våra arbeten*. ['Our work', trans.], are given as examples.

Första skoldagen takes place in 1865 and tells the story of Robert's first day of school. His family are tenant farmers who live in poverty. The same is true of many of the other pupils. The school is run by the church cantor Olof Andersson under the supervision of Bengt Bengtsson, who owns the school building. The text describes hardships in the school and how the pupils and adults interacted in the school setting of that time. The text contains certain words which may not be easily understood by the pupils: *fickrova* ['pocket watch', trans.], *smilfink* ['smarmer, fawner, smiler, backscratcher, oiler', trans.], *rote* ['district, rota', trans.], *ambulera* ['to ambulate', trans.] and *visthusbod* ['pantry, larder', trans.] are some examples. There is perhaps a disparity between some of the wording and texts read by young people today: "And he was fascinated by the way the congregation followed the slightest hint from the sanctuary."

There are 13 questions, seven of which are open-ended, following the text *Första skoldagen*. The pupil is asked to place five sentences in chronological order. Two of these five sentences are given. Another exercise involves associating a given adjective with the name of the person in the text for whom the adjective is most descriptive. The name of the aforementioned cantor is to be written after the word *inställsam* ['ingratiating, obsequious', trans.]. A third exercise calls upon the pupil to find what is written between the lines of the text. It says in the exercise: "It is implied in the text that Robert likes Carolina. Quote words from the text that substantiate this." The questions are primarily focused on what is happening and less on an interpretation of the text. And any pupil who has read the text should easily be able to answer correctly.

In *Våra arbeten* four people with different professions tell briefly about their work and the advantages and disadvantages of their job. The questions that follow are associated with the texts by two of these people: Ulla has written about her life as a cashier and Ylva has written about her work as a lawyer.

Våra arbeten questions

Questions about Ulla's work

- 1 Sometimes customers squabble with Ulla. Why? Tick one alternative
(A: The goods are old, B: Ulla is slow, C: The goods are expensive, D: Ulla is unfriendly)
- 2 Ulla learned two things when she began her job. Which? Tick one alternative
(A: To smile at the customers, B: To use the price codes,
C: To handle the goods, D: To put the goods on the belt)
- 3 How could the customers, according to the text, help Ulla? Tick one alternative
(A: By having even change, B: By shopping in the morning,
C: By placing the goods correctly on the belt, D: By using the shopping carts)

- 4 Sometimes Ulla makes mistakes. How so? Tick one alternative.
(A: She doesn't weigh the goods right, B: She adds wrong,
C: She come too late, D: She enters the wrong price)
- 5 What is it that makes Ulla like her job? Tick one alternative.
(A: Nice colleagues, B: Nice customers, C: Good salary, D: Good work hours)

All of the questions are limited to choices. There are some open questions in conjunction with *Våra arbeten*, but these are in the minority. Question 1 calls upon the pupil to find a passage in the text that explains why customers complain. In the text there is no mention of the age of the goods, Ulla's speed or possible unfriendliness. The answer is instead explicit in the sentence: "when a customer holds up the queue by complaining to me about the total being too high." This sentence comes after a text that comments the high prices on food. In question 2 the pupil is asked to read the sentence: "I learned two main things then: that we usually use codes when we enter the price of goods, and that it's important to be friendly." Ticking alternative A can possibly require some interpretation, since *vänliga* ['friendly, kind', trans.] has different connotations for different individuals. The context together with the text should, however, provide the answer. Question 3 is perhaps more intricate in that the author of the text indicates that customers should compare the prices of goods, since the writer means that it is sometimes difficult to note the correct price. However, considerably more is said about the placement of the goods on the conveyor belt, and this discussion ends by saying: "I wish that more people were considerate in this way." The shopping hour and the size of the bills and coins are not mentioned at all. And strictly speaking, the writer of the text is not helped short-term by a customer who pays with exactly the right (not too much) change. In question 4 the technique of using two distracters, which are in no way represented in the text, is repeated. The correct answer is D which can be ticked by one who has read the sentence: "Sometimes of course I enter the wrong price." Question 5 is understood by reading the last sentence in the text: "But I like my job a lot, probably mostly because of all the nice people who work there." Friendly customers have been mentioned previously in the text, but the concise reason for why the writer is happy with her job is actually presented.

All of the questions that follow the text about Ulla are considered fairly easy to solve. The main reason being that the answer can be found explicitly in the text. The fact that there are distracters that refer to things that are not mentioned at all, neither explicitly nor "between the lines", helps.

Questions about Ylva's work

- 1 What does Ylva think is her most important work task?
(A: To judge in court, B: To convict a criminal,
C: To help people to agree, D: To help people with their tax returns)
- 2 What book do you think is most important for a lawyer?
(A: Sveriges Rikes lag ["book of Swedish law", trans. com],
B: Deklarationshandboken ["Guide to filing tax returns", trans. com],
C: Domarboken [Biblical "Judges", trans. com],
D: Polisunderrättelser ["Swedish police intelligence", trans. com])
- 3 A lawyer must keep up with a lot of laws. There is another difficulty too. Which?
(A: That there are so many criminals, B: That laws often change,
C: To start one's own company, D: That people quarrel)

- 4 When suspects are brought before the court, lawyers have three tasks. Which?
5. What does Ylva work with most?
 (A: To help people divide up an inheritance, B: To help people do their tax returns,
 C: To help suspects before the court , D: To help people to start their own company)

The questions associated with the lawyer Ylva's text are like the ones above. One question (4) is, however, open-ended. It doesn't ask the pupil to reflect more closely on the text, but to find explicit information in the text. The other exercises are similar to those that followed the text about the cashier. In exercise 2, for example, there are three distracters (*B, C, D*) none of which are mentioned in the text. Exercise 1 is somewhat more intricate since the answer is not explicit in the text. Finally, Exercise 3 also offers distracters that are not represented in the text.

There is a certain difference in the way the texts *Första skoldagen* and *Våra arbeten* are handled in the exercises. The exercises for the first text require more that the pupil can, for example, interpret and make inferences. For students in the ninth school year this is undoubtedly no greater problem. In comparing questions in PIRLS, PISA and national tests of Swedish, we note that the NU tests differ from the others. The three others are directed more toward exercises that involve interpretation, (relatively advanced) inference, and reflection. In order to understand the NU exercises it is important to bear in mind that they are designed to work in the fifth as well as the ninth school year. It is also stated in the NU report that the questions are not "representative of ... current ways of preparing texts". (p. 26) The report also gives detailed comments about the exercises' low degree of difficulty: "Both texts are therefore very easy for many students in the last class." Their use is justified by reference to the main purpose of NU-03 (i.e. to note differences over time). In addition, the quality of the tests was considered capable of providing "the basis for testing the reading ability of pupils who have reading difficulties." (p. 26)

5.3.3 Results NU-03

NU-03 is, as previously said, comprehensive. In the present report, only the measurements taken on *reading comprehension* are discussed. The primary purpose of the national evaluation was to measure results over time.

The results of the NU tests convey a picture of reading comprehension over time. The first NU test, *Första skoldagen*, was given first in 1992 and the second, *Våra arbeten*, in 1995. The differences over time are striking. In 2003 pupils perform much worse than pupils did in 1992 and 1995. It is pointed out in NU-03 that the texts that are presented were not representative of current demands for difficulty in the ninth school year. In the 1992 survey the results were therefore:

extremely and negatively unevenly distributed , "ceiling effects" appeared. In this year's evaluation the distribution is more even, which means that a considerably larger number of pupils than in the previous evaluation had poor results. [...] The distribution in 2003 is large, which confirms that the reading comprehension of all the more pupils has become worse.

Similar results are found in a comparison between 1995 and 2003. Even here the results are, as a consequence of the low demands of the test, negatively and

unevenly distributed. And even here the results indicate a decline in reading comprehension over time. All questions were answered correctly by 20 percent of the pupils in 1995; this figure dropped to 13 percent in 2003.

Girls performed better than boys in the NU tests, as well as in all other aforementioned tests. The major decline for boys is in the comparison of results between 1995 and 2003.

5.3.4 Framework and exercises relative to the syllabus and national test

As previously stated, no theoretical framework equivalent to that of PIRLS and PISA, is accounted for in NU-03. We can, however, assume that NU-03 intended to measure what was stated in the syllabus for Swedish. We must bear in mind that *Första skoldagen*, like the texts used in the trend study, was first used while the Curriculum for the Compulsory School (Lgr 80) was still applicable. The latter differs from Lpo 94 with regard, for example, to the formulation of the goals to be attained and the degree of concretion regarding what is to be included in the teaching of Swedish.³ Thus in the section which follows, information will be viewed only in light of the syllabus and the national test. Since the NU test is intended primarily for pupils in their fifth school year, information should be viewed in relationship to the syllabus' goals to attain for that year. In the goals to attain for the fifth school year (unlike those for the ninth school year) nothing is said about the age for which the literature read by the pupil should be adapted. But given the goals for the ninth school year, even the fifth school year should deal with literature suitable for the age group. Pupils shall, as previously mentioned, be able to read fluently to themselves as well as aloud and be able to report a course of events and reflect over texts. *Första skoldagen* was considered in 2003 to be of such poor quality that it could not be used in the national test (Elmeroth, 2005). The archaic use of language in the text, for example, is notable. The word *fickrova* ['pocket watch', trans.], for which one of the exercises requests a definition, is neither in the Swedish Academy's glossary of Swedish, nor in the glossary of the Swedish national encyclopaedia. Nor is it represented among the millions of words in the running text of the Swedish daily press in the 20th century (Språkdata, GU).

In principle, the exercises associated with *Första skoldagen* ask only to find information that is explicit in the text. The distracters in the multiple-choice questions are often only seemingly distracting, which is illustrated by the ceiling effects in the results. The more advanced exercises to a large extent call upon the pupil to interpret a text, or explain the significance of a word with the help of surrounding text. In order to better correspond with the syllabus, it would have been valuable to have had a number of exercises that required reflection. The literary quality of the texts has not been discussed at all in the present report, though it is noteworthy that *Första skoldagen* was used despite the fact that the constructors of the national test thought the quality of text was so low as to render it unsuitable. It would, of course, have been difficult to note differences over time had the text been replaced, so the question remains what would have been best: to use a poor text to measure, or refrain from measuring?

³ One explanation for the design of the NU test is that it was produced while Lgr 80 was still being used. The NU test should be designed in relationship to the applicable syllabus, since the national test and other tests are constructed to fulfil the goals of the latter.

The text group *Våra arbeten* is followed by exercises that adhere to the principles stated above. The open questions are in a clear minority and these ask basically for information that is expressed explicitly in the texts. Even more so than in *Första skoldagen*, the distracters do not really succeed in distracting. The goals to attain according to the syllabus are dealt with to a certain extent, but discussion and reflection over the texts is missing.

The texts and the way of constructing the texts in the NU tests does not measure up to today's. This is also commented on in the report (ibidem, p. 26). There are two studies that indicate that reading literacy has declined in Sweden. The way of dealing with a text in the trend study differs considerably from that of PIRLS and the NU tests. Once again, the issue of test design *can* be a factor that affects the results.

6 The Agency's assessment

Of five different kinds of tests (including the national test) two stand out. Both of the trend surveys, the trend study and the NU test, are different from the others with regard to design. They are also the ones that supply the most dismal results. Both show that reading literacy has declined over time. And yet the NU test reports, understandably, good results with regard to the number who perform well. But there is no doubt about the fact that literacy has declined. The authors of the two reports explain the results in part differently. Elmeroth (2005) talks about the lack of motivation among the pupils. She points out that today's method of measuring reading is not represented in the NU test. By accounting for exercises in this study we have also been able to make a note of differences between the NU tests and national tests.

The differences between the trend study and the national test of Swedish are similar to those which were reported for the NU tests. A certain way of measuring is dominant in the NU test and the trend study, but not in the national test. Furthermore, Gustafsson and Rosén (2006) are of the opinion that one explanation for the poorer results in the trend study has to do with the duration of the test, i.e. that the pupils in the trend study, as opposed to pupils in PIRLS, took the test in less time.

It is certainly valuable to be aware of the time factor as a possible explanation for differences in the results between the trend study and PIRLS. On the other hand, pupils undoubtedly had the same time frame in 1991 as in 2001. Time can hardly explain the decline in literacy. Likewise, the arguments that were brought out in the NU report – that the texts, due to their low level of difficulty and archaic language were poorly motivating – are also understood with an awareness that the texts were *probably* just as poorly motivating in 1992 and 1995. Searches in the language database at Göteborg University indicate that words that were experienced as archaic in 2001 were not represented in everyday speech in 1992 either.

It may well be true that the design of the NU test does not correspond to the current method. We must, however, consider the degree of difficulty. The test is so easy that the design itself should not be used to explain the poor results. In the case of the trend study, however, the design offers a fundamental explanation. The question is whether or not the school system has so totally transformed the way of working with text, that (poor) pupils cannot relate at all to the trend study.

There are fewer question regarding PIRLS and PISA. Both of these assessment surveys present well-poised definitions and measure reading comprehension in a way that is comparable to that of the national test. The definitions of literacy which PIRLS and PISA present were undoubtedly written in the beginning to be applicable in many countries. It is of course positive that the international organisations and researchers who have contributed to these have views that converge with Sweden's. We should, however, reflect on how the definitions have managed to converge. It is easy to read Sweden's more explicit definition, for example, into the general definitions in PIRLS and PISA. If Sweden's definition of reading comprehension, i.e. the points covered in the syllabus, had

specified the exact content of the education and the goals then this convergence would not have been as easy to achieve. Initially it is questions of definition and points of departure that make participation in international studies at all of interest for Sweden. The relatively open definition in the Swedish syllabus, i.e. of goals to attain, is applicable to other definitions and vice versa. On the other hand, too narrow a definition might ultimately have limited any possibility for convergence. If the syllabus had described the content of goals and education on a detailed level (e.g. which authors the pupil should read etc.) there is a risk that it had not converged with any other definition. With the general definitions, however, there is a risk that the definitions converge, while the instruments (cp. the trend study and PIRLS) do not.

Thus we have seen assessment surveys, which all have the same goals, but different instruments. It can prove to be somewhat risky to rely completely on definitions and goals, as said previously. The trend study and PIRLS present similar definitions and clearly similar objectives, but are designed differently. There is every reason to scrutinize the results of different assessment surveys by taking a stand on the design of the test. Even if the ambition of two assessment surveys is the same, the instruments can differ and thus yield different results from the same population. It would have been devastating to have accepted without question results from the NU test and thus risked believing that it reported a true and complete picture of the literacy of Swedish pupils. Because the NU test was so easy we are unable to say that pupils in their ninth school year performed well.

Pupils in Swedish schools are trained to read texts and deal with them in accordance with the instructions in the national test and current syllabus. For them to then encounter another procedure can yield results that are not completely true. As we noted in the introduction, reading and dealing with a text is not a given, but something we largely learn.

The purpose of this report is to note similarities and differences between the assessment surveys on one hand and the syllabus and its implementation on the other. Real and major similarities between PIRLS, PISA, and the implementation of the Swedish syllabus have been elucidated by presenting the concrete content of these surveys. What we can agree to, having studied the aforementioned surveys, is that there is international consensus regarding our point of departure as well as our way of dealing with a text. The focus is on the same kinds of problems: how are pupils best prepared for a life as active members of a society? And we agree to the answer: by allowing the pupil to confront different kinds of texts, and by allowing the pupil to interpret and reflect, as the principal instrument in the reading comprehension process. It is, of course, always good news when Swedish pupils perform well in international contexts. It is even more pleasing when these contexts appear to bear major similarities with the Swedish steering documents. The convergence, with regard to definitions as well as content, between the assessment surveys and the syllabus and its implementation, is something to bear in mind.

Finally, we conclude with the words of a pupil in an early school year whom we have consulted who explained the benefit of being able to read by saying:

you've gotta learn things like that ... 'cause when you grow up then you have your own kids who also have to learn that they have to be able to read too... 'cause then they can't read for example any book and you have to be able to read fairy tales and stuff like that (interview with a pupil in Andersson & Lundgren, 2005, p. 27)

Some final reflections

(What the Swedish National Agency for Education says about the results)

Sweden is currently participating in a number of international studies (PISA, PIRLS and TIMSS), which are largely governed by international organisations and consortia. The possibilities for Sweden to exercise an influence on the design of such studies is naturally limited. At the same time it is important from a national perspective to assess pupils' knowledge and knowledge development. An alternative to participation in international studies would be for Sweden to implement its own national study, as was the case with the national evaluation in 2003 (NU-03), for example.

Implementing studies on the national level has the advantage of being able to be adapted to Swedish conditions and that the results can be explicitly related to the goals of the Swedish school system. This is an advantage of nationally designed studies. The disadvantages are that the costs of designing the instruments and doing the analyses must be carried alone by Sweden, and that the results cannot be compared with those of other countries. This speaks favourably for participation in the international studies. The risk that the Swedish goals become subordinate to the goals and framework that are formulated for the international studies and that the Swedish goals fall by the wayside speaks against this.

A conceivable alternative might be to use results from the national test to assess levels of knowledge and knowledge development. A critical review of the various national tests shows, however, that they do not work especially well in such contexts. This is simply because they have not been constructed for these purposes, but more as a support for teachers in their interpretation of learning goals and grading criteria (if courses are to be graded) in their efforts to promote just and equal grading. The choice is thus between specially developed national instruments or participation in national studies.

How well do the frameworks for the international studies correspond with the Swedish curricula and syllabuses? Do we risk losing our distinctive national character? Or are our pupils' results being judged on misleading premises because, for example, they have never been given the opportunity in school to learn the demands of the various studies, because Swedish goals differ from what these studies test?

In order to answer these kinds of questions and be able to grasp and interpret the results of national as well as international studies, the kinds of analyses that are presented in this report are important, if not essential. These analyses are also crucial to decisions regarding participation in various international studies as well as for possible planning of national assessments.

The results presented here on reading comprehension are interesting in many ways and lead to (at least) three clear conclusions:

Firstly, there appears to be a great degree of international consensus – which includes Sweden – on the significance of good reading comprehension. This is important knowledge that implies that international studies of reading comprehension are highly relevant to the extent that their framework and their concrete content, i.e. their exercises, reflect the Swedish syllabuses.

Fear of a future risk that international organisations take over and become the ones to formulate global syllabuses and curricula has been expressed in discussions about the possible influence of international studies on national

steering documents. The results show that the concordance among the Swedish syllabuses, their concrete form in the national tests, and (modern) international frameworks is rather good. This is not so surprising in light of the globalisation and internationalisation of our day and age. The empirical evidence is nevertheless important.

Secondly, this study shows that trend measurements are not without problems. That an interest in trends in knowledge development has become increasingly pronounced in recent years is reflected in national as well as international studies. When international studies were first conceived – some 40 years ago (ref) – trend studies were not considered, at least not explicitly and not incorporated into the design of the studies. The trend concept was incorporated from the very start in PISA, and the trend study is one of the explicit objectives now of TIMSS and PIRLS.

”If you want to measure change, don’t change the measure” is a well-known and often-quoted expression in conjunction with trend studies. Unfortunately this is not so easy, as this study shows. It is not always such a good idea to use old instruments to measure new knowledge, since steering documents and thus perhaps even the approach to knowledge as well as methods used to register pupils’ knowledge acquisition have changed over time. It is important to obtain knowledge about possible changes in knowledge over time, but it is worth considering the extent to which instruments can be modified to make them relevant to existing steering systems and capable of adequately measuring trends. This is, however, a problem that concerns both national and international trend studies.

Thirdly, the analyses show that when it comes to studies of reading literacy, the results in the present report indicate that the international studies well fulfil their objective of testing proficiency levels and development in reading even from a Swedish national perspective. More tangible differences in approach do not exist between the international framework for PIRLS and the Swedish syllabuses. This suggests that Sweden should prioritise participation in PIRLS rather than develop its own national instruments to test literacy on a nationwide level. This applies to current assessment of progress as well to trend. It is, however, important that Swedish diagnostic and test material is available to support the pedagogical work of the teachers. The international studies test only reading literacy, and thus cannot be used as a measure of goal attainment, which is important to keep in mind.

7 Literature

Andersson, Kerstin & Lundgren, Ingegerd (2005): *Från avkodning till läsförståelse. Läsutveckling utifrån pedagog- och elevperspektiv*. [‘From decoding to reading comprehension. Reading development from a pedagogical and pupil’s perspective’, trans.]. Unpublished, undergraduate paper.
Norrköping: Institutionen för tematisk utbildning och forskning,
Linköpingsuniversitet, LiU Norrköping, Sweden.

Edfeldt, Åke W (1982): *Läsprocessen* [‘The reading process’, trans.].
Grundbok om läsforskning [‘Basic guide to research on reading literacy’, trans.]
Stockholm: Liber.

Elmeroth, Elisabeth (2005): *Nationella utvärderingen av grundskolan 2003. Svenska och svenska som National Evaluation of the Compulsory School 2003. Swedish and Swedish as a second language*. (Subject report for report 251)
Stockholm: The Swedish National Agency for Education.

Gustafsson, Jan-Eric & Rosén, Monica (2006): ”The Dimensional Structure of Reading Assessment Tasks in the IEA Reading Literacy Study 1991 and the Progress in International Reading Literacy Study 2001.” In: *Educational Research and Evaluation*. Unpublished.

Hasselrot, Titti (2003): *Barns läskompetens i Sverige och i världen. PIRLS 2001*. [‘Children’s’ reading literacy in Sweden and the world, PIRLS 2001.’, trans.].
Stockholm: The Swedish National Agency for Education.

Liberg, Caroline (1993): *Hur barn lär sig läsa och skriva*. [‘How children learn to read and write’, trans.] Lund: Student literature.

Lundberg, Ingvar (1984): *Språk och läsning*. [‘Language and reading’, trans.]
Stockholm: Liber.

Rosén, Monica & Gustafsson, Jan-Eric (2004): *Förändringar i läskompetens. En jämförelse över tid och länder*. [‘Changes in reading literacy. A comparison over time and between countries.’, trans.] (Forskning i fokus no. 22)
Stockholm: The Swedish National Agency for School Improvement.
ISSN 1651-3460.

Rosén, Monica, Myberg, Eva & Gustafsson, Jan-Eric (2004): *Läskompetens i skolår 3och 4. En jämförelse mellan 35 länder*. [‘Reading literacy in the 3rd and 4th school year. A comparison among 35 countries’, trans.] *Progress in International Reading Literacy Study*. (PIRLS). (Forskning i Fokus no. 21) Stockholm: The Swedish National Agency for School Improvement. ISSN 1651-3460.

Taube, Karin, Pettersson, Astrid & Molander, B-O (2001): *PISA 2000*. (Swedish National Agency for Education report 209)
Stockholm: The Swedish National Agency for Education.



Skolverket
Swedish National Agency for Education

www.skolverket.se