The IEA Reading Literacy Study:

Technical Report

Richard M. Wolf (ed.)

International Association for the
Evaluation of Educational Achievement

The Hague, The Netherlands
1995

Forward

As part of IEA's committment to make the results of its multinational efforts known and accessible to researchers and policy makers worldwide, IEA is proud to present this volume on the IEA's Reading Literacy Study.

IEA owes a debt of gratitude to the Study's Directors, researchers, and all those who have contributed to this major international effort, especially to the schools and teachers who ultimately made this effort successful.

Dr. Tjeerd Plomp
Chair, IEA

Dr. W. Frank Hull IV
Executive Director, IEA

April 1995

Table of Contents

Chapter	Title	Author
1	Introduction to the IEA Reading Literacy Study	Richard M. Wolf
2	The Organization and Conduct of the IEA Reading Literacy Study	T. Neville Postlethwaite
3	The Measurement of Reading Literacy: How the International Tests of Literacy Were Developed	Warwick B. Elley
4	Sample Design Procedures for the IEA Study of Reading Literacy	Kenneth G. Ross
5	Questionnaire Development	Hans Wagemaker
6	The Study Archive of the IEA Reading Literacy Study	Andreas Schleicher and Jean Yip
7	The Development of the International Constructs	Ingvar Lundberg and Dirk Hastedt

Appendices

A	Questionnaires-Population A
B	Questionnaires-Population B
C	Variable List of the IEA Reading Literacy Study

Chapter 1

Intrduction to the IEA Reading Literacy Study

Richard M. Wolf

The IEA Reading Literacy Study: An Introduction

Introduction

The Reading Literacy Study conducted by the International Association for the Evaluation of Educational Achievement (IEA) is the largest single subject study of educational achievement ever carried out. The study involved the testing of over 200,000 students in more than nine thousand schools in thirty-two school systems. In addition, over ten thousand teachers of the tested students completed questionnaires which furnished a wealth of information about teacher characteristics and teaching practices in the area of reading. The major results of the study have been reported in a series of publications (Elley, 1992,1994, Lundberg and Linnakyla, 1992, Postlethwaite and Ross, 1992, and Wagemaker, forthcoming).

The purpose of this volume is to provide researchers with a body of information to enable them to conduct secondary analyses of the data collected from the study and to furnish technical information about the study. While the publications produced by the study present a considerable amount of information, it was simply not possible to address each and every question that could be answered by the data. It is hoped that interested researchers and policy makers will carry out additional analyses to answer additional questions that may be asked of the data.

The Study

The IEA Reading Literacy Study is, on the surface, easy to describe. Probability samples of nine and thirteen year olds were selected in each of the thirty-two participating school systems. In some cases, the school systems represented entire nations with a single centralized school system. In other cases, the situation is more complex. In some nations, e.g., Germany, West Germany and East Germany were treated as separate school systems since, at the time the study was planned and carried out, these were separate school systems. However, in other countries, e.g., Switzerland and the United States, each canton or state has its own educational system, but these countries elected to participate as a single school system.

Each student selected to be in the study took a reading test that covered two class periods and spent another class period completing a student questionnaire (see

Appendices A and B) which supplied information about the student and his home. Teachers of the selected students completed a teacher questionnaire that elicited information about the teachers and their teaching practices (see Appendices A and B). The principal of each school that participated in the study or his/her designee completed a school questionnaire that supplied information about the school (see Appendices A and B). Finally, each participating school system completed a national questionnaire that provided information about the organization and administration of the school system. Information about the development of these questionnaires is presented in Chapter 5 while information about the tests and how they were developed is presented in Chapter 3 of this volume. Details on the organization and conduct of the study are presented in Chapter 2 while pertinent information about how the sampling was planned and carried out is presented in Chapter 4.

The study generated a wealth of information. Each response to each test question or questionnaire item has been preserved in a data archive. Details on the handling and archiving of the data from the study are presented in Chapter 6. Interested persons who wish to obtain a copy of the data archive for further analysis may write to the IEA Headquarters in The Hague. The address is:

IEA
Sweelinksplein 14
2517 GK The Hague
The Netherlands

Special Features of the IEA Reading Literacy Study

There are four special features in the IEA Reading Literacy Study that may not be apparent to the typical reader of the publications resulting from the study. Each is important in its own right, but, together, they make this a model for how an international study should be planned and carried out. These need to be noted. Each will be presented and discussed in turn.

1. The IEA Reading Literacy Study was extremely well organized and carried out. The effort that was required to organize and conduct a worldwide study was monumental. Problems of languages differences, wide geographical separation, and differences in culture and research traditions presented formidable problems for the researchers who were involved in the study. That these problems were conquered is a testament to the dedication of the professionals involved. Chapter 2 presents a considerable amount of detailed information as to how this was accomplished. It

should be read carefully to gain a full appreciation of the challenges that the researchers faced and overcame.

2 The tests that were developed to measure reading literacy at the nine and thirteen year old level are among the best ever developed for use in comparative studies. They are not only technically sound, but take into account different language and cultural traditions. They are worthy of study in and of themselves. Chapter 3 details the procedures that were followed in developing the tests. They furnish an excellent model for the development of test instruments.

3. The researchers who worked on the IEA Reading Literacy Study developed and tested an explicit model for the development of reading literacy. The model is presented in Elley's report (Elley, 1994) and is also included in Chapters 5 and 7. Figures 2 and 3 in Chapter 5 set forth the model while Chapter 7 details the composition of the constructs. Results of analyses presented in the major publications of the study employ weights derived from an international model of the development of reading literacy. However, if national analyses are undertaken, it would be appropriate to use nationally derived weights for the components of the model to test a national model. This should provide a set of rich opportunities for additional analyses.

4. The IEA Reading Literacy Study was an international cooperative venture. Each participating school system (in some cases, nations) had a representative to the study. The representative was called the National Research Coordinator (NRC). The NRC's met once or twice a year over the course of the study to review and discuss the work to date and plan for future work. It was out of these meetings that the design of the instruments and the conceptual model of the study were developed. The continual interchange between the NRC's, the members of the International Steering Committee, the staff of the International Coordinating Center and invited experts in various fields resulted in a high quality effort that attended to the needs of the participating school systems and criteria of intellectual excellence. It meets the highest standards of intellectual cooperation.

The chapters of this volume present a wealth of information, much of it technical in nature. This is entirely appropriate for a technical report. It should enable readers to acquire the information to carry out additional analyses of interest and to understand

the procedures followed in the research presented in the major publications of the study.

One further comment is in order. The readers of this report will quickly note that type style, bibliographic style, and organization vary considerably from chapter to chapter. Each chapter was prepared by a different author who adopted his/her own convention for preparing a chapter. This has led to some unevenness. To standardize the material would have required a considerable amount of time. More important, any effort to standardize the chapters could have resulted in the introduction of errors. Accordingly, it was decided to present the various chapters in each author's style (after appropriate review) so that the reader will not be burdened by errors that could easily have resulted by attempting to put the volume into a single style.

References

W.B. Elley (1992) *How in the World Do Students Read?* Hamburg, Germany, Univ. of Hamburg.

W.B. Elley (1994) *The IEA Study of Reading Literacy: Achievement and Instruction in Thirty-two School Systems.* Oxford: Pergamon Press.

I. Lundberg and P. Linnakyla (1992) *Teaching Reading Around the World.* The Hague: International Association for the Evaluation of Educational Achievement.

T.N. Postlethwaite and K.N. Ross (1992) *Effective Schools in Reading.* The Hague: International Association for the Evaluation of Educational Achievement.

H. Wagemaker (ed.) (1995) *Gender Differences in Reading.* The Hague: International Association for the Evaluation of Educational Achievement.

Chapter 2

The Organization and Conduct of the IEA Reading Literacy Study

T. Neville Postlethwaite

ORGANIZATION AND CONDUCT OF THE IEA READING LITERACY STUDY

The aims of this chapter are to describe briefly the organization and conduct of the study (Section 1) and then to comment in more detail on problems encountered and the lessons learned (Section 2).

Section 1: Organization and conduct of the study

The initiation of the study

In 1985, the IEA General Assembly, which consists of the senior representatives of the research institutions in the systems of education which are the members of IEA, agreed that a study of Reading Literacy should to be carried out. This was based on a proposal written by Professors Elley, Lundberg, Guthrie and Purves. It was the responsibility of the IEA General Assembly to raise the funds for the study, establish an International Coordinating Center (ICC), appoint a coordinator, and a Steering Committee.

Initial funding. In late 1988 early 1989, initial funding was granted by the MacArthur Foundation and the Mellon Corporation.

Establishment of a Steering Committee. In late 1989, the IEA appointed a Steering Committee consisting of Warwick Elley (New Zealand), Ingvar Lundberg (Sweden), John Guthrie (USA), Ken Ross (Australia), and Francis Mangubhai (Fiji). Each of these persons was eminent in either the study of reading, reading test construction, or probability sampling for educational surveys. Alan Purves (USA), who had been active in the drafting of the proposal became the General Assembly's ex-oficio member of the Steering Committee in 1989. Nelson Trujillo-Rodriguez (Venezuela) was a member of the Steering Committee for the first year of the operation but had to drop out because of other commitments. Albert E. Beaton (USA) joined the Steering Committee in 1990. Peter Allerup (Denmark), and Nadir Atash (Afghanistan and working at Westat in the USA) joined the Steering Committee in 1991.

It should be pointed out that none of these persons was paid for the work they undertook for the study and yet their investment of time was enormous – both between meetings and at meetings. In retrospect it would have been highly desirable to buy their services from their institutions for longer periods of time. However, given the budget for the study, this was, unfortunately, not possible.

Participation in the study. In late 1989, some 40 IEA National Centers indicated that they were interested in participating in the Reading Literacy study. As with many international studies, some dropped out primarily because of a lack of national funding and one or two joined after the study was in progress. Each National Center appointed a National Research Coordinator (NRC) who was the person to be in charge of the day to day work of the study in the National Center. Each National Center had to pay the travel cost for the NRC to attend NRC meetings. The host institute for an NRC meeting covered the cost of board and lodging for one NRC from each country.

Establishment of an International Coordinating Center (ICC). A small International Coordinating Center (ICC) was established at the University of Hamburg in November 1988. The IEA General Assembly invited Neville Postlethwaite, who worked at that university, to be the international coordinator. Neville Postlethwaite accepted the task on condition that he orchestrated the work but could employ an assistant coordinator who would be responsible for the day to day work. As with the members of the Steering Committee, Neville Postlethwaite had no remuneration for the work. The notion was that the assistant coordinator would grow into the job and be able to take over the coordinators job after about two years.

It was important to have a data processor/analyst involved in the study from the beginning. This was Andreas Schleicher (also the author of Chapter 6). In turn, he established the ICC data processing unit (see Chapter 6 of this report). In January 1992 he replaced Neville Postlethwaite as international coordinator.

A secretarial staff was also established under the assistant coordinators: Dieter Kotte (Nov. 1988 – July 1990) and Dr. Elaine Degenhart (Sept. 1990 – Oct. 1992).

First meetings

The first meeting was hosted by the National Center for Educational Statistics (NCES) in Washington, DC (USA) in November 1988. It was at this first meeting that the major aims, overall research design and procedures of the study were agreed upon. To a very large extent, these remained as had been proposed with only minor amendments.

The major aims of the study were to:

- estimate the levels of literacy achievement of 9- and 14-year-olds;
- describe the voluntary reading activities of 9- and 14-year-olds;
- identify the relationships between policies and instructional practices on the one hand and reading literacy achievement and reading activities on the other hand.

There was much debate on the definition of Reading Literacy. There are many schools of thought. In the end, the definition agreed upon for this study was "Reading Literacy is the ability to understand, and use, those written language forms required by society and/or valued by the individual". The focus of the study was on *reading* literacy. Tests of reading in three domains were to be developed:

1) *Narrative* texts in which the writer's aim is to tell a story.
2) *Expository* texts in which the aim is to describe or explain something.
3) *Documents* such as forms, charts, maps, directories in which the task is to locate facts within a structure.

A picture-word matching test was also to be constructed to test the word recognition of 9-year-olds.

There was also much discussion on the problem of which student, teacher, and school variables to include in the study. Each variable was to be justified theoretically and also in terms of how it would be used in the analyses.

Following this first meeting, the Steering Committee requested items from the National Centers for the domains to be tested. Nearly all National Centers provided passages and items. Three members of the Steering Committee met in February and March 1989 to select (and, where necessary, edit) the passages and items. The results of this meeting were reviewed in a full Steering Committee meeting in April 1989. At the same time, work continued on the background questionnaires.

A draft sampling manual was written. This was later revised and then used by all NRCs for designing their probability samples (see Chapter 4).

The ICC produced a series of manuals for the pilot testing (one for National Centers, one for School Coordinators for the testing, one for the test administrators, and two for data entry at the National Centers). A Data Entry Manager program and manual were written specifically for the study. The program was written for the PC. The decision was also taken to conduct all data

processing and run all analyses on PC's at the ICC. With the increased capacity of PC's it was felt that to rely on operators mounting and dismounting tapes at a main frame computer would not be a desirable situation.

In July of 1989 a second NRC meeting was hosted by NCES in Washington, D.C. The aim of the meeting was to review the general procedures for the conduct of the pilot study and to finalize all of the instruments. Following this meeting, the instruments and manuals were prepared and sent out to the NRCs who had to translate or arrange for the translation of all of the instruments and prepare for their national pilot study data collection.

It was very useful that Warwick Elley, at his own expense, was able to visit nineteen of the participating research institutes in 1989. This helped NRCs and their national committees on points of clarification.

The Pilot Study

Following the July 1989 NRC meeting, NRCs prepared for the pilot testing. For most NRCs the major task was the translation of the instruments. This was typically done by the members of the National Committees in such a way that two persons independently translated each item. A set of principles developed by the Steering Committee guided the translation work. Where there were discrepancies, they were resolved by the committee as a whole. The translated versions of the pilot instruments were sent to the ICC where they were checked. At the same time, NRCs sent in documentation on the problems they had encountered in translation or on deviations they had made in background questionnaire items because of the way in which their system was organized. All of these comments and deviations were compiled in a master list at the ICC.

All instruments were trial tested on judgment samples at both population levels in all countries participating at that time. In the pilot study, four forms of test items were rotated because of the large number of items used. The first systems (Botswana and New Zealand) tested in late September/early October 1989 since their school years ended in early to mid December. Other systems conducted the pilot testing between January and March 1990. In theory all pilot data sets were meant to be at the ICC by the end of March 1990 so that the cleaning of the data and the production of the item analyses for test items and univariates for all background variables could be produced for a meeting of the Steering/Technical Committee to review all of the work on 7 May. Not all NRCs met the deadline of the end of March. Indeed, the last data set (the Netherlands) arrived on May 5. Despite this, the Rasch and classical item statistics as well as all univariates were ready by the beginning of the Technical/Steering Committee meeting. Decisions were then taken on which test items to drop, and on the rewording of other test items and questionnaire items. Checks were also made on the effectiveness of translations by studying the patterns of item difficulty indices across nations and items (see Chapter 3).

Revised versions of the instruments were produced by mid June and sent to all NRCs in time for them to meet with their National Committees before coming to the July 1990 meeting of NRCs in Frascati, Italy. This meeting was hosted by the Centro Europeo Dell'Educazione.

Final Preparations for main testing

The 1990 Frascati NRC meeting was *the* meeting upon which the rest of the study depended. It was at that meeting that:

1) all instruments were finalized;
2) all national sampling designs were to be finalized, i.e. the sampling referee was to agree to a final sampling design with each NRC;

3) the conduct of the main testing was established in detail and Manuals 1, 2, and 3 (for NRCs, for School Coordinators, and for Test Administrators) were finalized;
4) the DataEntryManager program was explained to those not having used it in the pilot study. It was now clear to all NRCs that it was advantageous to use this program;
5) "final" lists of hypotheses to be tested were planned.

Three NRCs were not able to attend the meeting because of lack of local (national) funding. In two cases, this could be solved by the ICC or by another NRC instructing the missing person. In the third case (Nigeria) no money was available for a member of the ICC data processing staff to visit Nigeria. The result was that there were serious deficiencies in the Nigerian sample and data entry. The student file was salvageable but this was not the case for the teacher and school files.

Upon returning to their National Centers, the NRCs had to check or revise the translations of each item in the tests and questionnaires as well as Manuals 2 and 3.

One of the questions had to be devised nationally. This concerned "possessions in the home". NRCs had to select 10 possessions (e.g. video recorder, car, washing machine, etc.) such that when the 10 items were added together pupils could be placed on a scale ranging from 0 – 10, indicating the "wealth" of their homes. This required that there was variance on the possession of each item. In other words, there was little point in including items possessed by more than 90 percent of homes or possessed by fewer than 10 percent of homes. The analyses of the pilot data identified those items with very little variance in a particular country. This information was given to the NRCs with the request that they remove such items and, on the basis of consumer information in their country, replace them with items where there would be variance. Most NRCs accomplished this task.

The NRCs also had to draw the sample of schools or classes and two parallel samples from which replacement schools could be drawn, contact the schools and achieve a final sample such that there was a 100 percent response rate of schools or classes. This exercise is described in the Sampling Chapter later in this volume.

The main testing

The main testing was meant to take place in the first week of the eighth month of the school year in each school system. Because of the different starting dates of the school year in different systems the testing took place between September 1990 (Botswana) and June 1991 (Germany). Table 1 presents the date(s) of testing, the age of entry to school, the grade level tested and the mean age and standard deviation (in years) for each sample in each country.

From this table, it can be seen that the date of testing was not at the same point of the school year in each system, nor was the grade level tested the same. This variation is to be expected given the differences in the age of entry to school. The mean ages of the student samples were also different. This weakens the comparability of systems within each population level. The reasons for this are explained later in the Sampling Chapter (Chapter 4). The issue of comparability is discussed in Section 2 of this chapter.

Table 1. Date of testing, age of school entry, grade level, mean age, and standard deviation for each sample in each country

	Date of testing 1991	Age of school entry	Population A			Population B		
			Grade tested	Mean age	SD age	Grade tested	Mean age	SD age
BEL/FR	15.2–15.3	6	4	9.83	0.61	8	14.75	0.94
BOT			—	—	—	7	14.69	1.11
CAN/BC	21–28.3 8–12.4	6	3	8.89	0.47	8	13.93	0.54
CYP	3.5 A 29–30.4 B	5y, 6m	4	9.75	0.39	9	14.77	0.45
DEN	4–8.3 11–15.3	6	3	9.75	0.36	8	14.75	0.36
FIN	5–15.3	7	3	9.72	0.35	8	14.74	0.32
FRA	12–17.4	6	4	10.08	0.73	9	15.42	0.86
GER/E	1.3–15.6	6	3	9.45	0.38	8	14.41	0.40
GER/W	1.3–15.6	6	3	9.44	0.52	8	14.59	0.61
GRE	Dec–Jan A Feb–Mar B	6	4	9.26	0.33	9	14.38	0.44
HK	Feb–Mar	6	4	9.96	0.65	9	15.23	0.84
HUN	4–8.3	6	3	9.34	0.56	8	14.11	0.48
ICE	8–12.4	6	4	9.79	0.29	9	14.80	0.31
IND	14–20.2	7	4	10.77	1.10	—	—	—
IRE	Mar–Apr	6	3	9.23	0.48	8	14.45	0.56
ITA	13–25.5		4	9.86	0.49	8	14.05	0.64
NET	Mar–Apr	4 (6)*	3	9.22	0.46	8	14.34	0.62
NZ		6	3	10.00	0.33	8	15.03	0.37
NIG	Feb–Apr	5 or 6	—	—	—	9	15.27	1.23
NOR	1–14.3	7	3	9.75	0.34	8	14.76	0.35
PHI	Feb–Mar	7	—	—	—	8	14.48	0.29
POR		6	4	10.38	1.07	9	15.62	1.17
SIN	Sep, Nov 1990	6**	3	9.32	0.32	8	14.41	0.69
SLO	Apr. 91	7	3	9.72	0.41	8	14.70	0.39
SPA		6	4	10.00	0.49	8	14.20	0.67
SWE	11–15.3	7	3	9.75	0.31	8	14.75	0.33
SWI	March	6 or 7	3	9.73	0.55	8	14.85	0.68
THA	Feb.	6	—	—	—	9	15.20	0.63
T/T	Mar–May	6	3	9.57	0.58	8	14.42	0.56
USA	Feb–Mar	6	4	9.99	0.56	9	15.01	0.67
VEN		6	4	10.70	1.29	9	15.52	1.07
ZIM	Mar–Apr	6	—	—	—	9	14.49	1.04

NOTE: * Compulsory school entry at 4, instruction in reading begins at 6
** Not compulsory

All data sets were to be returned to the ICC by the end of July 1991. For the most part they came in between January and August 1991. One National Center – because of severe illness of the NRC – had mishandled the data entry. Upon return to work, the NRC saw what had happened and re-entered all of the data in a very short space of time and extremely accurately.

The cleaning, weighting, and merging of the data

Special programs were written to clean the data. The sample weights were calculated (school and student) and the data files were merged. The initial cleaning was completed by October 1991 and that, at an NRC meeting, hosted by the Danish National Institute for Educational Research in early October 1991, NRCs checked their univariate data for unexpected values and odd results. NRCs reported to the ICC data processing staff how systematic errors might be corrected. A document (see Appendix D) was compiled variable by variable, indicating the percentage of missing data and special comments. These included such matters as the variable having to be dropped because it was mistranslated, it was misinterpreted by respondents, respondents answered what the authorities might expect from them rather than the reality of the situation (the "compliancy" effect), and explanatory comments needed by authors and users of the data.

Systematic errors were subsequently corrected at the ICC. The sampling weights were calculated and files were merged.

A final item analysis

A second full item analysis (classical and Rasch) was undertaken in January 1992. As with all forms of test construction the aim was to have a measure or measures which are valid for what pupils are intended to achieve. However, when achievement is scaled for the purpose of showing the amount of differences among groups, then the scaling of the underlying traits of the different types of reading must be unidimensional. This sometimes creates a conflict, in terms of the items to be included in scores between the defined construct and the achievement of unidimensionality. How this problem was dealt with in the area of Reading is presented in Chapter 3. On the basis of the item analyses decisions were taken on:

1) which countries to drop for the total calibration exercise for a particular domain;
2) which items to drop for all countries from the calculation of a domain score; and
3) which items to drop within a country and then estimate using the Rasch procedure.

Following this procedure, Rasch item statistics and person scores were calculated for each domain at each population level.

Construct Analysis of Background Questionnaire Items

The Steering Committee made an *a priori* allocation of questionnaire items to constructs. A series of factor-analytic approaches were used (varimax rotation, procrustean bed, and principal components) in an attempt to create the constructs desired by the Steering Committee. Dirk Hastedt of the ICC data processing team, together with Prof. Ingvar Lundberg of the Steering Committee, were responsible for the final constructs' analyses. The analyses undertaken are presented in Chapter 7. They were begun in September 1991 and finalized in March 1992.

Production of data summaries and files for all NRCs

In February and March 1992, the ICC produced summary univariate statistics for all background variables and all scores. For any one variable either the mean, standard deviation and distribution or the percentages selecting each option in a question were presented for all countries/systems on one page. (For more detail see Chapter 6.) These were sent out to all NRCs. NRCs were also able to receive their cleaned, weighted, merged files either in SAS format or in raw data file format. These files, together with documentation, were prepared at the ICC and sent out in February and March 1992.

Data analyses and international reports

A decision was taken early in the study that instead of waiting till a final international research report was ready, the study would produce a set of booklets dealing with initial results during the second half of 1992, but that the large international report and the technical report would be prepared for the end of 1992.

For each booklet and the international report, a detailed outline was prepared together with a detailed set of data processing instructions. These went through a series of reviews by the Steering and Technical Committee and were amended several times. Authorships and the number of booklets to be written were also revised as situations changed.

In the end the following booklets, book and technical reports were produced:

1) How in the World do Students Read? A study of the Reading Literacy of 9- and 14-year-olds in 32 countries.
2) Effective Schools in Reading: Implications for Educational Planners
3) Teaching Reading Around the World
4) Gender Differences in Reading Literacy in 32 Countries
5) The IEA Study of Reading Literacy: Achievement and Instruction in Thirty-Two School Systems
6) IEA Reading Literacy Study: A Technical Report

The booklets and books were written by either members of the Steering Committee and/or ICC members, and/or NRCs. Each publication was reviewed by one or more NRCs, selected members of the Steering/Technical Committee and the IEA Editorial Committee.

It was agreed at an early point in the study that National Reports published after the publication of the first booklet could also include the univariate background statistics and student reading literacy scores from other countries. It was, therefore, important to set a firm date for the publication of the first booklet. This was set for 13 June 1992. Several NRCs, therefore, planned to publish their national reports in the week beginning 15 June 1992. The date of 13 June for the international report thus became critical. In turn, this required that the data processing and data analysis timetable (which was, in any case, tight) had to be strictly kept. This, as will be seen later in the "Lessons Learned" section, did involve some unexpected problems and there was a delay of three months. The analyses for the first booklet were produced at the ICC and written up by the author both at his home base in New Zealand and at the ICC.

The analyses for the other booklets were executed according to the data specification submitted by authors. The analyses were sent to the authors who, in turn, required further analyses. The scheduled publication dates of the booklets were:

Booklet	1	September 1992
	2	November 1992
	3	January 1993
	4	April 1993

These booklets were prepared for publication by the ICC and they were also printed by a local firm in Hamburg. Arrangements were made with the International Reading Association with headquarters in the United States to sell the first booklet. At the time of writing this chapter, the IEA HQ was negotiating with the International Reading Association about the distribution of the

other three booklets. It is clear, however, that the IEA HQ needs to develop a general policy about the distribution and sales of booklets if future projects are to produce booklets.

The international Report followed a similar set of steps. The editor of the International Report, Prof. Warwick Elley, designated specific persons involved in the study to be responsible for specific chapters in the report. Data analysis specifications were made and sent to the authors. Some of the analyses were replicated within system analyses.

It was clear that multilevel analyses would be required. Two members of the Reading Literacy Study (a RL Steering Committee member Prof. Ingvar Lundberg, and Dr. Ingrid Munck who worked with the Swedish National Center) took the initiative of applying for Swedish funds to allow the ICC to perform multilevel international analyses at the University of Gothenburg. Not only was a high power computer available but also a number of experienced staff under the leadership of Prof. Jan-Erik Gustafsson, and Prof. Bengt Muthén. The models were developed by the Steering Committee during the first six months of 1992 and tested in June and September 1992. The writing of the International Report was undertaken in the period April to November 1992. A first draft was reviewed by the NRCs in October 1992 at an NRC meeting in Madrid, hosted by the Spanish National Center (Centro de Investigación, Documentación, y Evaluación) at the Ministry of Education and Science. The NRCs gave their comments to the general editor.

Rewriting was undertaken in Hamburg at the end of October by Warwick Elley and Ingvar Lundberg. Some further analyses were required. These were completed in November, 1992. Further writing and editing took place in November and December, 1992. The final draft manuscript was sent to the IEA Editorial Committee in January, 1993. Further revisions were undertaken on the basis of the reviewers comments. The final version of the international report was sent to the publisher in March, 1993.

Overall Timetable

Table 2 presents the overall timetable and meeting places of the Reading Literacy Study from November 1988 to the end of 1992.

Table 2: Brief Summary of Overall Timetable

1988	
November	First meeting of NRCs, Steering Committee and ICC members (Washington DC – hosted by NCES)
December	National Committee members meetings and begin to submit test items via the NRCs
1989	
January	Submission of test items. Writing of Sampling Manual. Initial work on background questionnaires
February	Subcommittee meeting (Singapore) of test committee for screening (about 5,000 items), selecting, and modifying passages and items
April	Subcommittee meeting of test construction committee (San Francisco – hosted by NCES)
May – June	Writing of Manuals. Review of tests by National Committees
July	Second Meeting of NRCs, Steering Committee and ICC members (Washington DC – hosted by NCES). Pilot versions of tests and questionnaires produced. First sampling plans reviewed
August	Translation of tests, and questionnaires into national languages. Writing of Codebooks and data entry program September Pilot testing in countries. Preparation for pilot analyses
1990	
Jan – Mar	Data cleaning and verification at the ICC
April	Classical and Rasch analyses of items
May	Steering Committee meeting. Selection of passages and items for final tests. Preparation of questionnaires
June	Typing of final international tests and questionnaires. Review by National Committees
July	Third NRC meeting (hosted by CEDE – Frascati, Italy). Final Review of all instruments and final changes made. Final Sampling plans reviewed
August	Final version of international instruments and Manuals, 1, 2 and 3 sent to all NRCs
September	Samples drawn. Main Testing in first set of countries. Writing of final codebooks and Manual 4, 5 and 6. Visit by ICC to some countries
December	Return of first country's data to ICC
1991	
Jan. – May	Visit by ICC to six countries. Main testing in remaining countries. Submission of some data sets to ICC and Valid Range Forms and national documentation
April – July	More data sets submitted to ICC. Cleaning of data files
July	All but one data file received
August	Cleaning of data files continued. Production of univariates and item analyses
November	Fourth NRC meeting (Hosted by Danish Institute for Educational Research). Checking of univariates: list of problem variables produced)
1992	
January	Merging of data files for each target population. Sampling weights calculated. Construct analyses begin
February	Production of univariate statistics for all variables, all populations, all countries. First analyses for first booklet. Analyses for second booklet. Further analyses
March	Meeting of Steering Committee (Hamburg). Decision on scoring and scaling. Meeting of Gender Booklet Committee
April – May	Calculation of scales and recalculation of scales
June	Preparation begun of first booklet. More analyses for second, third and fourth booklets. Most constructs finalized. Data covariance matrices produced
July – Sept.	Further work on all booklets and international report. First booklet published. Constructs finalized
October	Fifth NRC meeting (El Escorial – hosted by Spanish Ministry of Education) Booklets reviewed. International report reviewed. Further analyses specified. National analyses reports given. Lessons learned described
November	Further analyses and writing of booklets, international report and Technical Report. Booklets 2 and 3 published. Project closed down at end December

Financing

Table 3 provides an overview of the total international income and expenditures.

The original budget for the four years was estimated at US$ 2.5 million. It was not possible to raise all of the funding. Further to this, there were delays encountered in the reception of the NAS money with the result that the IEA HQ had to lend money to the Reading Literacy Study. The lack of money, as will be seen in Section 2 of this Chapter meant holding fewer of the minimally foreseen number of meetings, fewer visits of the ICC to National Centers, no visits of the Sampling referee to National Centers, no visits of NRCs to the ICC in 1992 for work on both international and national analyses, and no meetings of the Steering Committee for the multivariate analyses.

The International Coordinator, the Steering Committee members and the Technical Advisors did not receive any salaries or honoraria for the many months of work they devoted to the study. The University of Hamburg, site of the ICC, provided rooms for the project and covered the telephone (including fax) and postage expenses. Continuous direct communication is essential to the successful operation of an international project which is based on the cooperative input of its far-flung participants. Therefore the telephone and postage bills are quite high and would have added considerably to the ICC budget.

Table 3: Overview of Total Income and Expenditure

Income	**US$**
MacArthur Foundation	168,000
Mellon Foundation	135,000
US National Academy of Sciences	557,000
European Commission	40,000
UNESCO	5,000
National Centers	152,000
Other (including interest)	29,000
Total	1,086,000

Expenditures	
Personnel (including benefits)	678,000
Office Costs	74,000
Computer purchase and maintenance	122,000
Meetings	140,000
Other*	72,000
Total	1,086,000

* Other includes moving expenses for the second assistant international coordinator and the printing of the booklets.

As can be seen, the total international costs were just over US $ 1 million covering the period from November 1988 to December 1992. Just over half of the funding came from the US National Academy of Sciences. However, if the MacArthur and Mellon foundations are included, then some 80 percent of the international funding came from the United States of America.

In terms of expenditures, the largest item was personnel followed by meetings and computer purchase and maintenance. Of the total, 13 percent was spent in 1989, 26 in 1990, 23 in 1991, and 38 in 1992.

Section 2: Lessons Learned

As already mentioned, there is many a slip 'twixt cup and lip. This section examines the types of problems encountered in the course of the study and makes suggestions, either directly or indirectly, about what might be done to improve the conduct of future international studies.

Planning a Study

Although a proposal must be prepared by IEA in order to obtain funding before a study can begin it is nevertheless extremely important that the researchers in each country are involved in the conceptualization of studies. Thus, there is a balance required between the writing of the proposal for obtaining funding and the modification of the design (and even some aims) by the national researchers themselves. Without the cooperation of the national researchers no study will take place. Many of the national researchers are experienced not only in conducting their own national surveys but also in other international studies and even in previous IEA studies. Their input can raise the quality of a study from a theoretical point of view and, at the same time, ensure that the study is "doable" in the sense that sampling and the data collection procedures will work in all countries. Thus, the writing of a new study proposal should involve some experienced national researchers.

Any study proposal should include a clear statement of the general and specific aims of the study and a priority stated among these aims. It should be accompanied by an extensive review of past international research in the area and relevant national research. It is the prioritized aims that determine the research design. The exact definition(s) of the target population(s) should be given. The measurement model for the background instruments should be clear and the outcome measures (at different levels of aggregation) should be clearly specified. Furthermore, the definition of the achievement scores and scaling methods to be used should be clear.

There should also be a statement about the ground rules for the study. These include, *inter alia*, minimum participation conditions, minimum sample accuracy requirements, international compulsory and international option parts, the amount of work required of the national researchers and of national committees.

In short:

1.Involve selected national researchers in the writing of any new study proposals;
2.Ensure that the general and specific aims are clear and in order of priority;
3.Include a review of all past international and relevant national research;
4. Provide clear definitions of the target population;
5.Explain the measurement model for the background questionnaires (i.e. which indicators for which purpose associated with which aims and, where appropriate, theories);
6. Provide information on minimum participation;
7. Provide information on minimum sampling accuracy requirements;
8. Provide clear information on the outcome scores (cognitive and affective) to be producesd and the form of scaling to be used.

NRCs

A special comment should be written about NRCs. Ideally, an NRC should have experience in international survey research (or at least large scale national survey research) and be a subject matter specialist in the subject under study. If it is not possible to identify someone with both sets of skills then someone with the former skills should be identified. This point should be made absolutely clear by the IEA HQ. It is the IEA General Assembly that admits national centers to IEA

on the basis of their proven survey research experience. It is the heads of National Centers who designate NRCs. These NRCs must have an overall knowledge of test construction, attitude scaling, sampling, data management and data analyses as well as of project management. It is the NRCs who are personally responsible for the day-to-day management of the study. And, still, training should be built in via seminars and longer NRC meetings.

Length of study

The study was planned to last 4 years from the first meeting of the National Research Coordinators to the completion of the final publication. It is to be noted that the first IEA mathematics study took from mid 1962 to mid 1966. Subsequent IEA studies took longer.

However, it would seem that 4 years was an appropriate time period for the Reading Literacy study. But, much depends on:

1) The organization and orchestration of the study being very well planned and executed;
2) the key persons being well-selected, committed to the study, and able to work hard between meetings; and
3) there being sufficient funding to ensure that these persons can be freed from commitments to their normal work, and that the ICC can employ sufficient full-time staff from the beginning of the study.

It should be noted that money does not buy commitment and hard work. Persons involved in the study should possess these "virtues" but adequate funding can facilitate their work.

Participation: National systems or subsystems

Some nations such as Singapore, the Scandinavian countries and Hungary have one national education system. Others have several administratively independent systems. Countries such as Belgium have the Flemish system and the French system; Australia has 6 states and 2 territories; Canada has 10 provinces and 2 territories, and Switzerland has 26 cantons. In this study, each country or system was allowed to decide whether it entered as one entity or many. Thus, the United States (with 50 states) was one entry and British Columbia (as one province of Canada) was also one entry.

Furthermore, the IEA Chairperson informed the ICC that the ex-East German and ex-West German systems were to be analyzed and presented separately in all Reading Literacy publications, even though the testing took place after reunification. The reason given was that the students were the products of two systems and that this was a unique opportunity to assess the effects of the two systems separately. However, Switzerland wished to be represented as one nation. Since these nation and system values are entered as rows in tables, then the average of the column values (with equal weights because a system is a system) from the grand mean, the problem arises that each row carries the same weight whether it be one part of Germany, one Canadian Province, the United States or Indonesia.

This "political" problem is also related to the practical problem of the forms of multivariate analyses to be undertaken. To formulate and test path models (by which ever method of analysis) is a time-consuming task. As more subsystems are allowed into a study so the *time* required for multivariate analyses within countries is increased. When many national systems and subsystems participate in a study the aim of formulating and testing one large international model as opposed to replicated within system analyses must be decided from the outset. Alternatively, the study could decide on a maximum number of systems to be allowed into the study. But, what would this number be? And, would such a restriction be politically acceptable?

As can be seen, the mixture of subsystems and national systems can be an uneasy one. This also seems to be a politically sensitive issue but it is incumbent on any study to have some rational guidelines about this from the very beginning.

Definition of the target population

Some studies have demands for strict comparability of the target populations tested. Thus, for example, *all* 14-year-olds in one country should be compared with *all* 14-year-olds in other countries. In this study, the modal grade of regular schooling for 14-year-olds was identified and then a probability sample of all pupils in that grade was tested. The age means and distributions were – as was to be expected – not identical, or even similar, in all countries. It was at the first NRC meeting that this problem was taken up. The NRCs argued for testing at the modal grade only because of the administrative difficulties of test administration across grades. The consequence of this decision – namely the problem of strict comparison – was known. It was interesting, therefore, in postulating "lessons learned" at the end of the Reading Literacy Study, the NRCs recommended that future IEA studies define their target populations in such a way that age and grade effects can be separated, for example, by testing multiple grades.

There are several problems when setting the defined target population. First, the pupils in an age-defined group can be split across at least two grades at the time of testing. Secondly, in systems practicing grade repeating, the pupils can even be spread across 4 or 5 grades. These are the pupils in regular schooling. The third problem is that the percentage of the age group in special education schools ranges in different countries from less than 1 percent to 9 percent. These pupils are typically excluded from the defined population so that even the defined populations are actually not comparable.

Absolutely comparable samples of pupils are, therefore, never likely to be obtained. The question then becomes "What is a reasonable basis for comparison?" Should there be age or grade sampling or both? If a study were to identify the modal grade group and then take only the 14-year-olds from that grade *plus* the 14-year-olds from the grade groups on either side of the modal grade group it could be expected that between 1 and 15 percent of the age group would be excluded. Would the resulting population then be "reasonable" for comparison purposes? Each future study will need to deal with this problem from the outset.

However, although the adjacent grade approach may help somewhat towards solving the problem of making fair comparisons there is still the problem of undertaking multivariate analyses where the teacher and school characteristics as well as pupil characteristics are examined in terms of their effect on differences in achievement between pupils or between schools. This requires that either an intact class is drawn and the teacher(s) characteristics for that class are related to their pupils' achievement or that the average of all teachers (for the subject matter being studied) is used. It can be argued that the teacher of a class this term or year is merely a proxy for all teachers in the school since the pupils have been exposed to other teachers in previous grades. This, of course, is the case for a cross-sectional study but not for a longitudinal study where the particular teacher would have to be used if the aim of the study was to examine, say, the effect of teacher behaviors on change in pupil or class achievement. However, the averaging of several teachers to form a proxy teacher can also be hazardous especially when attempting to average teachers' responses on such variables as their perceptions of strategies and methods of teaching.

As will be seen in the chapter on sampling (Chapter 4) and, as already mentioned, only the modal grade for the age group was taken and within that grade only one intact class was drawn. If there was only one class in the grade then that class was taken. If there were two or more classes, then typically one class was drawn at random. There were one or two NRCs who deviated from this plan in that they drew a random sample of classes rather than schools as the primary sampling unit. Or, in one or two cases per school (where there were two or more classes) were drawn.

The reason for drawing an intact class within the sampled school was that many NRCs said that it would be administratively impossible to draw pupils either at random from all classes in the modal grade or from adjacent grades. There were two reasons given. The first was that no extra rooms for testing were available in a school or that school principals would refuse to participate in the study because it would cause too much disruption in the school. The second reason was that some of the target population would be in other schools. For example, in Italy half of the 14-year-olds are in the last grade of Scuola Media but the other half are in many types of secondary schools. Thus, administratively, the burden would be insuperable. Indeed, it is already a difficult task to obtain a high response rate on the modal grade for an age group.

The definition of the target population is not an easy matter. Several of the questions to be asked at the beginning of a study include:

1) *To what extent is strict comparability across countries required? If so, what constitutes reasonable comparability of the operational definition of target populations?*

2) *What are the exact guidelines for defining the excluded population and reporting it? Will multivariate analyses be undertaken? If so, will this be on students drawn at random within schools across grades and, in this case, which teachers will be involved? Or, will a separate intact class sample be drawn? Or, should two intact classes be drawn in order to have some estimate of the between class within school vs. the between school effects. In this case, one would have two separate target populations: One for the "strict" comparison of scores across countries and a second for multivariate analyses as was done in IEA's first mathematics study (Husén, 1967).*

3) *Will the study be cross-sectional or longitudinal or a mixture of both? What then will be the implications for the definition of the target population?*

Sampling

A sampling manual is critical. Once decisions have been taken on the international definition of the target population the sooner a sampling manual is produced the better. The details of the sampling manual written for this study are given in the Sampling chapter (Chapter 4).

The lessons learned from this study are that:

1. *The provision by NRCs of age/grade tables is highly desirable, if not essential.*
2. *It is desirable to describe the drawing of one national sample first and then to describe the drawing of domain samples;*
3. *The provision of tables for the different steps in sampling is essential.*
4. *More information should be given on how to determine Roh for sampling when no information exists within a country;*
5. *The minimum sampling accuracy required must be laid down before countries decide to participate. A simple equivalent sample of about 800 seems reasonable;*
6. *The Sampling Referee must explain to NRCs both in plenary session and individually how to complete the sample design tables;*
7. *Funding must be available for the Sampling Referee to visit at least half of the National Centers to work through the tables and ensure correct completion. It was in about half of the countries that NRCs had difficulties either because they had never undertaken this kind of task before or, interestingly enough, because they were "guilty" of secondary ignorance, i.e. they did not recognize when they did not know something. Sampling is clearly an area where "a little knowledge is a dangerous thing".*
8. *The Sampling Referee should spend all the time required at the ICC when the sampling weights and standard errors of sampling are calculated. In this study the ICC was in Hamburg and the sampling referee was in Australia. Fax and express mail were insufficient for rapid and detailed exchange of comments. Although NRCs should be encouraged to calculate sampling weights at their National Centers, weights for all countries should be calculated at the ICC.*

   It is a crucial element in studies such as these that the sampling is carefully conducted in each country. It is also clear that there are few persons in the world who are well versed in the theory and practice of survey sampling in education and who have international experience. This study was lucky to have such a person. But, it is also clear that it is an enormous amount of work and requires a lot of time. For a four year study, it is estimated that the Sampling Referee needs to devote at least one full year to the task (the writing of the manual, the teaching of the NRCs, visits to half of the National Centers, spending time at the ICC for the weighting exercise and for the calculation of the standard errors, and, finally, for the writing of the report).
9. *The funding for all of the time and activities of the Sampling Referee must be built into the study from the beginning.*

The composition and work of the Steering Committee

The composition of the Steering Committee for studies is important. Which skills are needed on the Steering Committee? Should the committee only steer or steer and work?

Skills required on Steering Committee. Apart from the skills required for sampling, the skills for the construction of background questionnaires, tests and attitude scales are needed. Data processing and analysis skills are essential. These involve having:

1) a good knowledge of the subject matter and the theory of trait structure in the subject matter;

2) international experience in content analysis of syllabi and textbooks;

3) experience in conducting item analysis and scaling (both classical and IRT);

4) international experience in attitudinal scaling and construct analysis.

International experience is important in terms of knowing about the avoidance of culturally biased items and of the general expected difficulty and facility levels of the items in different countries. It sometimes occurs that there are good national specialists but without international experience. This often results in increasing the number of members on the Steering Committee to include different types of countries. Such an expanded committee can become expensive. Reality often demands an uneasy compromise.

Indeed, there is much to be said for having a small core group that is the Steering Committee and others could be coopted as specific skills are required. Much work falls onto the chairperson of the Steering Committee. This person should have a good grasp of all facets of survey research (test construction and scaling, attitude scaling, questionnaire construction, and various forms of multivariate analyses) as well as being au fait with the subject matter and being a good writer. Such persons are hard to find. Above all, however, it is the research skills that are important. Other people can be co-opted for the subject matter expertise when the above paragon cannot be found. Each member should have his/her own expertise as well as general skills but should also recognize greater competence in others for particular skills.

The national committees at each National Center should have a similar gamut of skills and experience at the national level as the international committee at the international level.

Steering versus steering and working. Although the primary purpose of the Steering Committee is to work with the ICC in determining the conduct of a study and the details of each step of the study it is quite clear than an enormous amount of the work also has to be undertaken by the members of the Steering Committee. This involves not only design questions but also the specification of content analyses for the NRCs, the selection and editing of items, and the finalization of all instruments. Although meetings of the Steering Committee are held regularly, the members must each undertake a great deal of work between meetings. During Steering Committee meetings the detailed work is reviewed and final products are agreed upon.

Needless to say, there is much time that has to be invested. There is a limit to the amount of time an individual is prepared to invest without remuneration. As already mentioned, no member of the Steering Committee in this study received any payment. Most of the member of the committee were university professors. With universities tightening up their regulations about working outside the university it will become incumbent on future international studies of this kind to ensure that funding is available to cover the time spent by members of the Steering Committee.

As already mentioned, one member of the committee acquired other commitments such that he was unable to undertake the writing expected of him. His very late withdrawal created a problem that was only partially overcome. It can also occur that skills needed at one point in the life of a study are no longer required at another point in the study. There is little sense for a member to attend a meeting when his or her skills are no longer required.

It is suggested that very careful thought should be given to the composition of a core Steering Committee and that other members are coopted for only short periods of time.

Relationship of the International Steering Committee to the National Committee. It is important to delineate from the outset the skills required on a national committee and also the tasks that the national committee will have to undertake. In this study the tasks of the National Committee included:

1) definition of reading literacy domains (by interaction with the Steering Committee)
2) supplying items
3) translating items and adapting them to local conditions
4) planning and undertaking trial testing items
5) assisting with the classification of items
6) undertaking national panel ratings for standard setting
7) planning and selecting national option items

Such tasks (which may differ slightly from study to study) must be specified to the NRCs before they select their National Committees.

Test Construction

Good test construction always takes a long time. The details of the conceptual thinking and the various steps taken in the test construction are described in Chapter 3.

As already mentioned in Section 1 of this chapter, there were the normal problems of defining the criterion measures where there would appear to be differences between what the reading specialists regard as different but important aspects of reading and what the those with empirical approaches to latent traits are prepared to recognize as different aspects. The NRCs insisted that three domain scores be produced and that there would be no total score. However, some of the NRCs who at the beginning of the study did not want a total score did so by the end of the study. Furthermore, there was a demand for a total score by international agencies such as UNESCO and OECD. This is an issue to which much more time should be devoted at the beginning of a study. There was also the problem of determining the mode by which Reading would be measured. However, when testing hundreds of thousands of children and with limited financial resources there was not a great deal of choice. The final decision taken was to use predominantly multiple choice items, about 20 completion items, and a few open-ended questions. However, more open-ended questions would have been desirable. All countries should submit items. Only about 20 countries out of 30 submitted items. The final items for pilot testing were selected and edited in February, 1989. However, one country which had many item to submit could not submit items because of governmental bureaucratic constraints until April. Strict deadlines should be adhered to.

There was much discussion about the form of scaling to be used. As will be seen in Chapter 3, both classical and Rasch item analyses were used as a basis for item and passage selection for the final tests. Even after the main data collection, both types of item analysis were run again. Very few items were dropped from the domain scales. These were communicated to the NRCs. Following these decisions in January 1992 several questions arose about the way of handling 'not reached' items in the calculation of scores. There was much debate about these matters and the final problem was not resolved until early April 1992.

Indeed the Steering Committee had unanimously agreed in January 1992 that the estimation of scores would be undertaken for those students attending one testing session but not the other session, for items where errors had occurred nationally (either translation or printing errors), and for "not reached" items. Scores were calculated for each student for each reading domain based on these assumptions. Three weeks later the Chairperson of the Steering Committee unilaterally decided that he was not prepared to work with scores involving the estimation for "not reached" items. It was recognized from the outset that each form of scoring was fallible. Nevertheless, to allow the study to proceed, a decision was taken to recalculate the scores omitting the estimation of "not reached". This took time and caused a delay of several weeks in the production of the first booklet. Clearly, it would be desirable to have the types of scoring and scaling to be used for tests decided at the very beginning of the study.

The lessons learned were:

1. *Have the key members responsible for the test construction work full-time during the item preparation and item selection periods.*
2. *Have all national centers submit items but be strict about the deadlines for submission.*
3. *Have more open-ended questions.*
4. *Decide on scoring and scaling procedures near the very beginning of the study.*
5. *Decide on domain and/or total scores carefully.*
6. *Foresee and take care about the balance between reflecting curriculum content and the production of scales.*
7. *Meet more frequently and for longer periods with the NRCs.*
8. *Have more visits to National Centers.*

Construction of Questionnaires

There is a developed and known set of technical procedures for test and attitude scale construction. It is often the case that those who believe that they can use their mother tongue well also believe that they can "write" questionnaires. It is also often the case that a great deal of time is devoted to the test and attitude scale construction and that the questionnaires are 'constructed' at the last minute.

Typically in these kinds of cross-sectional survey studies there are background questionnaires for pupils, for teacher, and for school principals (or an equivalent in primary schools in Switzerland which do not have school principals!) as well as a National Case Study questionnaire.

The background data to be collected are a function of the questions posed by the study. It is essential for a study of this kind to have the general aims, the specific aims and, as much as possible, the operationalized aims of the study written down at the outset. The operationalized aims should include dummy tables, i.e. the tables which will appear in the final publications but without the data filled in. By doing this, it becomes clear which variables will be used in which way. For example, it will be clear as to which variables will appear for what type of descriptive (univariate) purposes only and which will also be used in a construct or some form of multivariate analysis.

At this point a number of specific 'bits' of information that have to be collected can be listed. For each 'bit' a question or set of questions can be written. It is then extremely useful to *write* down both a theoretical justification for collecting that piece of information and a short description of how it will be used in the data analysis.

The ICC did, in fact, do this for examples of the three questionnaires before the first NRC meeting. At the NRC meeting a committee was formed to "construct" the questionnaires. This was considered important for two main reasons. The first was that the collective wisdom of the NRCs

In some cases NRCs had not prepared their national versions of the instruments with great care and had the errors which were caught during these visits not been removed there would have been many more difficulties experienced at the cleaning stage at the ICC.

An NRC meeting typically lasts five or six days. Most of the time at these meetings is spent on test construction and to some extent on background questionnaires and attitude scales. Typically, subgroups met for each of these instruments in parallel but NRCs tended to be much more interested in the tests than in the questionnaires. Little time is left for procedures. Given that the cost of having more NRC meetings is high, serious consideration must be given to increasing the number of days for an NRC meeting.

The lessons learned were:

1. *Take great care in the writing of manuals*
2. *Have a specially written data entry program*
3. *Spend more time at NRC meetings on procedures even if this means lengthening the NRC meeting from 5 or 6 days to 8 or 9 days*
4. *Have the ICC data processing staff visit the National Centers to work with the National Center data enterers and others.*

Revision of instruments, manuals, entry program and codebooks

The proposed final instruments were sent out in early June 1991 to all NRCs. Their task was to review them with their National Committees and come in early July to the NRC meeting in Frascati, Italy, to make final adjustments to the instruments. This time was very short. For those NRCs who had to have changes to test items and background questionnaire items translated in order to discuss them with their national committees, time was extremely short – too short.

Nevertheless, the NRC meeting did take place but, as already stated, time there was also short and a disproportionate amount of time was devoted to the Reading Tests and too little time to procedures and data entry.

The revision of the manuals, entry programs, and code books did not present a problem.

The lesson learned was:

Allow more time between the ICC production of proposed final instruments and the NRC meeting when all the instruments and procedures are actually finalized.

Main testing

There was much faxing between NRCs and the ICC as the NRCs were preparing their instruments. In many cases, this arose from not having read Manual I carefully (see the point above about more time at NRC meetings for dealing with procedures).

In principle, no NRC was meant to draw the probability sample of schools or classes until his/her sampling plan had been formally agreed to by the Sampling Referee. This was adhered to. However, it did occur that one or two countries failed to reach an 80 percent response rate. It is to be noted that the USA achieved a much better response rate in this study than it had achieved in previous IEA studies in which it had participated. At the same time, it was only just over 80 percent whereas most countries had response rates of over 95 percent.

All NRCs, with one exception, handed in their data sets, together with the accompanying documentation, on time. The one exception – already noted – was due to the illness of the NRC.

At the NRC meeting just before the main testing the NRCs decided that one further component would be added to the study for Population B only. This was the ratings by national panels of each passage and item in the tests in terms of the desirability of its being mastered. A manual was produced by one of the US National Center staff together with a computer program. Neither of these were piloted. The upshot was that only twelve countries undertook the exercise, the data came in very late and problems were experienced in processing the data. It was only at the last minute that the two cutoff points for nine of the countries were produced. The moral of the story is that it is extremely unwise to add components to a study without their having been trial tested.

The lessons learned were:

1) Allow more time for NRCs to prepare final national versions of all instruments
2) Never add a component to a study without that component having been piloted.

Main testing Data processing

As will be seen from Chapter 6, the data processing was well planned. The data cleaning had been well prepared. Nevertheless, the cleaning did involve contacting NRCs on many occasions to check on particular data. It did occur that some NRCs had not undertaken data correction at the National Center before sending the data files to the ICC. It also occurred that some NRCs *believed* that no answer (i.e. missing data) really indicated "No" as a reply to some variables but had no real evidence for this. Even as late as February 1992 some NRCs were requesting changes to be made from missing to "No". At the same time NRCs were changing the valid ranges for open-ended questions as well as asking for the removal or addition of specific students/schools.

As will be seen from the Sampling Chapter there were different sampling designs requiring different weighting formulae. Although contact was maintained with the Sampling Referee it would have been desirable to have had the Sampling Referee at the ICC as stratum/school weights were being calculated and recalculated.

Each booklet had a set of data processing specifications. These were to be undertaken at the ICC and then sent to the author. However, for the first booklet, a decision was taken to take the data on optical disk to the New Zealand National Center (where the author would also be) in order to conduct the specified analyses and then conduct further analyses that the author might request. Although the computing "environments" were said to be the same, they were not so, and this caused some delay. The cleaning, merging and weighting had also taken some weeks longer than expected. Finally, the Rasch statistics programs were also some weeks late in being produced. The problem of how to deal with "Not Reached" items has already been mentioned.

The multivariate analyses on the pooled data set were eventually undertaken in Sweden by Ingrid Munck and Ingvar Lundberg using LISREL. At the same time, a series of Partial Least Squares (PLS) path analyses were undertaken at the ICC in Hamburg. These analyses were used for certain chapters in the international report. These analyses were run in a short time. It would have been desirable to have had much more time and as a consequence, many analyses had to be left for other researchers to undertake as secondary analyses.

As with all data analyses in a study with strict time limits, there is never enough time to complete all of the analyses originally planned.

To this end a set of the clean weighted SAS and raw data files were delivered to the IEA HQ in the Hague. It is the task of the HQ to organize a data archive to be used by other researchers for secondary analyses.

The specific lessons were:

1. *Work out in detail the organization and procedures in advance.*
2. *Plan the data processing in minute detail;*
3. *Never attempt to take data from one computer environment to another without having tried it out first;*
4. *A high performance PC network with multitasking operation system seems to be an adequate solution for this kind of project. Indeed it is several times faster than many mainframe systems.*

The writing of the booklets

The chairperson of the Steering Committee initially designated authors for each booklet and each chapter of the international report. Outlines and data processing specifications were written and reviewed by the Steering/Technical Committee and NRCs.

Each author received the printout for his analyses and then proceeded to write his/her report. It is, of course, a well known experience that more and more data analyses are required by authors as they work. Where possible further analyses were undertaken and sent to them. However, the major lesson learned was that authors should spend at least two sets of two weeks at the ICC when the analyses and writing up are being done.

The ICC provided style guidelines for the writing of the booklets and major reports to authors. The authors were requested to use word processing programs which could be easily converted to the Macintosh Apple machines used in the ICC's administration unit. The authors sent the diskettes to the ICC. In some cases the guidelines were not followed. In other cases the program used to produce graphs were such that the graphs could not readily be used at the ICC. In other cases, numerical values used by the author were found to be incorrect at the ICC. All of these points took more time than planned to correct at the ICC.

The booklets were filmed and printed in Hamburg. The cost of the first booklet was twice that of the estimate.

The international report was typed at the ICC but was printed and published by Pergamon Press in Oxford, England.

This technical report prepared for duplication both at the ICC and at the IEA HQ in the Hague. Four hundred copies were produced which are held at the IEA HQ for use by those conducting secondary analyses and needing many of the technical points reported later in this report.

The lessons learned were:

1. *As much as possible arrange (financially and logistically) that the authors of booklets, books and the technical report spend time separately and together at the ICC for checking values and writing the reports.*
2. *Although fax and e-mail help, communication by these modes can never replace being at the ICC.*
3. *Ensure that the guidelines for format and style are "taught" to the authors.*
4. *Ensure that the graph programs used by the authors can be readily used at the ICC.*
5. *Arrange for one sole distributor of booklets.*
6. *Use professional advice for the production of press conferences.*

Comments on the ICC (Administration and data processing)

Full-time vs. Part-time personnel. The ICC relied on one full-time assistant coordinator and one full-time data processing manager both of whom were assisted by part-time staff. It was only in the last year that there was a second full-time person in the data processing unit.

It was an error to rely so heavily on part-time personnel since part-time persons tend not to be fully committed and tend not to be acquainted with all of the interrelated parts of the study. It would have been wise to have employed mostly full-time staff. This was not done in this case because of the incertitude of the funding and because every effort had to be made to stretch the money as far as possible.

Rooms and equipment. It is important to have sufficient rooms to house the persons and equipment. This study was based at Hamburg University in the Institute of Comparative Education which, in turn, was part of the Faculty of Education. After great effort two rooms were made available. These measured 6.5 x 3.5 meters each. In each of these rooms there were always at least three, and often five people working. In addition several PCs and one laser printer had to be accommodated and in one of the rooms a fast printer was running for much of the time. After two years, a third room measuring 6.5 x 2.1 meters was made available. The fast (and noisy) printer together with two PCs and two persons were moved to that room. But even this final situation was far from adequate. The university was able to supply only one IBM AT machine. All other machines had to be purchased and maintained by the study. All software had to be acquired by the study.

Personnel. It is desirable to have an international staff at an ICC. Although the working language of IEA is English it is very useful to have personnel who either have very good English as a second language or who have English as a mother tongue and speak one other language fluently. With their experience of having to communicate in a foreign language they are sensitive to the difficulty that some NRCs experience when trying to communicate in English and they know how to express themselves clearly and unambiguously in English. At the same time, researchers from different school systems can spot inadequacies of questions and problems and can suggest procedures for overcoming problems.

The personnel must be able to undertake both word processing and data processing. This may sound easy but it proved not to be the case. The mentality and skills required for data cleaning and data analyses are somewhat different. In general university students do not have the skills required. Indeed, even when they do have the skills there is usually industry nearby which attracts those students with good skills – even for part-time work while they are studying. Although the RL study had an outstanding data manager he could only allocate relatively routine tasks to others. It was difficult to find another person at the university who was of a sufficiently high standard – *and available* – to help him.

The accounts also had to be kept. In the case of this study it has been seen in Section 1 of this Chapter that the funding came from outside of Germany, the German governmental authorities having declared that they were constitutionally unable to grant money to nongovernmental organizations such as IEA. The Hamburg University accounts department works in German and according to German government rules. It was, thus necessary to create a nonprofit company in order to receive money and employ staff. In turn this meant that the ICC had to keep the accounts and deliver accounts and reports not only to the grant giving agencies but also to IEA and the German tax authorities.

Should ICC be located at a University? The RL study's experience is that the ICC should not be housed at a University. The experiences in Hamburg were unfortunate but from experience of other universities there are always the following problems:

- Lack of rooms for the staff
- Lack of meeting rooms when term is in session
- Acquiring and keeping competent staff for both data processing and data analysis
- Keeping the accounts

It is suggested that each IEA study has an administrative ICC but that the data processing (organization of codebooks, data entry program, cleaning, weighting and merging) for *all* projects be centered at the IEA HQ. This staff would also produce the univariate statistics and the specified correlation/covariance matrices of specified variables for others to work on. This assumes that the IEA HQ can acquire a sufficiently competent staff to undertake this work. This, in turn, assumes sufficient secure and continuous funding. If the IEA HQ cannot do this, then the work should be contracted out to a private firm with proven experience in international work of this kind.

Ideal staffing. The following are seen as the ideal attributes of persons employed at the ICC:

i) Coordinator: This person should have general educational survey research skills and proven organizational skills. The general research skills consist of good experience in research design, sampling, test/questionnaire/attitude scale construction, data collection, data cleaning, data analysis, and writing up skills.

The organizational skills consist of proven experience in making realistic timetables, running primarily a secretarial and accounting office, organizing international meetings and knowing when and how to kick, cajole, or charm Steering Committee members, NRCs, and others in English and, where appropriate, in other languages.

ii) Data processing staff: In any study, the amount of time needed for the data processing activities will vary at different points in time. Given this situation, there should be a good full-time core staff but others will be needed at particular times. The types of skills required are somewhat different for data management and for data analysis. Both sets of skills are required throughout a study. It is easy to underestimate the amount of time needed for the practical steps of processing and analysis. At least one member of the data processing staff *must* have good experience in knowing what is realistic.

An ideal staff for a study such as this one should consist of:

Data Manager (full-time). This person must have good research skills, a statistical background, and experience in the implementation of both routine and complicated procedures. Furthermore, he or she person must have data management skills to plan and supervise the work of the team. It is a bonus if this person also has good programming skills.

Good programmer/Statistician (full-time). This person should possess at least one data management language and be fully conversant with one good statistical package. These skills should be at a professional level.

One Data Processor (full-time). This person must be able to undertake routine work.

Data Processors (part-time). These persons should be able to undertake routine work as it is required.

Computer environment. The computer environment should be interactive and flexible in its configuration. It was advantageous to use a PC network in this study because the performance times reached, or surpassed, those of many mainframe machines. At the same time, it allowed the user to see the data immediately on screen and also allowed interaction so that the user could undertake modifications immediately.

It is important to have an operating system allowing multiple jobs and which is also independent of the software being used. To this end, the operating system was changed from DOS to OS/2.

It should be noted that it is easy to underestimate the amount of hard disk space required. It is desirable to estimate this correctly from the beginning of the study.

The major lessons learned for the ICC were:

1. Have a few full-time persons rather than several part-time; have one person designated for communication with all NRCs and one for managing all publications;
2. Have one full-time data manager and two full-time assistants;
3. Ensure adequate resources for the ICC before starting (this may be utopian , but there are limits!)
4. Ensure adequate accommodation;
5. Have the Steering Committee Chairperson and the ICC with easy traveling distance of each other;
6. IEA HQ should seriously reconsider setting up a central unit for the clearing, merging and weighting of data sets for all projects in one central place;
7. Have more visits of key ICC persons to National Centers at key points in the study to help and check on data entry, sampling, and instrument matters;
8. Have NRCs visit the ICC;
9. Have the ICC build in training for some NRCs as required.

Major Lessons Learned: a Brief Summary

The above has given details of lessons learned. The major lessons are listed below. These are based on what were perceived to be the strengths and the weaknesses of practices and procedures in the Reading Literacy Study:

1. *IEA studies should be collaborative in nature involving NRCs. All have inputs to make. They should never be treated as mere "gatherers" of data.*
2. *IEA should involve selected NRCs in ongoing projects in the writing of the proposals for new studies. In this connection, the IEA HQ should always ensure that past and existing methodologies from various projects are reviewed continuously and that all projects are aware of them; all NRCs should be informed about all ongoing projects.*
3. *Any proposal should include:*
   - *a clear set of general and specific aims with priorities set on them;*
   - *a thorough review of pertinent earlier international and national studies and relevant theories;*
   - *clear definitions of desired, excluded, and defined target populations; target populations should be defined so that age and grade effects can be separated;*
   - *minimum sampling accuracy requirements;*
   - *a measurement model (i.e. for assessing reliability) for background questionnaires;*
   - *descriptions of outcome measures (cognitive and attitudinal);*
   - *detailed definitions of scoring and scaling procedures to be used;*
   - *minimum participation rules;*
   - *guidelines of what an NRC is expected to know and do;*
   - *guidelines of work expected from national committees;*

4. *All decisions about steps and procedures in a study should be made collaborationally among NRCs, the ICC, and the Steering Committee. This involves clarification of points in the proposal used for raising funding and changes being made, where appropriate.*
5. *NRC meetings (especially earlier ones) should be more frequent than in the RL study and longer than one week to ensure that all points are thoroughly discussed and agreed on.*
6. *More care should be taken with the construction of the background questionnaires; it would be desirable to have:*
   a) *a core set of questions where comparability is established,*
   b) *international options, and,*
   c) *national options.*
7. *More care should be taken at early stages in a project about which scores will be produced (e.g. only domain or domain and total) with a full knowledge of the consequences of the decision. This also applies to rotation schemes and the time allowed for testing.*
8. *All projects should include two pilot studies: the first for test items and questionnaire questions; and, the second for final revised items and questions (and instruments) as well as procedures.*
9) *NRCs should be more involved in the international analyses; this requires adequate funding;*
10. *A relational link should be established, from the beginning of any study, between the international and national components of a study, for this purpose, more training will be needed in:*
    - *research methods and analyses;*
    - *research management;*
    - *data collection and operational issues;*
    - *computer use.*
11. *Every effort should be made for the ICC to support national analyses and National Centers where support is required.*
12. *The ICC should have full-time rather than part-time personnel and adequate accommodation; the Steering Committee chairperson should be located near the ICC.*
13. *As much funding as possible should be assured before a study begins. This should cover the costs for at least the first two thirds of the study.*

Chapter 3

The Measurement of Reading Literacy: How the International Tests of Literacy Were Developed

Warwick B. Elley

THE MEASUREMENT OF READING: HOW THE INTERNATIONAL TESTS OF LITERACY WERE DEVELOPED

Warwick B Elley
University of Canterbury, New Zealand

(a) INTRODUCTION

In drawing up guidelines for an internationally valid test of Reading-Literacy, which would be acceptable in more than 30 countries, the Steering Committee faced a number of conceptual and technical challenges. It was necessary first to reach agreement on such matters as:

(1) The definition of Reading-Literacy which would guide the selection of literacy tasks.

(2) The test specifications, outlining the major domains to be sampled and the weightings to be applied at each level.

(3) The item formats to use in developing appropriate literacy tasks - or items.

(4) The best procedures for writing, editing and pilot testing these literacy tasks.

Reaching consensus on these issues occupied much of the time and effort of the National Research Co-ordinators at their first two meetings in 1988 and 1989. A commitment to democratic procedures and maximum participation of member countries at the outset made for lengthy discussions. However, the design and instruments were eventually agreed to and completed within the deadlines set for Phase I of the project - up to the end of July 1990.

(b) THE CHALLENGE

Many would claim that the literacy act is so embedded in particular cultural contexts that it is not possible to develop a single measure of reading-literacy, that could be considered valid for all participating countries. (Levine, 1986).

It is certainly true that there are many reading materials and genres which are more familiar and important in some cultures than others. In this connection research has clearly shown that a reader's background knowledge (or schemata) is a major influence on reading comprehension (Freebody and Anderson 1983; Steffenson, Joag-dev and Anderson, 1979) and this does vary considerably across cultures, making cultural bias difficult to avoid.

Translation of test contents across languages introduces another problem. For instance, the extent of consistency between sound and symbol in reading materials is a potentially contaminating variable in a test translated into other languages. Some languages, such as Spanish and Finnish, for instance, are phonetically quite regular. Others, like English, are

notoriously irregular. The Chinese language is not alphabetic, while the Thai language is characterized by many silent syllables. Differences in syntax between languages are another obstacle to producing culturally equivalent tests, especially at the younger age levels.

In addition, sociolinguists would argue that the larger social context within which reading is practised in a community would undermine the international validity of the tests. The availability of books, magazines and newspapers, the censorship regulations, the style and the importance of testing in the school system, the relative status of languages in which people practise literacy - all of these factors, and others, influence the significance of literacy test scores, and throw doubt on the viability of preparing a single set of tests which could be regarded as a fair basis for assessing reading literacy across countries.

These kinds of challenges and constraints have been faced by many who have attempted international comparisons of achievement in the past. The procedures developed by IEA over the past two decades, for developing tests in any subject, go some distance towards alleviating some of these problems. Thus, tests are constructed and edited in democratic fashion, by an international committee, in such a way that all participants have an opportunity to submit passages and items, and to comment on, or veto, those materials selected. In this way passages which are considered unsuitable by virtue of an unfamiliar topic or genre are eliminated. Culturally appropriate modifications to names, places and measurement units, for instance, are made to others. These strategies were adopted by the Steering Committee at the outset, thus alleviating the first problem raised above.

In addition a policy was proposed for pilot testing all materials in all participating countries, so that all NRC's would have a chance to identify problems for their students before the final selection of items took place. This policy was implemented to good advantage in the current project, and most countries undertook a sizeable pilot study in their schools. Moreover, the Steering Committee was in a position to study patterns of difficulty across cultures, and to eliminate offending passages and items at this stage.

The specific problems raised above should be addressed more thoroughly, however. The fact that unfamiliar topics and genres were to be eliminated by any of 32 country representatives raised the consequential criticism that so much would have to be dropped that the final result would be satisfactory to no single country. The Committee met this criticism by planning an unusually long test at each age level, and by subsequently comparing the diversity of topics and genres of the finished products with a sample of reading tests drawn from several participating countries. This exercise, described later in this chapter, showed that there was indeed much in common in the school and societal experiences of children in all countries, and that no reading test, in any one country showed a greater diversity of materials than the IEA tests developed. It is a fact that a school timetable, for instance, is common in all countries, that it is interpreted in a similar way, and that a comprehension question about which subject students were studying on Thursday morning, first period, is equally sensible, and difficult in all countries. Likewise stories about a boy who wanted a pet dog, or about a frail old man whose family were unkind to him, or about conflicts between animals, all these proved to be universal in their appeal and their acceptability.

The expectation amongst students that specific kinds of questions might be asked about the test passages may be found to vary across cultures, but such problems can be identified in the editing and the pilot testing phases. As for the claim that no reading test can be equally fair because of variations in students' predominant background knowledge between

countries, the same case can be made against reading tests within countries. Many topics are avoided because they favour one sex, or rural children, or middle class children, or those with particular cultural experiences. The test constructor should be alert to these matters at all stages.

The criticism that languages are not equally difficult for reading purposes is based more on logical deduction than empirical research, and the possibility of clarifying this matter in the present study was not overlooked. The fact that languages differ in such characteristics as their flexibility of word order and in the number of synonyms available for key words and phrases did influence the test construction process. Thus cloze tests and vocabulary items were excluded from the IEA tests for these reasons, and a number of potentially useful measures of "enabling skills" were dropped once NRC's began writing them. Whether narrative and expository passages, with complex prose and with following questions, could be translated fairly was an empirical question, on which previous researchers had expressed optimism. (Thorndike, 1973; McCauley and Colberg, 1983). After his study of reading comprehension in 15 countries, Thorndike claimed that "an item is an item is an item, whether in English, French, German, Swedish, Polish, Finnish or Serbo-Croatian. The difficulty lies, it appears, not in the verbiage or syntax of a specific language, but in the structure of the ideas with which the item deals." (Thorndike 1973-4, p. 145). It is the art of the test constructor, in the first place, and the translator in the second place, to ensure that the verbiage and syntax do not intrude. So precautions had to be taken, at all stages, to minimise, and check for cases of difficulty which could be attributed to linguistic rather than conceptual factors. The Rasch score analyses generated from the pilot studies provided an essential check on the degree of success that attended these efforts. These are described later in this chapter.

To address the sociological arguments, that literacy practices are not comparable across cultures, the Steering Committee set out to clarify the context within which children in each country engaged in literacy tasks by means of a series of questionnaires to NRC's, school principals, teachers, and
pupils. Information was sought about the availability of resources, the occasions when students read, the popular reading materials, the styles of teaching, and the like. No doubt, differences do remain, but a survey of the kind envisaged was felt to have potential for identifying both the incidence and the significance of these varying practices.

The style of testing within each school system was another key subject which occupied much time in discussion and research, and the relative importance of such variables as test format and frequency of testing, in relation to the influence of substantive comprehension difficulties in the text, were explored during the study.

Finally, there are critics who claim that the artificiality of any formal standardized testing process is sufficiently unlike "real life" literacy tasks to make the exercise worthless. To those critics we can point out that policy makers in all countries call for the kind of information that such tests produce. And for good reason. There is plenty of factual evidence to show that results from individual reading tests, miscue analyses and oral retelling exercises correlate well with the results from well prepared standardized group tests. There may be some for whom the correlations are less than perfect, but if cultural bias and the logistics of mass testing are problems with standardized tests, they would be even greater with whatever passes for "real life" literacy tasks in most countries. The onus is surely on these critics to demonstrate that people who perform well on standard group tests of reading are not, in fact, competent readers in genuine reading tasks and vice-versa.

In sum, the obstacles faced by the Steering Committee in developing an international literacy test were many and complex, but the potential value of the exercise was widely agreed, by more than 30 country representatives, to be worth the effort of overcoming them. It is worth stating that each generation of researchers stands to benefit from the experience of the last. It is true that satisfactory models of reading comprehension are still being debated and developed by researchers, but this task of developing a common test for 32 countries did reveal, fortunately, a high level of consensus amongst NRC's on procedures and suitable contents.

(c) DEFINITIONS

Quality measurement requires, first and foremost, clear definition of concepts. The concept of literacy has a long and colourful history (see Graff 1981, Levine 1986, Guthrie and Kirsch 1984, Clifford 1984, Wagner 1989) and there is much debate over the meaning and importance of such items as basic literacy, functional literacy, cultural literacy, workplace literacy, and the like. The Steering Committee, mindful of the constraints imposed by the need to include thousands of participants, in many countries, using a multiplicity of languages, and operating in various cultural traditions, searched for a working concept that would be widely acceptable and would provide guidelines for a set of tests which would measure central aspects of literacy.

At the outset, writing ability was excluded from the definition. A recent project of the IEA had focussed on writing (Gorman, Purves and Degenhart, 1988) and the problems entailed in making international comparisons in this curriculum area had been fully explored.

The notion of functional literacy, with its connotations of being able to use one's literacy skill to function effectively within one's own society was popular in the early discussions, but some NRCs wanted to extend the notion beyond the basic levels needed for survival, to include higher-level thinking and the reading of good literature, for example.

In an attempt to embrace both of these viewpoints, and to provide some link with the earlier IEA study of reading (Thorndike 1973), the Steering Committee proposed the following definition, which was subsequently accepted by NRCs:

"Reading-Literacy is the ability to understand and use those written language forms that are required by society and/or valued by the individual."

The category of language forms required by society refers to those kinds of literacy tasks which are needed to cope with the business of living in an organized society - reading notices, directions, maps, graphs, and government circulars, to name a few; the category of language forms valued by the individual allows for the inclusion of leisure reading - narrative prose, or popular magazine articles. These may be valued by individual readers, but are less often required for survival in a society. The Committee proposed that the extent of value seen in the test materials for individuals in a country could be rated by NRC's, during the test development phase.

Participating NRCs were agreed that both "understanding" and "use" of these written forms were important in the definition, and various attempts were made to incorporate both aspects, although the constraints of mass testing and the assessment traditions existing in most countries inevitably meant that the major operational focus was placed on "understanding".

In the event, it was found possible to measure the students' ability to follow instructions correctly in a few tasks.

(d) DECISIONS ON WHAT TO MEASURE

During its earliest deliberations the Steering Committee proposed that three processing dimensions of reading-literacy should be tested - viz: decoding, locating information and comprehension. Several kinds of text materials - words, phrases, sentences, extended prose, documents, environmental print - were envisaged, to allow for testing these four processes.

Once these various proposals were discussed with NRCs from many countries, however, it became obvious that difficulties of translation and a variety of time and cultural constraints required a number of revisions in the initial plans. "Decoding" for instance was felt to be difficult to assess in non-alphabetic languages, and was replaced with an array of "Initial Enabling Tasks" such as Visual Discrimination, Segmentation, Word Order, Logical Inference, Cohesive Ties and Word-Picture Matching.

Subsequently, the testing of single words and short phrases in notices, was found impracticable because of major problems in translation and differences in language features, and most of the Initial Enabling Tasks were eventually dropped for such reasons. In addition, it was agreed that mastery of many of these tasks was complete at age 9, and that whatever variability was found to exist was often as much an effect of differing amounts of practice in reading, as a causal factor. Only the word-picture matching task survived the gauntlet of the NRC's views, once they had attempted item writing on these various enabling tasks. Thus, attention focussed on a classification scheme for the comprehension of continuous prose.

Debate on the underlying dimensions of reading comprehension has a long history. Davis (1944), for instance, starting with eight logically distinguishable reading sub-skills, surveyed 988 American University students, and after a factor analytic study of the results, reduced the eight skills to four: identifying word meanings, drawing inferences, identifying writers' technique and mood, and finding answers to questions. Thorndike (1971) reanalysed his data and reduced the four factors to three: of which the first (vocabulary) was the most important, and the other two difficult to label. A further re-analysis by Spearritt (1972) found that all factors were highly correlated (0.75 to 0.93), and that the vocabulary factor was clearly the largest.

A large British study (Lunzer and Gardner 1979) of 10-11 year olds using tests of eight postulated sub-skills yielded only a single general factor which accounted for 81% of the common variance, while another study of New Zealand 9-10 year olds revealed a similar finding (Elley & Reid 1969). A review of 12 other attempts to unravel this problem reported by Farr and Carey (1986) showed that six studies produced only a single general factor, while the other sub-skills identified differed considerably from one study to the next.

While the findings are not conclusive, there is a converging consensus amongst researchers that reading ability, interpreted as the ability to construct meaning from continuous prose, is, for students between 9 and 20 years, a unitary ability. While the sub-skills may be isolated logically for the purpose of item writing or teaching, the evidence points to the fact that students' ability transfers readily from one such skill to another. Certainly there are differences between students in background knowledge (see Anderson & Pearson 1984)

and interest (Asher 1980), but these differences probably cut across the hypothesized sub-skills in such a way as to reduce their identity in any test of their presence. Those studies which appear to identify separate reading skills are rarely replicated with similar results, suggesting that the pattern is very dependent on some characteristic of the passages used in testing. After a thorough analysis of these issues, Lunzer and Gardner are drawn to the view that:

> "It is not possible to provide a reliable differential test of reading comprehension skills" (Lunzer and Gardner, 1979, p.70)

Recent American studies point in the same direction (Zwick, 1987; Carroll, 1987; Stenner *et al*, 1988), although there has been debate on the finer points of the single proficiency scale of reading that NAEP researchers have used to report their reading survey results.

Faced with such findings, the Steering Committee accepted the proposal that a more defensible classification of prose reading tasks centred on the genre they represented (Bower 1980; Brooks & Warren, 1970). Students who habitually read narrative prose develop a conceptual framework or "story grammar," which enables them to read and recall narrative texts more easily (Stern & Glenn, 1979; Wilson & Anderson, 1986; Graesser *et al*, 1991), and thus to achieve higher scores on tests of fictional prose. Students who more often read expository prose, designed to explain something, are more likely to perform well on tests of such prose, and can be trained to do so. (Anderson & Armbruster, 1982; Dansevan, 1978; Bartlett, 1978; Weaver & Kintsch, 1991).

Brewer's logical classification (1980) into descriptive, narrative and expository prose has gained wide support amongst cognitive psychologists (see Barr, Kamil, Mosenthal and Pearson, 1991), and variants of it are widely used in reading test construction. (eg. NAEP in USA, APU in UK, Progressive Achievement Tests in New Zealand). After discussion, NRCs agreed to a division of prose passages into two basic categories, narrative and expository (embracing descriptive and persuasive genres).

Not all reading requires interaction with continuous prose passages. Real life literacy tasks, often require readers to carry out a physical task, to locate a fact or set of figures, to interpret a picture or graph, to skim a page and note its structure - tasks which do not require sustained language processing. Are these, too, part of the postulated indivisible unitary comprehension ability?

Research by Guthrie and Kirsch (1987) and the NAEP adult literacy study (Kirsch & Junglebut, 1986) suggest not. In both these surveys the correlations between comprehension of prose and the skill of locating information in charts, maps, directions and other documents was low, despite high test reliabilities. Apparently the cognitive processes required to search for and retrieve specific items of information from documents, tables and diagrams are sufficiently different from the more language-based processing required for working through a paragraph of text, with a view to remembering it, that the two abilities can be readily separated and independently reported. While the research is recent, it receives confirmation from the results of standardized tests of study skills. Reid *et al.* (1978), for instance, show correlations of 0.6 to 0.7 between Reading Comprehension and Locating Skills with maps, charts and diagrams, but correlations of 0.8 to 0.9 with Vocabulary Knowledge.

In the light of the conclusions drawn from such studies, the US NAEP exercises for recent national assessments of 9, 13 and 17 year-olds allowed for a category known as "document

literacy", along with the comprehension of literacy and informational prose. Research into the components of document literacy is a growing field (Guthrie, 1988; Kirsch and Mosenthal, 1988; Wright, 1988), but the research evidence on its separate identity and its frequency of usage in out-of-school reading was sufficiently strong to warrant a serious attempt to measure it, independently of reading comprehension. The tasks appeared to be different - one requiring the processing of large prose segments of continuous text and responding to questions about it; the other requiring a sizing up of structure, without complete reading, and using the structure to locate and process information within it.

It was also proposed at the first NRC meeting, that a cross-section of topic themes should be included, representing tasks that are likely to be encountered at Home, at School, in Society at large, and at Work. These were found to be helpful in ensuring a balance in two types of text materials - Documents and Expository Prose (but not for Narrative). Subsequently, the majority of the literacy tasks devised to tap Work contexts were vetoed as unsuitable for 9 year olds, and frequently also for 14 year olds. Thus, the Society and Work categories were eventually combined for both populations, and included in the Test Blueprint. As in the case of the mental processes being tested, there was too much subjective judgement entailed in deciding where to locate the various test items, once they had been prepared. Nevertheless, the cells of the Blueprint did provide a helpful guide for item writing, and for analysing and reporting results.

For the younger readers, in Population A, a Word Recognition test, measured by matching individual words with pictures, and administered under speeded conditions, was the only one of the beginning or enabling tasks retained for the pilot testing. This test was very easy, and was designed principally to establish whether weaknesses found in reading comprehension in a given group of students could be attributed to their inability to decode words of high familiarity. This Word RecognitionTest was not seen as a major domain in the test blueprint, and was not added to, or reported with the findings of the major factors. It was intended more as a covariate for the Population A results, and as an easier task which weaker readers may use to generate confidence for the tasks ahead.

The blueprint of the Reading-Literacy tests, agreed to by NRC's in July 1989 for the pilot tests conducted in 1989-90 is given below. The letters (A) and (B) refer to Populations A and B. The numbers in the cells represent the number of items pretested.

TABLE 1

TEST BLUEPRINT FOR PILOT TESTING (JULY 1989)

WORD RECOGNITION	NARRATIVE	EXPOSITORY	DOCUMENTS	THEMES
		(A) 20 (B) 20	(A) 30 (B) 20	Home
(A) 48 (B) -	(A) 100 (B) 85	(A) 45 (B) 45	(A) 30 (B) 55	School
		(A) 15 (B) 25	(A) 20 (B) 40	Society/ Work
TOTALS: (A) 48 (B) -	(A) 100 (B) 85	(A) 80 (B) 90	(A) 80 (B) 115	(A) 308 (B) 290

As it was intended to pilot test approximately three times as many items as were planned for the final tests, the numbers given in most of the cells for the pilot test are about three times greater than the target numbers for the final tests.

The final test blueprint is given in Table 7.

It is clear from the Blueprint that the School theme was accorded more weight than Home or Society/Work, although the latter was given more priority at Population B level.

As noted above, the major text domains, Narrative, Expository and Document were presumed to cover the major types and varieties of reading materials commonly encountered in reading in most countries. Poetry was excluded from the outset, as being largely untranslatable, and other genres such as persuasive and descriptive text, being less common, were subsumed under Expository. Although there was often debate on the issue of the independence of the three major domains, the majority of NRC's finally agreed that the three dimensions were conceptually separate, and an analysis of the kinds of questions typically set in these domains supported the distinction. Furthermore, results of the pilot tests subsequently provided some confirmation for the continued separation of the domains. The classification also coincided with that of other major test projects in literacy, and with the views of the Steering Committee members.

By way of definition, the Narrative domain was intended to include continuous text materials in which the writer's aim was to tell a story, whether fact or fiction. Narrative texts are normally designed to entertain or involve the reader emotionally, are written in the past tense, and usually have people or animals as their main theme. Common features of

narrative text are illustrated in the now-familiar concept of a story grammar - setting, theme, conflict and resolution. (Mandler and Johnson 1977, Thorndyke 1977, Kintsch 1977).

By contrast Expository texts are designed to describe or explain something; they may be written in the present or past, and the style is typically formal and impersonal, highlighting such features as definitions, causes, classifications, functions, contrasts and examples, rather than a moving plot with climax.

The third domain, Documents, is a relatively new category, popularised by Guthrie and Kirsch (1987). Documents refer to such things as forms, charts, labels, graphs, recipes, maps, directories and sets of instructions (Kirsch and Jungeblut, 1986). The reading tasks associated with Documents normally require the student to skim the materials to identify the structure, and use it to locate the required information. Thus, word for word reading is not needed, or desirable.

(e) DEVELOPMENT OF PASSAGES

After lengthy discussions by the Steering Committee and NRC's the following guidelines were proposed to assist test item writers in each country prepare suitable passages and items.

1 Quantity

Each NRC was invited to submit to the Steering Committee two items for each of six initial enabling tasks initially envisaged for Population A; two sets of comprehension items based on short prose passages for each Population - one narrative and one expository; two sets of items based on long passages and two sets of items based on documents for each Population. Bearing in mind the difficulties of undertaking translation, these requests were reduced somewhat for non-English speaking members. No suggestions were made as to the theme of the items submitted, but NRC's were asked to keep in mind a balance amongst items which might be encountered at home, at school, in society, and at work.

2 Criteria for Suitable Passages and Items

In addition to the usual basic principles about selection and preparation of suitable material for reading tests, it was felt that reading tests which were to be internationally valid would require an additional set of criteria for the selection or writing of suitable passages, and items. The following criteria were agreed to:

Suitable Passages

(i) They should be unseen by the students, i.e. they should not be part of a current school textbook, or basal reading series.

(ii) They should not date quickly, i.e. they could be used again in 10 years time.

(iii) They should be suitable for all countries, languages, ethnic groups, socioeconomic levels, sexes.

(iv) They should be well written and interesting.

(v) They should usually be self-contained, i.e. they could normally stand alone as a genuine reading experience in their own right.

(vi) They should not contain so much common knowledge that students could answer the questions without reading the passages.

(vii) They should range from simple (for weak 9-year-olds) to complex (for competent 14-year-olds). Some questions were planned to be common to both tests.

Suitable Items

(i) They should be multiple-choice or one word answers in most cases, but there would be provision for some extended response type. Some might require underlining, or filling in forms. Matching sets of items could be used occasionally.

(ii) Multiple-choice items in Narrative, Expository and Document cells should normally be 4-choice for Population A; 4 or 5-choice for Population B.

(iii) The wording of the items should be as simple as possible so that the difficulty of the question resides in the passage, not the question itself.

(iv) A short passage might have 3-6 questions in a set; a long passage might have 6-12 questions. Some of these would drop out after pilot testing.

(v) The locating information tasks should vary from simple one feature tasks where the child searches to locate a simple number or word, to more complex tasks in which several features (or organizing principles) are required to solve the problem. Some of these items may require manipulation of figures (numbers).

(vi) In testing comprehension, the following kinds of questions should be considered for inclusion:

(a) Literal questions, where the answer is explicitly stated in the text.

(b) Main idea questions, which determine whether students can distinguish between key ideas and minor details, or can recognize the summary of the main theme in a phrase or title.

(c) Prediction questions, which require students to infer likely outcomes.

(d) Cause-effect questions; if-then questions.

(e) Inference questions about the writer's aim, or viewpoint, in all or part of the passage.

(f) Inference questions about the motives of characters.

(g) Evaluation questions about the extent to which the writer, or certain characters, achieved their aims.

(h) Vocabulary questions, which detect students' ability to infer meaning from context.

(i) Application questions which show whether students can follow instructions, fill in a form, and do something to show they have really understood.

Further suggestions were made about standard item-writing practice, and several specimen items were provided in each category.

Screening of Items Submitted

Between November 1988 and March 1989, 20 countries submitted test items to the Steering Committee for consideration, and nearly 2000 were considered for each Population level. Many of these were considered to be unsuitable because they were taken from standardized tests still in use, or because they were of doubtful quality or judged inappropriate for use in all participating countries.

The Steering Committee screened and rated the items in two meetings held during March 1989 in Singapore, and in April in San Francisco. Many of the passages and items which showed promise were revised before and during these meetings. Additional passages were written or adapted by the Steering Committee where obvious gaps existed, while others were taken from the previous IEA Reading Study, so that a link with that study would be made possible.

At the time when the items were distributed to NRC's for review in April 1989, 18 of the 20 countries that had submitted test materials had at least one set included in the proposed item pool for the pilot test. At that stage, the original set of nearly 2000 at each Population level had been reduced to approximately 325 at each level, or 650 altogether.

Major changes were recommended by the Steering Committee in the initial enabling tasks, many of which had been found untranslatable by some NRC's, or too simple, or biased culturally, or of dubious validity for 9 year olds. Another category, that of locating information in prose passages, was also found to be unpopular with member countries, and very few items in the category of unfamiliar or fictitious vocabulary, to be inferred from context, were submitted. Both of these categories were finally dropped, by consensus of NRC's.

As a possible replacement for the initial enabling tasks, the Steering Committee proposed an innovative approach to measuring "Environmental Print". Stop signs, exit signs, and other notices such as are found in shopping centres and public places were devised, but these too were eventually excluded. The problems envisaged in translation, and the diversity of cultural traditions and expectations revealed during the international meetings inevitably meant that some items which would be acceptable and operational in a few countries had to be dropped following the veto of others. The main criteria used in excluding items at this stage, were those listed above for the item writing phase.

Editing of Items by Participating NRC's

In keeping with the principle of maximum participation in test development all NRC's received from ICC in April 1989, a complete set of the test materials proposed by the

Steering Committee. In consultation with their National Committees, the NRC's reviewed the items in preparation for the second NRC meeting to be held at the end of June 1989. Meanwhile, Steering Committee members undertook preliminary small-scale pilot studies on selected passages and item types in several countries, and were thus able to eliminate some of the faults in the items.

At the June 1989 meeting NRC's were furthermore asked to rate all remaining passages on a 3 point scale and to comment whenever they thought them unsuitable. While most passages were judged suitable by the majority of NRC's, a few were ruled out at this stage. For instance, some passages were found to be already familiar in certain countries, and others were felt to favour students in some countries at the expense of others. Most remaining items were subjected to considerable editing in the course of that NRC meeting.

Other relevant decisions made at the June 1989 meeting were that the proposed simple Word Matching test of 48 items should be included in the Population A pilot tests, and speeded to such an extent that students might reveal their level of automaticity of word recognition. The preferred format of most NRC's was multiple-choice, in which pupils had to select the one of four pictures which matched a familiar, high frequency word.

An alternative proposal, in which one picture was to be matched with one of four words, was judged unworkable, as there would be too much variation in word orthography and helpfulness of options across languages, after the items had been translated. In a speeded test - 90 seconds for 48 items - this variability could well be a significant source of difficulty.

To allow for enough easy comprehension items for weaker 9 year olds, a set of 20 3-choice items based on a simple story was adapted from the 1970 IEA Reading Speed Tests for inclusion in the pilot tests.

Another decision made by NRC's at this stage was the omission of all vocabulary items. Again the problems of variation in difficulty after translation were felt to be insuperable, even when the words were tested in context. An experiment with artificial words, trialled in Singapore, met with partial success, but was finally voted unsuitable for a valid test of literacy.

While test items of several formats had been requested from NRC's, most of those submitted were multiple-choice, and the majority of NRC's expressed their preference for this format in a large-scale survey. Where thousands of tests were to be scored in limited time, usually with scarce resources, the advantage of an objective format was appealing. However, a number of NRC's made a case for including some open-ended response items where students were required to generate their own answers. In addition, some 20% of the items in the pilot tests were of the completion or fill-in type, where students filled in a one or two-word answer, which was to be scored right or wrong.

After much debate, it was agreed to pilot test about 10 open-ended items, requiring 2-3 lines of writing, at each Population level, and to investigate the issue at greater length while the pilot studies were being conducted. These investigations are described below.

(f) PLANNING OF PILOT TESTS

In preparation for the pilot testing phase, a number of decisions were made about the format and sequence of the tests. It was agreed to pilot test about three times as many items as would be required for the final tests (i.e. 300 items to reduce to 100, at each level); four tests parallel in length and difficulty would be administered, with a common core of about 40 items at each Population level (plus the Word Recognition test for the 9 year olds); that two passages, or 10-12 items would be included in each domain as "bridging passages" between the two populations; that these forms would be rotated over students in each classroom; that about 240 students would be tested on each item where possible, so that a Rasch analysis could be conducted on the results (in addition to a classical item analysis); and that the pilot tests would be administered to judgement samples, carefully selected to represent the range of achievement in each country; that the time limits for the main testing sessions would be 30 minutes for Population A, and 45 minutes for Population B, but that some flexibility would be allowed in these limits so as to provide complete data where possible.

To achieve these objectives, the pilot tests were prepared during and after the June 1989 NRC meeting, and distributed in the internationally-prescribed format by the ICC during July. Guidelines were added for translation procedures in non-English speaking contexts, and for cultural adaptations. Two independent forward translations were requested for all test passages and items, and copies of all translated tests were sent to ICC for checking once they had been completed and administered.

The following kinds of cultural adaptations were recommended to allow for difference in conventions, test traditions and local idiom.

Examples of Permitted Cultural Adaptations to English Versions of Four Test Passages

CALENDAR:	The National Holiday could be changed to a locally suitable one; "Mum" could be changed to "Mom", "Mother", or Mamma" etc.
WHALE:	"30 metres" and "100,000 kg" could be converted to imperial equivalents.
MARMOT:	"Feet" could be metricized. In Southern Hemisphere countries, "south" could be changed to "north" (with consequent change in later items); "August" could also be changed, to "February".
MAP:	Street names could be changed to others more locally appropriate, but of equivalent length and difficulty. "Petrol" could be changed to "gas", or "benzine", etc.

Approximately 90 such modifications were suggested altogether, and were widely adopted by NRC's. Any other proposed changes were to be cleared with the ICC.

While the effects of such changes are sometimes difficult to predict, some reassurance was gained after a small study undertaken by the Steering Committee. In one country, the National Advisory Committee considered that the passage about Marmots was unsuitable, as these animals are unknown to their students. The Committee recommended that Marmots be replaced throughout the passage with Rabbits. To clarify the effects of such a change, two parallel forms of the passage and items were prepared, one including marmots and the other

including rabbits. The tests were administered, in alternating sequence, to the pupils of two classes of 14 year olds in a large co-educational high school, in February 1990. The students' test papers were scored and results for the two forms compared. The mean scores for the two formats were found to be almost identical, with the small difference favouring the Marmot passage. It was concluded that the change had no effect on difficulty levels.

Rasch Modelling

The decision to use Rasch analysis in the pilot tests was not taken lightly. While there was considerable support from the Steering Committee and many NRC's, there were a number of reservations expressed about the assumptions of the Rasch model in tests of reading comprehension. The required assumption that the test items should fit a unidimensional scale in each domain was unacceptable to some, as it could well exclude many useful questions, even passages. Moreover, the samples required for Rasch analyses were larger than those typically used in classical item analyses, making the exercise more expensive for participating countries.

Regarding the first issue, there has been a long-standing empirical debate on the structure of traditional reading comprehension tests (see above). More recent analyses, using latent trait analysis, by Andrich and Godfrey (1979), and Hillocks and Ludlow (1984), showed that reading scores did fit a single Rasch model. Haertel (1980) used latent class modelling to show that all nine of his postulated reading subskills conformed to a single underlying trait, and studies by Zwick (1987) and Davey and Macready (1990) confirmed this finding at several different age levels.

While the last word has not been said on this issue, the potential advantages of using Rasch modelling for quickly identifying poor fitting items across large and diverse samples of students was felt to outweigh the possible disadvantages. Persuasive arguments in favour of using Rasch for the pilot tests were:

(1) With a Rasch analysis, the opportunity to identify disturbances in the pooled pilot test data, which may signal problems of translation or cultural difference across countries was given considerable weight.

(2) In using three domains, the IEA test developers were allowing for three independent dimensions, rather than the single dimension which has been the major source of criticism.

(3) Classical item analysis itself, assumes the same kind of unidimensionality as the Rasch model if items are selected primarily on the basis of discrimination indices within each domain.

(4) When total scores are calculated for individuals by summing the results obtained from a wide diversity of reading test items, the assumption of unidimensionality, of each domain for which this aggregation occurs, has already been accepted.

(5) The fact that Rasch analyses produce sample-free indices of difficulty made the method appealing when it was planned to pilot so many items in rotated format on judgement samples within each country.

(6) It was agreed at the NRC meeting in 1989 that item statistics should guide rather than dominate the selection of items, and that other factors such as the test plan must[1] be considered in attempting to maximise validity.

After consultation with those who had used Rasch modelling in other studies of reading comprehension, it was agreed that the pilot test results should be subjected to Rasch analyses.

(g) CONCURRENT STUDIES OF ITEM FORMAT AND LAYOUT

While the participating countries were undertaking their pilot studies, the Steering Committee arranged for further investigation of issues that were of continuing concern to NRC's.

(1) Item Format

The question of item format in reading comprehension tests, particularly about multiple-choice versus open-ended items, was a source of debate from the outset, yet the empirical research on this topic, confined as it was to students over 14 years, was felt to be inconclusive for younger students in an international survey. Earlier correlational studies conducted in England, USA, Canada and Netherlands (e.g. Vernon 1962, Choppin and Purves 1969, Traub and Fisher 1977, Van den Bergh 1988) showed that both formats were measuring the same underlying ability. Multiple-choice tests tended to be easier, faster, and more popular with students, but the same students performed well or badly on each format. Factor analyses did not differentiate between the two formats at all.

After reviewing this research, an empirical investigation was designed by the Steering Committee on 9 year olds, and carried out by two of its members in four schools, two in a small Australian city, and two in a larger New Zealand city. The details of this project are presented in Elley and Mangubhai (1992). The main features are summarised below.

The design of the study was straightforward. Ten typical passages, representing the narrative and expository domains of the Population A pilot tests were chosen to be administered in two sessions of 35 minutes each. In the first session, 5 passages were included, and all students took these tests in multiple-choice format. In the second session, four to seven days later, 5 more passages were used. Half the students in each class, randomly selected, took the multiple-choice items, while the other half answered similar questions in open-ended format. Nineteen of the items were parallel in the two formats. The remaining 13 items were deliberately designed to be different so that the strength of each item type could be more fully exploited. Students were given either one or two lines to write their answer in the open-ended format. Half-marks were given for answers that were considered partially correct. A reliability check of the first 50 answer sheets showed high agreement between the scoring of the two researchers.

It had been hypothesised that, if format makes a difference, then students who took the two multiple-choice tests would show a higher correlation between their scores on the two occasions than students who took the two different formats. It was also hypothesized that

mean scores would be higher in the multiple-choice format, that the students would prefer it, and that the open-ended format would take longer to complete.

The first hypothesis was not supported. The correlation (N = 143) between the two tests of like format was 0.78; that between the two unlike tests was 0.81. In other words the students who performed well on a multiple-choice test were just as likely to perform well on an open-ended test as they would on another multiple-choice test. Apparently the two tests were measuring the same qualities, regardless of format. This finding confirmed that of previous studies conducted on older students.

The other hypotheses were all supported in this investigation. The multiple-choice items were easier, by approximately 20% (69% v 49%); over one-quarter of the students taking the open-ended questions failed to finish in the allotted time, while all students except one, taking the multiple-choice test, did complete the test. As for student attitudes, there was a clear-cut preference for the multiple-choice format, 88% to 12%. It was noted that many students responding to the open-ended format omitted and wrote incomplete or unintelligible answers. It is probable that their writing ability, rather than their reading ability was the major constraint on their performance. Finally, no gender differences were found in the pattern of results for the students in this investigation.

The findings of this study were reported to all NRCs. In a survey-type reading comprehension test, item format appeared to be irrelevant for determining the relative performance of groups of 9 year old students. Of course, this result was obtained with English-speaking students, who had had some, but only limited exposure to both formats before. Nevertheless, the findings from eight other studies on older students, in a variety of countries, all pointed in the same direction.

It was noted by the researchers that the value of open-ended questions for other purposes was not undermined by this study. The advantages of having students reveal their thinking in open-ended answers are many, in diagnostic testing and in numerous instructional settings. However, for large-scale one-off surveys, where attention focusses on relative attainment of groups under standardized conditions, the finding of this study, that multiple-choice tests provided the same kind of results as open-ended items did, in less time and with less cost, offered strong support for the NRCs' decision to use that format predominantly in the IEA Reading-Literacy project.

(2) Study of Layout Factors

Concern was expressed by several NRCs that the decentralization of control over the preparation of test booklets by the National Committee in each country could give rise to differences in print size and layout that would be a source of variance in student performance between countries. The Steering Committee examined this problem, and agreed to advise NRCs on a standard type size and layout to be used where possible. Text illustrations were also required to be identical, as they too were believed to affect difficulty levels.

On the question of line length and print size, opinions differed about the optimum standards for 9 year olds. There was some research (Tinker 1966, Watts and Nisbett 1974, Zachrisson 1965), but very little, that students can tolerate a wide diversity of standards, accommodating quickly to different types without any problems. Others claimed that large print and short lines did assist young students, so that strict control should be exercised over these factors. Once again, the Steering Committee resorted to an empirical study.

An investigation was conducted under the supervision of the Chairman of the Steering Committee, in a large New Zealand city involving 132 8-9 year old students from three state co-educational elementary schools. These students were administered two standardized tests of reading comprehension, under systematically varied circumstances. The first test was given under normal standardized conditions, using the tests as they are commonly printed. The seven passages read (in 40 minutes) were presented in 12 point Geneva font, in double columns. The second test, administered a week later, consisted of four alternate forms, prepared by adapting the text layout of the original test, a parallel form of the first test. The four versions were adapted as follows:

Form A : short lines (8.5 cm) and small print (10 point)
Form B : short lines (8.5 cm) and large print (14 point)
Form C : long lines (17.5 cm) and small print (10 point)
Form D : long lines (17.5 cm) and large print (14 point).

The four forms were distributed in systematic multiple-matrix fashion, so that 25% of the students attempted each form. Students were not aware of the variations in the presentation. All other layout features - print, font, leading, margins, justification, and page size were identical. The hypothesis was that there would be no difference in the performance of subjects taking the four adapted forms of the second test.

Table 2 shows the raw score means and standard deviations on each test for all subjects, classified by the forms (A to D) taken by students in the second test. The possible score in each test was 36.

TABLE 2

Means, Standard Deviations and Difference Scores on Reading Comprehension Tests with Different Print Size and Line Length

	TEST 1		TEST 2		DIFFERENCE
FORM	M	SD	M	SD	M
A	19.27	9.57	20.55	9.91	1.27
B	19.35	8.58	20.81	9.05	1.46
C	20.06	8.93	21.42	9.26	1.36
D	20.24	8.44	21.42	8.34	1.18

Clearly the differences between the four groups were trivial, and an ANOVA confirmed that there were no significant differences: $F(3, 128) = 0.03$.N.S. A similar result was found when students were compared within three ability groups. It appears that neither print size nor line length have a significant influence on the reading comprehension of 8 and 9 year old children reading under normal, relatively unspeeded test conditions in English. There may be some uncertainty about whether these results apply in other languages, but the Steering Committee saw no reason to change their decision to allow NRCs to prepare their own forms of the tests used in the IEA Reading-Literacy survey. Nevertheless the format and print types were strongly recommended by ICC and the sequence of passages, items and illustrations were standardized.

All NRCs received hardcopy versions of the tests as well as Macintosh-formatted diskettes. The basic font recommended was New Century Schoolbook (12 point) from ADOBE. A few tables were presented in Helvetica (12 point). To achieve maximum print quality, pixel graphics were drawn or scanned with 300 dpi. All NRCs were urged to ensure that, whenever possible, items faced the relevant text, in the same way as in the models distributed by ICC. Copies of the final versions of the tests, as used in each country, were subsequently sent to ICC for checking.

(3) Length of Reading Passages

Another issue on which NRCs expressed different views was the question of long versus short reading passages. In most countries, short reading passages of 100-200 words were traditional, covering a wide range of topics. In other countries, a campaign to make reading tests more authentic had resulted in a call for more extended passages, containing complete stories, rather than brief isolated extracts which were felt to be artificial. In the present study several passages, over 600 words in length, were submitted for possible inclusion in the test. Apart from their greater authenticity, it was claimed that such passages gave students a

greater opportunity to build up new schema, and use them to interpret subsequent events, an ability not easily tapped in short passages.

On the other side, proponents of short selections argued that a wider range of topics and genres could be included in the same testing time if long passages were avoided, an important consideration in an international test where cultural biases were to be minimised. For instance, one long passage proposed on the value of plants to human life, was criticized not just for the time it would take to read, but also because it would favour students who were better informed about this topic. A very long story on the relationship between two adult sisters was rated unacceptable by several NRCs because the topic would disadvantage boys, and some girls, who had little interest in the theme, and who would thus find it a daunting task to complete the reading. Other points were raised about the extent to which young people typically did read long or short selections, and whether there really was any difference in test results from the two kinds of passages.

This was regarded as an empirical question, which the Steering Committee felt obliged to follow up. Only one authoritative study of the problem was located and this was conducted on 3565 US high school students by Modu and May (1977). They compared the comprehension results obtained from short (approximately 100 words), medium (200 words) and long passages (over 400 words), using three different multiple-choice tests. The correlations between the short passages, and a mixture of short, medium and long were extremely high (.946 and .947 on two independent samples) after correction for attenuation, and the researchers concluded that they were measuring the same ability. Comparison of the same items in short or long passages showed that students had more difficulty answering in the latter type. They also found greater reliability per unit of testing time in the case of short passages, and greater correlation with vocabulary tests.

In view of the logical and empirical arguments favouring short passages, the Steering Committee decided to make extensive use of them in the proposed tests. However, the lack of research using younger students and longer narratives, and the need to provide face validity in countries where longer passages were widely used, provided additional reasons to retain one or two longer passages at each age level.

(h) PILOT STUDY TEST RESULTS

The pilot studies were conducted in each country between October 1989 and March 1990. The selection of samples of students was guided by instructions sent from ICC in the National Co-ordinator's Manual. The purpose of these guidelines was to arrange for all countries to select a judgement sample which would allow for over 200 typical students to attempt each test item. Later checks showed that most samples were well chosen in this respect.

Financial constraints prevented a few countries from conducting a full pilot survey, while others joined the study too late to make the necessary arrangements. Nevertheless, 31 countries undertook the pilot exercise, most of them at both population levels. All but two of these countries managed to code and send their data sets to the ICC in Hamburg by the end of March 1990. The whole exercise proved very useful in most countries, as it provided an opportunity for less experienced NRCs to go through all the procedures needed for the final testing. Each NRC sent in a report on their experience, together with copies of the tests and manuals used.

(1) Analytic Procedures

In preparation for the item analyses, a small Technical Sub-Committee[1] which had been established by the Steering Committee met in Hamburg in October 1989 to identify priorities, to recommend computer equipment, and to spell out the details of data presentation in such a way that quick decisions and recommendations could be made by the Steering Committee about the selection of items for the final tests.

When data sets were received by the ICC, they were subjected to careful checking and cleaning, until they were considered ready for processing. Tables 3 and 4 present the Ns for each instrument used at each Population level.

1 The Sub-Committee members were Professor Albert Beaton, who was then Chairman of the IEA Technical Advisory Committee, and Professor Neville Postlethwaite, Dr. Ken Ross, and Andrew Schleicher. Other members were co-opted to the Committee for particular purposes, including Professor Warwick Elley, Dr. Nadir Atash, Dr. Hans Wagemaker and Dr Peter Allerup.

Table 3

Numbers of Students, Teachers and Schools in Pilot Study, Population A

COUNTRY	STUDENT TESTS (N PER FORM)*	STUDENT QUESTIONNAIRE	TEACHER QUESTIONNAIRE	SCHOOL QUESTIONNAIRE
Belgium	260	995	56	31
Canada (BC)	223	872	-	-
Canada (Q)	246	930	32	13
Denmark	235	255	52	44
Finland	263	130	31	16
France	137	763	-	-
Germany (W)	128	298	20	13
Greece	183	718	36	24
Hong Kong	295	1120	30	13
Hungary	153	537	23	9
Iceland	255	998	49	22
Indonesia	181	720	34	29
Ireland	224	863	26	9
Italy	208	828	-	-
Netherlands	185	739	24	27
New Zealand	71	261	16	6
Norway	209	813	45	18
Portugal	204	794	21	21
Singapore	229	908	26	9
Slovenia	157	596	27	14
Spain	269	1022	35	25
Sweden	278	1054	52	33
Switzerland	212	202	42	-
Trinidad & Tobago	189	713	30	26
USA	275	1052	50	32
Venezuela	250	809	20	17

* There were four forms. Thus, for example, Belgium had a total of 1040 students taking the tests. (4 x 260).

Table 4

Numbers of Students, Teachers and Schools in Pilot Study, Population B

COUNTRY	STUDENT TESTS (FORM 1 ONLY)	STUDENT QUESTIONNAIRE	TEACHER QUESTIONNAIRE	SCHOOL QUESTIONNAIRE
Belgium	208	766	39	6
Botswana	323	317	34	20
Canada (BC)	212	797	-	-
Canada (Q)	229	236	25	4
Denmark	250	290	52	36
Fiji	50	121	-	-
Finland	260	116	35	12
France	247	173	-	-
Germany (W)	121	485	20	12
Greece	218	187	25	19
Hong Kong	274	1077	30	14
Hungary	161	102	23	9
Iceland	247	846	44	-
Ireland	206	762	27	6
Italy	178	670		-
Netherlands	176	678	24	20
New Zealand	98	368	28	4
Norway	443	1526	66	16
Phillipines	705	2815	23	24
Portugal	172	293	13	9
Singapore	295	540	35	12
Spain	324	1243	42	23
Sweden	290	596	47	13
Switzerland	193	738	42	25
Trinidad & Tobago	230	242	-	20
USA	283	1093	51	44
Venezuela	223	636	25	10
Zimbabwe	60	48	8	7

A huge data set was processed by the ICC Technical Team in a very short time, and the results for the Marking Tests were presented, one item per page, as follows:

(A) Identification

1 Name of item.

2 Domain classification:

i.e. I - Word recognition
S - Simple narrative
N - Narrative prose
E - Expository prose
D - Document.

3 Number of the test form (out of 4) in which the item was located.

4 Number and name of the test passage (or Set) in its test form.

5 Item number within the passage.

6 Item construct, or mental process, presumed to be used in answering the item:
i.e. V - Verbatim match between words in text and item
P - Paraphrase
M - Main idea, message
I - Inference
L - Locating information
D - Following directions.

(B) Results by Country

1 N of students (in each country) who were administered the item.

2 P value: i.e. the percentage of students who had the item correct.

3 Point-biserial index within each domain: i.e. the correlation between the item and the domain score, in each form of the pilot test.

4 Percentage of students choosing each option, in multiple-choice items.

5 Percentage of students 'Omitting' the item (OMIT).

6 Percentage of students 'Not Reaching' the item (NR).

7 Point-biserial index for each option in multiple-choice items, and for 'Omissions' and 'Not Reached' categories.

8 National Rasch item difficulty, equated over the 4 test forms, excluding 'Omits' and 'Not Reached'. The Rasch scale ranges from +3.0 (very difficult) to -3.0 (very easy).

9 Rasch deviation index: i.e. the deviation of the country's Rasch difficulty index from the average of all countries' index.

10 National Rasch fit index: i.e. the extent to which the item fits the Rasch model derived from the data for that country. An item with good fit (near zero) is 'behaving' in a similar way to other items in the domain. Statistically, it represents the standardized weighted mean square residual, for each country.

11 Standard error of the national Rasch difficulty index.

12 P values for males and females separately.

13 Point-biserial index of the correct answer, excluding 'Omits' and 'Not Reached'.

14 Rasch international fit index: i.e. the extent to which the item fits the Rasch model derived from the data of a pooled sample of 150 students drawn at random from each country. Indices near zero indicate good fit; large deviations from zero indicate 'disturbances' in the data. High positive values were often associated with large variations in p values and point-biserial indices, gender differences, cultural differences and the like. High negative values indicated unusually high point-biserial indices, and were often associated with longer passages with large numbers of 'Omits', or 'Not Reached' figures.

15 Means, across all countries, for p values, for point-biserial indices (with and without omits), and for Rasch difficulty indices.

16 Flags to assist in identifying deviations, such as:
(a) p value is over 90%, or under 2%
(b) p value of distractor is higher than that of correct answer
(c) point-biserial index is less than 0.25
(d) point-biserial index of distractor is positive and greater than 0.10
(e) p value of genders differs by at least 10%
(f) omits exceed 25%
(g) Rasch fit index is greater than 2.5.

In addition, a 'Table of Effects' was prepared to identify large residuals, when the variance associated with Rasch difficulty indices, across items and countries, was removed. This table was also useful in assisting with the identification of any items that were culturally inappropriate, and was used to alert NRCs to problems of translation.

(2) Interpreting the Item Analysis

An example of a full set of statistics from an item, drawn from the pilot study item analysis for Population B is given below.

POP B VOLCANO 2, FORM 4, Item 2

(See next page).

Reading from the top of the table it can be seen that this item is the second one in the passage named VOLCANO, and is referred to by the label "Destroyed in Eruption." The passage is located in the Expository domain (E), is part of the 4th parallel form of the pilot test, and is the 5th passage in that form. The construct of the item is classified as P, as it requires students to paraphrase part of the text.

The explanations for each statistic are given below the heading, and the actual results for the item in the lower part of the table. The 28 participating countries are listed in alphabetical order and 22 statistics are presented for each.

Thus, Belgium had 208 Population B students taking this item; 80.77 percent answered it correctly, and the point-biserial index for Belgium was 0.54. Both these statistics are slightly higher than the international means, given at the foot of each column. The small p in the Flag

B44VOL02 VOLCANO-02/DESTROYED IN ERUPTION (C)
Domain: E Form: 4 Passage: 5/VOLCANO Item-No.: 2 Constructs: P

Expl.:
N : Number of students who were administered the item
P-VAL : Percentage of correct answers based on all students who were administered the item
P.BIS : Point biserial correlation based on domain-score; P.BM : p.bis with omits excluded
FLAGS : Codes: P: p-value of option <2 or >90, D: attractive distractor, B: p.bis <.25 or p.bis of distractor > .1, S: sex-bias > 10%
O: more than 25% omits; F: not fitting the Rasch scale (fit > 2.5)
OMIT : Percentage omitting the item
NR : Percentage not reaching the items where this is based on the test form
DIFF : Equated difficulty index from the Rasch Model
DEV : Deviation of the countries difficulty index from the average difficulty index
FIT : Item fit statistic
ERR : Maximum likelihood standard error of the adjusted item difficulty
P-MAL : Percentage of correct answers for male students
P-FM : Percentage of correct answers for female students

					Percentages for each alternative						P.BIS for each alternative						Rasch						
COUNTRY	N	P-VAL	P.BIS	FLAGS	A	B	*C	D	OMIT	NR	A	B	C	D	OMIT	NR	DIFF	DEV	FIT	ERR	P-MAL	P-FM	P.BM
BELGIUM (FR)	208	80.77	0.54	P---	9.6	1.4	80.8	1.4	6.2	0.5	-.25	-.13	0.54	-.21	-.36	-.17	-1.22	-0.03	-0.74	.25	79.4	86.7	0.43
BOTSWANA	297	57.24	0.37	----	12.8	12.1	57.2	13.5	4.4	0.0	-.26	-.08	0.37	-.08	-.20		-1.52	-0.32	-1.42	.13		100.0	0.34
CANADA (BC)	206	79.61	0.47	P--F	11.2	1.5	79.6	2.4	5.3	0.0	-.17	-.12	0.47	-.29	-.34		0.00	1.20	-3.48	.20	77.0	80.9	0.38
CANADA (Q)	199	77.39	0.42	P---	13.6	3.0	77.4	1.5	3.5	1.0	-.13	-.27	0.42	-.16	-.21	-.27	-1.49	-0.29	-0.58	.21	78.0	76.2	0.35
DENMARK*	216	84.26	0.40	P---	7.4	1.4	84.3	0.0	6.9	0.0	-.13	-.09	0.40		-.10		-1.56	-0.86	0.75	.25	89.1	84.7	0.22
GERMANY (W)	127	77.95	0.33	P---	18.9	0.8	78.0	0.8	1.6	0.0	-.20	-.00	0.33	-.17	-.34		-0.69	0.51	-0.42	.24	81.7	73.4	0.27
FIJI	48	83.33	0.24	.B--	4.2	10.4	83.3	2.1	0.0	0.0	-.09	-.25	0.24	0.04			-2.32	-1.12	0.29	.45	78.6	82.6	0.24
FINLAND*	0			----																			0.13
FRANCE	149	40.27	0.85	P.B.O.	5.4	0.0	40.3	0.0	53.0	1.3	0.20		0.85		-.88	-.16	-1.16	0.04	0.34	.42			0.27
GREECE	213	70.42	0.29	P---	20.2	4.2	70.4	0.5	4.7	0.0	-.16	-.17	0.29	-.07	-.14		-0.68	0.52	1.44	.18	72.1	70.5	0.35
HONG KONG	268	78.73	0.36	P---	17.5	2.2	78.7	1.1	0.0	0.4	-.24	-.14	0.36	-.19		-.23	-0.73	0.47	0.87	.17	74.1	82.8	0.35
HUNGARY	132	81.82	0.43	P.S.	11.4	2.3	81.8	0.8	3.8	0.0	-.25	-.17	0.43	-.02	-.31		-1.25	-0.05	0.06	.27	92.3	46.2	0.35
ICELAND	220	75.45	0.44	P.S.	10.9	5.5	75.5	1.4	6.8	0.0	-.01	-.31	0.44	-.12	-.28		-1.06	0.14	0.61	.20	82.1	70.0	0.33
IRELAND	208	82.21	0.49	P---	10.6	3.4	82.2	1.0	1.9	1.0	-.26	-.22	0.49	-.07	-.28	-.26	-1.57	-0.38	-0.54	.22	86.5	79.2	0.41
ITALY	163	83.44	0.34	P---	11.0	4.3	83.4	1.2	0.0	0.0	-.31	-.03	0.34	-.23			-1.12	-0.22	-0.11	.23	79.5	87.7	0.34
NETHERLANDS	173	79.19	0.27	P.S.	19.1	1.2	79.2	0.0	0.0	0.0	-.25	-.05	0.27		-.07		0.00	1.20	-1.81	.19	84.8	73.0	0.26
NEW ZEALAND	92	73.91	0.58	P---	16.3	3.3	73.9	0.0	6.5	0.0	-.23	-.40	0.58		-.40		-0.65	0.55	0.09	.31	79.5	72.3	0.49
NORWAY	443	75.62	0.48	P---	12.2	1.8	75.6	0.2	9.8	0.9	-.22	-.07	0.48	-.00	-.35	-.24	-1.47	-0.27	-0.41	.15	73.0	77.8	0.34
PHILLIPINES	698	61.32	0.33	----	16.6	18.3	61.3	3.7	0.0	0.0	-.16	-.18	0.33	-.15			-1.03	0.17	-0.31	.08	59.8	62.3	
PORTUGAL	171	84.80	0.47	P.S.	10.5	0.6	84.8	0.6	8.5	0.0	-.22	-.14	0.47	-.10	-.46		-1.15	0.06	-0.83	.26	-20.6	73.8	0.34
SINGAPORE	286	86.36	0.20	P.B.	10.5	2.4	86.4	0.0	0.7	0.0	-.07	-.18	0.20		-.20		-1.38	-0.18	2.26	.20	86.5	86.2	0.16
SPAIN	296	90.20	0.21	P.B.	7.1	1.0	90.2	0.3	1.4	0.0	-.19	0.00	0.21	-.05	-.12		-2.03	-0.83	0.75	.24	89.2	91.0	0.18
SWEDEN	289	80.97	0.40	P---	16.8	0.7	81.0	0.3	1.4	0.8	-.30	-.11	0.40	-.19	-.21	-.16	-1.20	0.00	0.06	.17	92.2	79.4	0.36
SWITZERLAND	187	81.28	0.37	P---	11.8	2.1	81.3	0.5	3.7	0.5	-.15	-.18	0.37	-.19	-.22	-.24	-1.11	0.09	-0.17	.23	79.8	83.1	0.30
TRINIDAD TOB	215	86.98	0.31	P---	7.9	4.7	87.0	0.0	0.0	0.0	-.22	-.17	0.31		-.26		-1.49	-0.29	0.84	.23	88.9	87.5	0.30
USA	280	77.86	0.31	P---	17.1	4.3	77.9	0.4	0.4	0.0	-.15	-.30	0.31	-.13	-.09		-0.79	0.41	1.02	.17	80.3	76.5	0.31
VENEZUELA	200	72.00	0.60	P---	11.0	4.0	72.0	1.0	10.5	1.5	-.20	-.13	0.60	-.12	-.43	-.29	-1.47	-0.27	-1.15	.22	69.9	74.4	0.43
ZIMBABWE	52	80.77	0.34	P.S.	13.5	5.8	80.8	0.0	0.0	0.0	-.20	-.28	0.34				-1.65	-0.45	0.71	.40	89.5	75.8	
MEANS INTERNAT. SCALE		77.19	0.40														-1.20		-1.32	.05			0.32

column indicates at least one of the multiple-choice alternatives attracted less than 2% of students, which was the case for all countries except Botswana, Fiji and Philippines. (This particular item was not taken by Finland.)
Under the percentages for each alternative, it can be seen that 9.6% chose A, 1.4% chose B, 80.8% chose the correct answer C (which is asterisked) and 1.4% chose D. A further 6.2% omitted the item, and 0.5% did not reach it. The next 6 columns show the point-biserial indices for each alternative, which are positive for the correct answer (0.54) and negative for the remainder, as would be expected.

The Rasch figures show that this item is relatively easy (Diff. of -1.22) and that it is closely in line with other countries in this respect. The item shows a relatively good fit to the Rasch model, at -0.74, and the standard error is within tolerable bounds (0.25). Female students (86.7%) scored higher on the item than males (79.4%) in Belgium, and the point-biserial index dropped slightly to 0.43, when the omits were excluded.

Inspection of the results for all listed countries reveals that most students in every country found it relatively easy, the p values clustering around 0.8. The exceptions to this pattern were Botswana and Philippines, which tended to show lower scores on all items, and France, which had a very high omit rate on this item - for reasons associated with test administration rather than with student ability.

The point-biserials were consistently positive, and moderate in magnitude, except for France. The flags show three capital B's, indicating low point-biserials, and one lower-case b, (for France), indicating a high one. In four cases there was a sex difference of 10%, all favouring boys. In Hungary these figures denoted a large difference, which was due to the nature of the sample selected for testing, rather than to any characteristics of the item. The single f, shown in the Flag column, indicates that this item does not fit the Rasch model well in British Columbia, and the O beside France flags the high number of students that omitted the item.

The next 6 columns show that A was the most popular distractor in all countries except Fiji and Philippines, and that B and D were consistently unpopular, especially D. The 'Omits', and 'Not Reached' columns showed that most students attempted this item, except in France.

Examination of the 6 point-biserial columns shows the expected positive indices for C, the correct answer, and consistently negative indices for the remainder. The only exceptions are France for Distractor A, Spain for Distractor B, and Fiji for Distractor D. The gaps in these 6 columns indicate that nobody selected those options.

The Rasch Difficulty column shows only negative signs, signifying an easy item in all countries, and the Deviation column shows only minimal variation also. The item appears to fit the Rasch model well in all cases except British Columbia, and the standard error is below 0.30 in all countries which tested N's over 100. The gender columns show close agreement in all except 5 countries, at least 2 of which were due to biassed sampling. The final column shows predictable declines in the point-biserial indices following the deletion of 'omit' and 'not reached' cases. Because of the large number of omits in France, the index dropped substantially, from 0.85 to 0.13.

The means at the foot of the table confirm that the item is easy (77.19% correct) and discriminates moderately well (0.40). The Rasch mean difficulty index confirms the difficulty level, and the point-biserial mean, after the exclusion of 'omits', shows a small decrease from 0.40 to 0.32. Finally, the Rasch international fit index of -1.32 signifies an item that is well within the expected bounds for the Rasch model, based on pooled data from all countries.

The general impression gained from these indices is that the item behaved in very similar fashion in all countries, except where there was evidence of an atypical sample of students, or when the item was not attempted by many.

(3) Descriptive Data on the Complete Pilot Tests

Normal checks for reliability within each domain were carried out on the four separate forms of the test, using the KR 20 formula. The reliability indices on the three major domains for Form 1 of the test of each population are given in Table 5.

Table 5

Pilot Test: KR 20 Reliability Indices for Form 1, Populations A and B

COUNTRY	EXPOSITORY		NARRATIVE		DOCUMENTS	
B.	Pop A.	Pop B.	Pop A.	Pop B.	Pop A.	Pop
Belgium (Fr)	.84	.85	.79	.86	.83	.87
Canada (BC)	.86	.87	-	.81	.86	.87
Canada (Q)	.83	.81	.76	.84	.85	.85
Denmark	.92	.87	.87	.86	.89	.90
Finland	.82	.84	.86	.82	.93	.88
France	.80	.79	.79	.79	.62	.85
Germany (W)	.88	.86	.84	.85	.88	.83
Greece	.84	.77	.83	.75	.83	.80
Hong Kong	.80	.71	.78	.77	.79	.67
Hungary	.85	.82	.83	.84	.85	.82
Iceland	.88	.78	.85	.83	.88	.82
Indonesia	.79	-	.72	-	.84	-
Ireland	.86	.82	.74	.84	.84	.81
Italy	.89	.83	.80	.80	.91	.87
Netherlands	.85	.72	.79	.74	.84	.80
New Zealand	.91	.88	.88	.85	.89	.91
Norway	.87	.85	.84	.84	.84	.88
Portugal	.82	.84	.83	.82	.84	.83
Singapore	.88	.77	.84	.71	.86	.78
Slovenia	.85	-	.81	-	.81	-
Spain	.84	.74	.82	.79	.78	.84
Sweden	.90	.86	.86	.84	.87	.91
Switzerland	.83	.83	.79	.75	.83	.86
Trinidad & Tobago	.85	.81	.78	.85	.86	.80
USA	.87	.86	.84	.85	.82	.85
Venezuela	.91	.91	.85	.89	.93	.91
Mean (All countries)	.86	.82	.82	.82	.84	.84
Mean (English speaking)	.87	.83	.81	.82	.86	.84
No. of items	36	38	24	28	30	36

Inspection of Table 5 shows that most reliability indices for Form 1 cluster between 0.80 and 0.90, which was considered satisfactory for short tests containing many relatively untried items. The other forms showed similar figures. More encouraging was the consistency of the figures across all countries. The only indices which fell below 0.75 were known to be due to problems of administration or scoring. It is of interest to note also that the mean indices for

the six English-speaking countries are almost identical to the remainder, which would be expected if the items were behaving in similar fashion in different languages.

Another analysis performed at this stage was designed to investigate the inter-relationships between the major domains. For all students, total scores on each domain within each test form were calculated and correlated with the scores obtained by the same students on the other domains. For Population A, the mean correlation coefficients across all four test forms are given in Table 6. Mean reliability indices are given (in brackets) in the diagonals.

Table 6

Mean Inter-Correlations and Reliability Indices for Population A Pilot Tests, by Domain

DOMAIN	EXPOSITORY	NARRATIVE	DOCUMENTS
Expository	(.83)	.71	.70
Narrative		(.81)	.63
Documents			(.87)

As the inter-correlations cluster around 0.70, it is clear that the students who performed well in one domain of reading tended to perform well in others, but the correlations were considerably lower than the mean reliability indices (over 0.80). The evidence was not strong, but sufficient to warrant maintaining the distinction between the three domains for the final tests. Similar trends were observed in the Population B results.

The Word Recognition Test was not included in these analyses because the scores were very high, close to the ceiling in most countries. As stated earlier, the purpose of this test was more to check for basic decoding skills, as a possible factor in explaining pervasively low scores across certain groups or sub-groups, and to provide some comfort for weaker readers who may have found little they could respond to in the remainder of the test.

(i) SELECTING THE FINAL TESTS

In order to select passages and items for the final tests, the Steering Committee, augmented by four NRCs, established a set of criteria at the outset. While most believed that the statistical criteria of adequate discrimination and Rasch model fit were important, it was agreed that content validity should also be given due weight. It was not possible, within the tight time and financial constraints of the study, to conduct a second pilot study. Therefore, it was felt that major changes to items, to bring them into line with the psychometric needs of the project were not feasible. If an item failed to discriminate, even though it was acceptable for other reasons, it usually had to be dropped. Such constraints made a certain amount of compromise inevitable when criteria were in conflict. Fortunately, there were a large number of items which appeared to be operating well by all criteria, so these decisions were not often difficult. It was pleasing to the researchers to find so many passages and items which not only showed adequate discrimination, but also behaved in such similar fashion across nations, as indicated by the Rasch fit statistics.

(1) Ten Principles of Item Selection

After a thorough discussion of the issues, the following 10 criteria for selecting items for the final tests were agreed to:

1 The passages and associated items should be of sufficient quality, in terms of writing style, interest and technical characteristics, to justify inclusion in an international test.

2 There should be a wide range of topics and genres represented in the passages chosen, i.e. topics should show a balance of topics related to home, school, society or work, and there should not be multiple passages on similar topics.

3 There should be a reasonable balance of skills or mental processes represented in the test items, i.e. NRCs expected to see items testing verbatim, paraphrase, main idea, inference, directions and locating information in documents, as these were all approved for the pilot tests. It was agreed too, that narrative passages would contain more items concerned with character interpretation, feelings, motives, attitudes, while the expository passages would have more items stressing facts, sequences, functions, cause-effect relations and generalisations about content.

4 There should be a range of difficulty levels in each test, i.e. There would be a few items easy enough to cater for weaker readers, and others difficult enough to extend the able students.

5 Discrimination indices should be consistently high, i.e. the point-biserial indices calculated within domains and forms should be above 0.30, preferably higher.

6 Items should show "good fit" to the Rasch international model derived from the scale which was generated by the remainder of the items in each domain, both within and across countries, i.e. the fit index should be near zero, or low negative in most cases. High positive fit indices (above 0.3) should be especially avoided, as they indicated disturbances in the student response patterns, due to culture bias, translation problems, gender differences or other difficulty.

7 Items with large gender differences should be avoided. A check on p values by gender for each country enabled this variable to be examined statistically.

8 Test passages in the narrative domain should vary in length from short to long. More specifically, there should be at least one narrative passage greater than 600 words in each test.

9 Cultural sensitivities should continue to be observed, i.e. the comments of NRCs about the individual items and passages trialled in the pilot tests should be carefully considered in selecting items.

10 There should be at least one bridging passage with the other population level in each domain, i.e. it should be possible to link the results for Population A with those of Population B by means of a common set of items in each domain.

Occasionally these criteria were in conflict, and the Steering Committee spent much time in making small changes to a few items to produce an acceptable balance. Inevitably, some very good items had to be omitted from the final forms, because they duplicated others, or because several items in the same passage set were unsuitable. Any passage with fewer than three or four good items was felt to be inefficient in terms of student testing time.

Fortunately, the original decision to try out three times as many items as were targetted for the final test proved to be a sound one, and consensus was reached, during the selection meeting on virtually all items to be recommended for the final tests.

Of the 300 items pilot tested for each population nearly 200 were dropped at each level. A few passages which showed acceptable psychometric properties had to be deleted because they were vetoed by one or more countries on moral or cultural grounds. These passages had not been identified earlier, apparently because some NRCs had not consulted as widely as they might have. A few other passages were criticized because the text or items were considered to be of indifferent quality by a few. Still others were dropped because they were competing with very good sets in the same category. Most, however, were deleted because they did not have enough international "good-fitting" items to warrant inclusion. None were deleted because of obvious gender bias. Those items which were selected were assembled and distributed to NRCs in June 1990 for further review and approval.

(2) Other Decisions on Proposed Final Test Format

In addition to the selection of particular passages and items, the Steering Committee made a number of further resolutions, based on the results of the pilot tests and the written comments of NRCs which were received at ICC after the exercise was completed.

1 The domains proposed for the pilot test blueprint; viz: Narrative, Expository and Documents (plus Word Recognition for Population A) were retained, as they each showed that some statistical variance was non-overlapping.

2 The Word Recognition Test was, provisionally, reduced from 48 items to 30, as a few items showed unacceptable differences across countries. The time limits were shortened from 90 seconds to 60 in consequence.

3 The easy narrative items adapted from the 1970 IEA reading tests were dropped, as they were found to be unnecessary, difficult to translate fairly in a few languages, and psychometrically inadequate in several countries.

4 Of the 20 free-response items, calling for extended (one or two-sentence) responses, the majority proved to be problematic. Large numbers of students omitted them, and NRCs from several countries objected to the time and cost of scoring them. However, a number of NRCs argued for some open-ended items, chiefly on the grounds of face validity and national policies within their country. After discussion, it was agreed that a single extended free-response question should be included in the Narrative and the Expository domains at each population level (i.e. four such questions altogether). These were to be included in every country's test, but the results were not to be included in the main international data set. Those NRC's who wished to were invited to participate in a separate international exercise organized by the United States NRC .

5 It was decided not to use rotated forms, either at Population A or B level in the main testing. The reasons advanced for avoiding rotation were increased adminstration difficulty, inflated cost, anticipated problems in secondary analysis, and strong appeals from various NRCs, based on their pilot test experience. The main disadvantage of this decision was that there would be fewer items on each domain.

6 The proposed tests were to be given in two sessions of about 40 items for Population A, and two sessions of 50 items at Population B level. These proposals allowed for about 27 items per domain at Population A and 32 per domain at Population B. The items from the three domains were to be intermingled, and the questionnaires for students were intended to be administered in a separate session, after the tests. The length of the tests was influenced by the desire to obtain three reliable domain scores, for each student. In many countries NRCs wished to confine testing to no more than two complete sessions of 35-40 minutes at Population A level, and 45-50 minutes at Population B level. After discussing many options it was decided to aim for two test sessions of about 40 items (plus Word Recognition) at Population A level and 2 sessions of about 50 items at Population B level. These figures assumed slightly more than one item per minute - which was found realistic in most countries that reported on timing. The document items were found to take less time than this, the longer prose passages slightly more.

7 The decision to include at least one longer narrative passage at each level was confirmed, as several NRC's argued that it increased the face value and authenticity of the whole reading exercise.

Several minor changes were made by the Steering Committee to the practice items and to the timing of the tests and small changes were made in the passages and items selected. Most of these changes were editorial, to improve parallelism or style or sequence, or to remove small sections of irrelevant material from the passages, but alterations of some substance were made to at least one option in 9 items of the proposed Population A tests, and in 6 items of the Population B tests. Most of these changes were due to excessively attractive distractors.

(3) Approving the Final Tests

NRCs reviewed the proposed tests during June 1990 and assembled at Frascati, Italy in July for final decisions. NRCs were given the opportunity to express any reservations they or their national committees had about format, timing, or the inclusion of particular items. In a round-table discussion, most NRCs expressed satisfaction that the best possible tests had been produced, but a few concerns remained.

(a) Several NRCs sought to reduce the number of items for Population A, as the test was felt to be too long.

(b) Some wished to drop long passages for young readers,

(c) Some asked for more practice items,

(d) Specific suggestions were made for improving particular passages or instructions to students.

(e) There was still some concern expressed about item format, but most NRCs were prepared to accept the decision to include two extended answer questions at each population level, the answers to be scored locally, and not included in the international data set.

(g) One NRC had undertaken an extensive study of the Word Recognition Test results across countries, and had proposed changes in items, timing and instructions. (Gil, 1990).

To accommodate the desire for more time at Population A level, the time allowed for the Word Recognition Test was increased from 60 seconds to 90, but the number of items was also increased from 30 to 40. The main tests were reduced in length by deleting two passages at each population level, and extending the time limits slightly. This allowed for more than one minute per item which was more liberal than the timing of standardized tests in many participating countries. Reports from pilot tests in several countries showed that this timing was ample. Population A students were required to complete 68 items in 2 sessions totalling 75 minutes; Population B students to complete 91 items in 95 minutes. All NRC's finally agreed to these decisions.

Sub-committees also reviewed the instructions, the practice items, and the item sequence, and minor changes were made to these. It was agreed that the Document items should be placed together in two blocks within each test, rather than scattered throughout, as these items did require a different mind set from the continuous prose passages. However, the majority of the changes made were cosmetic. The final tests represented, as well as could reasonably be expected, the consensus approval of all countries. No NRC expressed strong reservations about the completed instruments.

(4) Translation Procedures

As 22 of the participating countries were conducting their surveys in languages other than English, it was important to secure some standardization on procedures. Previous studies conducted by IEA had confronted this matter in a variety of ways, but most required NRCs to arrange for careful translations, by educators competent in English, followed by independent back-translations into English to ensure sufficient rigour.

In the case of reading, it was a particularly crucial issue as important changes in nuances of meaning and of difficulty could be conveyed by slight but arbitrary differences in wording.

In the first sessions with NRCs on the matter, members of the Steering Committee identified and discussed several items from the earlier IEA Reading Comprehension Survey (Thorndike 1973) which revealed differences in difficulty apparently due to translation problems. For instance, in one passage about a Desert Siege, four of the six items showed stable p values across four selected countries, two of which tested in English and two in Dutch. However, on one item requiring an understanding of the word "stampede", the Dutch pupils performed significantly better than expected on the basis of the English pupils' results. The most likely explanation for the discrepancy was that the English word "stampede" was translated into the Dutch equivalent for "panic" which was considered to be more familiar. Where much hinges on a single word, such discrepancies can be serious.

Other problems revolved around the extent to which the words of the test item matched the relevant words in the text. In some cases there was a deliberate, word-for-word, verbatim match between item and text. In other cases the item required the students to paraphrase the wording of the text. If the English text required students to transform "frightened" to "terrified", while the non-English text used the same word for "frightened" in both places, the item would clearly be easier in the latter case.

Numerous recommendations were made to NRCs about ways of handling such issues, and these procedures were tried out during the pilot phase. In the subsequent discussions, many called into question the validity of back-translation. Most argued that two good forward translations, followed by negotiation between translators on the discrepancies, was more effective.

Cross-cultural psychologists (e.g. Hulin, 1987) also question the value of back-translation in obtaining equivalence of meaning.

In the final guidelines to NRCs, the Steering Committee drew on previously prepared guidelines in other IEA studies, the wealth of experience in the Reading-Literacy researchers, and the results of the item analyses from the pilot study. The following points were stressed:

1 The aim of the translation is to convey the intended meaning of the author as closely as possible, without changing the difficulty of the text. No simplification or clarification of ideas should be made.

2 Independent translations by at least two different expert translators, familiar with age-appropriate linguistic demands are recommended for all passages and items. In cases of disagreement consensus should be aimed at either by direct negotiations or by a third expert making the final choice.

3 Particular attention should be paid to how the wording of questions matches the wording of the text in the relevant sections of text.

4 It was recommended that at least one passage at each level (narrative or expository) is subjected to back-translation. The translation back into English should be done as accurately as possible, and the result compared with the English version, to check for omissions, additions, unsuspected changes in meaning, or other problems. It was recommended that the back-translations be carried out by someone unfamiliar with the original text, and that it be done literally, rather than in polished idiom. There is no simple algorithm for evaluating the degree of discrepancy between the back-translation and the original English version. However, serious problems are easily detected. In cases of large deviations, additional passages should be selected for further back-translations, to see if the problem is common.

5 Item analyses of the pilot tests identified items with poor indices of fit to the RASCH-model. One probable reason for the poor fit is related to translation. Each NRC with translated tests was required to examine the flagged items. Problems with wording were suspected in these cases. Thus, for example, instead of a paraphrase question in the original version a verbatim question might have been created in the translated version. That was easily checked.

As a further check on the quality of translations, all NRCs were asked to make a report on difficulties they experienced in translating particular passages and items, and to send their final test versions to the ICC for examination.

In the final analysis the translations were believed to have been competently and fairly carried out, as the items behaved in similar fashion, with minor exceptions. The evidence for this conclusion comes from the good Rasch fit figures, internationally, and from a conventional analysis of the p values for all items across countries. This analysis showed that the p values correlated with a mean of over 0.92 with the international mean p values at each age level (See Table 13).

PREPARATION OF MANUALS

In preparation for the printing and administration of the final tests, a series of manuals, outlining standardized procedures for NRCs, School Coordinators and Test Administrators, was sent to all countries. These manuals provided details about the directions for assembling, translating and printing the tests, the timetables and conditions for test administration, the recommended procedures for contacting schools, the training of test administrators, and the documentation to be returned to ICC.

Flow charts were provided of the various actions required of all NRC's, and dates for testing and returning data were listed for all countries. Southern hemisphere countries administered the tests between September and November 1990, and Northern hemisphere countries administered theirs between January and April 1991.

PROPERTIES OF THE FINAL TESTS

The main characteristics of the final tests, as administered in all countries, are given in the Tables below. Tables 7 and 8 present the blueprints for both populations respectively, classified by domain and by mental process.

Table 7

Final Blueprint for Population A: Test Skills

Domain	Verbation	Paraphrase	Inference	Locate Information	Locate and Process	Total
Narrative	1	11	10	-	-	22
Expository	7	9	5	-	-	21
Documents	-	-	-	11	12	23
Total	8	20	15	11	12	66

Table 8

Final Blueprint for Population B: Test Skills

Domain	Para phrase	Main Idea	Infer ence	Locate Information	Locate and Process	Tot al
Narrat ive	9	3	17	-	-	29
Expos itory	15	4	7	-	-	26
Docu ments	-	-	-	14	20	34
Total	24	7	24	14	20	89

1. Item Statistics:
This section describes in some detail the contents of the final test items by mental process, format, theme, difficulty level, discrimination level and international equivalence.

(i) Mental process
The judgements about the mental processes tested by each item were determined provisionally by the Steering Committee, and verified by means of a classification exercise conducted by 27 suitably qualified raters in 12 countries. Definitions and examples of items in each category were sent to NRCs by the ICC, and the items of both tests were classified by two (or three) raters in each of the 12 countries, under the supervision of NRCs. A check on the extent of agreement amongst raters was carried out by calculating the percentage agreement between each rater and the modal rating of all raters. This analysis showed 80-82% agreement for both populations.

Two structural changes were made to the provisional classification made by the Steering Committee:

Categories with fewer than 5 items were combined with another. (Pop A: Main Idea to Inference; Pop B Verbatim to Paraphrase)

Locating Information in Documents was subdivided to make an additional category in which students had to locate then calculate, compare or process the information.

(ii) Item Format

On item format, the ratios of different item types were weighted in favour of multiple-choice at each population level, with four completion items (requiring one or two words) in the Population A test and 20 in the Population B test.

Four items required an extended response in the students' own words.

(iii) Item Themes

In the Document and Expository Domains, items were classified by theme. The final Population A test contained 2 passages relating to the home, 11 to the school and 2 to society. For Population B, the figures were 3 for home, 8 for school, and 5 for society/work.

Although school themes were given more weight than was originally proposed, the Steering Committee accepted this balance as realistic in a test designed for students in school.

(iv) Range of Topics

One potential weakness of an international test designed to meet the needs of students in 30 countries, is that in trying to eliminate topics which are vetoed by any one of those countries, the test contains such a narrow range of topics that it has no face validity for any single country. To investigate this problem, a study was made of the topics included in a sample of six recently standardized reading tests, drawn from a variety of countries. Table 9 shows the range of topics included in the passages of these tests, alongside those of the Population A Reading-Literacy Test, final form. The 15 topics of the latter test appear to cover at least as wide a range as that of any single national test, designed for students of similar age. A similar comment could be made about the Population B test. The numbers in brackets refer to the number of items on each passage.

(v) Item Difficulty Levels

The test plan provided for a wide range of difficulty in the reading tasks prepared. To accommodate the needs of students in the developing countries, a number of relatively easy items were included at each age level. Indeed, the item selection criterion of a good Rasch fit internationally meant that difficult items rarely survived the pilot tests as they did not function effectively in countries where reading levels were low, and guessing was common.

Table 10 and 11 presents distributions of Rasch logit values for each item, broken down by domain in the two tests.

Table 9

Analysis of Topics Contained in IEA Population A Test and Six Standardized Reading Tests

IEA 9 YEAR TEST		HONG KONG STANDARDIZED READING TEST		STANFORD DIAGNOSTIC READING TEST (USA)	
Narrative		Narrative		Narrative	
Bird & Elephant	(5)	King's Judgement	(23)	Family Conservation About Milk	(5)
Boy & Dog	(6)	Talking Tree	(6)	Girl Finds Birds Egg	(5)
Greedy Shark	(5)	Catching a Robber	(6)	Boy Attends Scout Camp	(7)
Frail Old Man	(7)	The Taxi Driver	(20)	Expository	
Expository		Expository		Moving a Book-cart	(4)
Post-card	(2)	Cable Car	(15)	A Ghost Town	(9)
Walrus	(7)	How Bananas Grow	(6)	Cloze Completion	
Quicksand	(3)	Vocabulary	(15)	(24)	
Marmots	(4)	TOTAL	(85)	TOTAL	(60)
Age of Trees	(6)		Sentences		
Documents					
Map	(4)				
School Timetable	(3)	FIJI STANDARDIZED READING TEST			
Bar Graph	(4)	Narrative			
Bus Schedule	(4)	Boy's Ambition	(5)		
Table of Contents	(3)	Naughty Boy Punished	(6)		
Temperature Chart	(5)	Old Woman & Bird	(6)		
Word Recognition	(30)	Girl Shopping	(10)		
TOTAL	(98)	Expository			
		Krishna's Farm			
		Vocabulary	(30)		
		TOTAL	(62)		

ACER PRIMARY READING TEST		NZCER PROGRESSIVE ACHIEVEMENT		IOWA TEST OF BASIC SKILLS (USA)	
Narrative		Narrative		Narrative	
Pippi Longstocking	(5)	Maori Legend	(5)	A Dog That Sings	(6)
A Strange Creature	(6)	Tom Sawyer	(5)	Two Indian Boys	(9)
Elephant & Crocodile	(5)	Maori Legend	(5)	A Boy on a Pony	(10)
Expository		Expository		Expository	
Cranes	(5)	Meteors	(5)	A Metal Sharpener	(7)
Baby Seal	(5)	Hawks	(5)	Cave Drawings	(8)
Imaginary Rocket Trip	(6)	Beaver	(5)	Blacksmith at Work	(11)
Birds at Dawn	(5)	Climbing Perch	(5)	Queen Mary Liner	(8)
		White Heron	(5)	Potlatch Feasts	(9)
			Vocabulary	(38)	
TOTAL	(38)	TOTAL	(40)	TOTAL	(106)

Table 10

Distribution of International Logit Values for Items in Pop A Test

Logit Value	Narrative	Expository	Documents	Total
2.00 or more	0	2	3	5
1.50 to 1.99	1	1	2	4
1.00 to 1.49	2	4	0	6
0.50 to 0.99	5	4	3	12
0.00 to 0.49	3	1	2	6
-0.00 to -0.49	5	3	3	11
-0.50 to -0.99	3	1	4	8
-1.00 to -1.49	0	2	3	5
-1.50 to -1.99	1	1	2	4
-2.00 or less	0	1		1
Total	20	20	22	62

Table 11

Distribution of International Logit Values for Items in Pop B Test

Logit Value	Narrative	Expository	Documents	Total
2.00 - 2.49	0	0	1	1
1.50 to 1.99	1	0	3	4
1.00 to 1.49	3	3	3	9
0.50 to 0.99	5	5	4	14
0.00 to 0.49	6	3	5	13
-0.00 to -0.49	4	6	5	15
-0.50 to -0.99	1	3	5	9
-1.00 to -1.49	2	3	3	8
-1.50 to -1.99	3	1	3	7
-2.00 or less	1	0	0	1
Total 26	24	32	82	

(vi) Item Discrimination Levels

All test items showed positive discrimination within domains. Internationally, most produced a mean point-biserial index of between 0.35 and 0.50; only two fell below this level the Population A test, and none in the Population B test. In the few cases where an item in one country was found to have a deviant discrimination index - usually because of a misprint, or a fault in translation - it was identified as not fitting in Rasch scale, and was deleted from that country's data set. The Rasch scale was able to accommodate such omissions and still provide accurate score estimates. Tables 12 and 13 present the mean point biserial indices for all countries for Populations A and B.

Table 12

Distribution of Point-Biserial Indices for Items in Population A Test

Point Biserial Index	Narrative	Expository	Documents	Total
.60 or more	2	0	0	2
.55 - .59	4	5	1	10
.50 - .54	5	5	6	16
.45 - .49	7	2	5	14
.40 - .44	2	4	4	10
.35 - .39	0	1	4	5
.30 - .34	0	1	0	1
.25 - .29	0	1	1	2
Below .25	0	0	0	0
Total	20	19	21	60

Table 13

Distribution of Point-Biserial Indices for Items in Population B Test

Point Biserial Index	Narrative	Expository	Documents	Total
.55 or more	5	0	0	5
.50 - .54	3	2	1	6
.45 - .49	5	6	2	13
.40 - .44	3	7	9	19
.35 - .39	6	5	14	25
.30 - .34	4	4	6	14
Below .30	0	0	0	0
Total	26	24	32	82

(vii) International Equivalence of Items

One important feature of an acceptable international test is its item equivalence across nations. To what extent are the test items behaving in similar fashion in all countries? Are the most difficult in one country the most difficult in all countries?

To investigate this matter, the difficulty indices for all items were correlated with the mean difficulty index for all countries in an attempt to identify any items which deviated from the expected pattern. For this purpose, international and national p. values were used, but the Rasch logit scores would have shown similar patterns as the two indices were correlated at over 0.99.

Of course, the Rasch analysis itself provided confirmation that the test was functioning in similar fashion in each country. However, this additional analysis was undertaken to investigate the extent to which translation may have distorted the scores. If all English-speaking countries showed more similar patterns than those revealed in cross-language sets of countries, for instance, one would have to conclude that the translation process was having an influence on the scores.

Table 14 shows for each population the correlation between the difficulty values of each country on each item and the mean (international) difficulty value, based on a pooled random sample of 200 students from each country. The median r's for all countries were correlated and compared with the median r's for the six English-speaking countries.

Table 14

Correlations Between p Values for Each Country and International p values

Country	Pop A	Pop B
Bel/Fr	.96	.92
Bot	-	.77
Can/BC	.96	.92
Cyp	.94	.91
Den	.93	.93
Fin	.91	.85
Fra	.96	.91
Ger/E	.93	.90
Ger/W	.94	.92
Gre	.94	.89
HK	.90	.82
Hun	.91	.88
Ice	.94	.89
Ire	.95	.94
Ita	.93	.91
Net	.95	.91
NZ	.95	.94
Nig	-	.85
Nor	.95	.89
Phi	-	.84
Por	.94	.87
Sin	.91	.87
Slo	.95	.90
Spa	.96	.86
Swe	.96	.91
Swi	.96	.93
Tha	-	.74
T/T	.95	-
USA	.91	.91
Ven	.87	.89
Median	.94	.90
Median English countries	.95	.91

The median correlations in Table 13 are sufficiently high, with minor exceptions, to generate confidence that the translation process did not have a marked effect on the results, and that the tests were functioning in similar fashion in each education system. The median r's for English-speaking countries were only slightly higher than those for all countries.

Test Statistics

This section outlines data on the reliability, validity and domain intercorrelations of the final tests.

(1) Reliability of Test Scores

Most conclusions drawn in this survey were based on grouped data, often of many hundreds of students, so the question of individual score reliability is of lesser significance. However, the individual reliability figures are relevant for some of the multivariate analyses and are given below for domains (KR 20) and for total scores (Cronbach Alpha and Equivalent Forms).

Table 15

Reliability Coefficients for Domains and Total Scores: Pop A

Country	Narrative KR-20	Expository KR-20	Documents KR-20	Total Cronbach Alpha	Total Equivalent Forms
Bel/Fr	.83	.79	.77	.91	.84
Can/BC	.85	.83	.79	.93	.87
Cyp	.85	.80	.78	.92	.88
Den	.91	.91	.89	.96	.88
Fin	.81	.75	.75	.90	.85
Fra	.83	.77	.72	.90	.84
Ger/E	.85	.87	.82	.93	.87
Ger/W	.85	.86	.82	.94	.86
Gre	.83	.78	.78	.92	.84
HK	.84	.73	.77	.91	.85
Hun	.82	.85	.79	.93	.89
Ice	.86	.85	.81	.94	.90
Ind	.73	.69	.68	.87	.84
Ire	.86	.81	.79	.93	.90
Ita	.84	.82	.80	.93	.86
Net	.83	.81	.77	.91	.84
NZ	.88	.82	.82	.94	.91
Nor	.89	.86	.84	.95	.90
Por	.81	.77	.79	.91	.84
Sin	.84	.73	.76	.92	.90
Slo	.84	.83	.80	-	-
Spa	.84	.83	.80	-	-
Swe	.87	.88	.85	.95	.91
Swi	.85	.85	.80	.93	.88
T/T	.85	.82	.79	.93	.90
USA	.86	.77	.73	.92	.87
Ven	.78	.76	.73	.89	.78
Median	.84	.81	.79	.93	.87

Most domain score reliability coefficients for Population A clustered around 0.80 to 0.85 while the coefficients for the total scores clustered around 0.90 - 0.95 (Cronbach alpha) or 0.85 to 0.90 (Equivalent halves). The coefficients for equivalent halves were estimated by correlating scores from the two testing sessions, and correcting by the Spearman Brown formula. The coefficients were typically lower for low-scoring countries, where guessing appeared to be common.

Table 16

Reliability Coefficients for Domains and Total Scores: Pop B

Country	Narrative KR-20	Expository KR-20	Documents KR-20	Total Cronbach Alpha	Total Equivalent Forms
Bel/Fr	.87	.80	.82	.93	.87
Bot	.48	.46	.73	.77	.80
Can/BC	.86	.82	.81	.93	.87
Cyp	.84	.82	.81	.93	.91
Den	.87	.82	.81	.93	.94
Fin	.82	.70	.72	.90	.90
Fra	.82	.74	.74	.89	.88
Ger/E	.88	.81	.79	.92	.86
Ger/W	.86	.79	.77	.93	.90
Gre	.81	.78	.77	.91	.82
HK	.80	.75	.71	.90	.88
Hun	.83	.79	.78	.92	.92
Ice	.86	.82	.81	.93	.92
Ire	.88	.82	.83	.94	.91
Ita	.85	.78	.81	.92	.89
Net	.86	.79	.83	.93	.84
NZ	.89	.85	.84	.95	.95
Nig	.77	.73	.84	.91	.76
Nor	.83	.79	.80	.92	.82
Phi	.79	.77	.82	.92	.90
Por	.79	.73	.73	.89	.88
Sin	.78	.74	.71	.89	.90
Slo	.79	.71	.71	.89	.85
Spa	.85	77	75	.91	.87
Swe	.86	80	81	.93	.91
Swi	.86	.79	.77	.92	.91
Tha	.86	.78	.84	.93	.72
T/T	.88	.85	.86	.95	.92
USA	.88	.84	.80	.94	.92
Ven	.79	.76	.72	.89	.70
Zim	.70	.69	.83	.89	.85
Median	.85	.79	.80	.92	.89

For Population B the domain reliabilities clustered around 0.80 to 0.85, except for a few countries with very low scores, where guessing appeared to be common. Again the reliability of the total scores was higher, around 0.85 to 0.95 in most countries.

LESSONS LEARNED FROM THE EXPERIENCE OF CONDUCTING A MULTI-NATIONAL FOUR-YEAR STUDY

Much was learned by those who participated in the study, about both process and products. Apart from the hard data produced on the international means and standard deviations, we revealed many unexpected findings, and confirmed many hypotheses about which there was only uncertainty before.

For instance, the consistent superiority of Finnish students across all subtests and most items; the repeated discrepancy in gender scores, at age 9, across all countries, favouring girls; the consistently negative relationship between television viewing hours and reading achievement, except in those countries where children regularly read sub-titles on the screen as they view foreign films; the importance of ready access to a range of books in all countries; the large gaps still existing between developed countries and the lesser-developed countries; the fact that Singapore children can achieve at high levels, though learning in a second language at the outset; the fact that countries which begin reading instruction at age 7 are amongst the highest achievers at age 9.

These and many other important outcomes are reported in the various national and international reports.

As for lessons about process, we learned much about the need for lengthy consultation about aims, models, procedures before decisions were made. There was never enough time at international meetings, and in retrospect many NRCs regretted this fact. This constraint was an inevitable consequence of insufficient funding for the project, a variable not uncommon in IEA studies.

Many NRCs would have like more opportunity for greater levels of interaction between Steering Committee, Technical Committee and National Research Coordinators. Several felt the need to develop, refine, and share a clearer conception of the theoretical model and hypotheses at the outset. Of course, there will probably always be some tension between national and international needs.

It was unfortunate that some ideal aspects of the design had to be dropped for lack of support. For instance, the Steering Committee believed that it would have enhanced the design to include pretests at the beginning of the school year. It would also have been preferable to conduct a second pilot study on the test items so that revised items could be trialled. This would have enabled the testers to choose from a greater range of test items, and to check further on the time allowances - which become a source of contention later on. The questionnaires also needed more trialling than they received, so that more work could be done on latent traits.

There was general agreement that the use of the Rasch Scale was a great step forward in identifying poor-fitting items, and establishing underlying traits. We also felt that Thorndike's

claim about the relatively minor intrusion of translation factors was confirmed. The translated items did behave similarly in most cases.

The fact that there were few generalizable findings which emerged from the multiple regression studies is partly a reflection of the fact that literacy is more firmly embedded in the diverse cultures of the various countries which participated - and is therefore less likely to throw up uniform common denominators. But it is also, we suspect, partly a function of the insensitivity of questionnaires to real differences in input variables in the classroom. There is a strong case for using more direct observational methods in classrooms if these studies are to reveal more useful findings about associations between teacher behaviour and school policies on the one hand and achievement levels on the other. The use of hierarchical linear modelling techniques may yet reveal more links than have been exposed to date.

REFERENCES

Anderson, R.C. and Pearson, D. (1984) A schema-theoretic view of basic processes in reading comprehension, in Pearson, D., Barr, R., Kamil, M. and Mosenthal, P. (Eds) Handbook of Reading Research, Vol 1. White Plains, New York, Longman.

Anderson, T.H., and Armbruster, B. (1982) Reader and text studying strategies,in W. Otto and S. White (Eds) Reading Expository Material. New York, Academic Press.

Andrich, D. and Godfrey, J.R. (1979). Hierarchies in the Skills of Davis' Reading Comprehension Test, Form D. An Empirical Investigation Using a Latent Trait Model. Reading Research Quarterly, 14, 182-200.

Archer, A.J. and Elley, W.B. (1993) Do print size and line length affect children's reading comprehension? Research in Education. 49, 75-76.

Asher, S.R. (1980) Topic interest and children's reading comprehension, in R.J. Spiro, B. Bruce and W.F. Brewer (Eds) Theoretical Issues in Reading Comprehension. Hillsdale, NJ, Erlbaum.

Barr, R., Kamil, Mosenthal, P. and Pearson, D. (Eds) (1991) Handbook of Reading Research, Vol II. New York, Longman.

Bartlett, B.J. (1978) Top-level structure as an organisational strategy for recall of classroom text. PhD Dissertation. Arizona State University, Tempe.

Brewer, W.F. (1980) Literary theory, rhetoric and stylistics; Implications for psychology. In R. Spiro, B. Bruce and W.F. Brewer (Eds) Theoretical Issues in Reading Comprehension. Hillsdale, NJ, Erlbaum.

Brooks, C. and Warren, R.P. (1979) An analysis of story comprehension in elementary school children, in R.O. Freedle (Ed) New Directions in Discourse Processing, Vol II. Norwood, NJ, Ablex.

Carroll, J.B. (1988). The NAEP Reading Proficiency Scale is Not a Fiction. Phi Delta Kappan, 69, 10, 761-4.

Choppin, B.H. and Purves, A.C. (1969). A Comparison of Open-Ended and Multiple-Choice Items Dealing with Literacy Understanding. Research in the Teaching of English, 3, 1, 15-24.

Clifford, G.J. (1984). Buch und Lesen: Historical Perspectives on Literacy and Schooling. Review of Educational Research, 54, 4, 472-500.

Danserau, D.F. (1978) The development of a learning strategy curriculum, in H.F. O'Neill Jr (Ed) Learning Strategies. New York, Academic Press.

Davey, B. and Macready, G.B. (1990). Applications of Latent Class Modelling to Investigate the Structure Underlying Reading Comprehension Items. Applied Measurement in Education, 3, 3, 209-229.

Davis, F. (1968). Research in Comprehension in Reading. Reading Research Quarterly, 3, 499-545.

Elley, W.B. and Mangubhai, F. (1992). Multiple-Choice and Open-Ended Items in Reading Tests: Same or Different? Studies in Educational Evaluation, 18, 191-199.

Elley, W.B. and Reid, N.A. (1969) Progressive Achievement Tests of Reading Teachers' Manual. Wellington, New Zealand Council for Educational Research.

Elley, W.B. (Ed.) (1994). The IEA Study of Reading Literacy: Achievement and Instruction in 32 School Systems. Oxford, Pergamon.

Farr, R. and Carey (1986). Reading: What Can be Measured? Newark, DE: International Reading Association.

Freebody, P. and Anderson, R.C. (1983). Effects of vocabulary difficulty, text cohesion and schema availability on reading comprehension. Reading Research Quarterly, 18, 277-294.

Gil, G.A. (1990). Study of the Results in the Word Recognition Test on the Pilot Test of the Reading Literacy Study. Madrid: Center for Educational Research and Documentation.

Gorman, T., Purves, A.C. and Degenhart, E. (1988). The IEA Study of Written Composition, I: The International Writing Tasks and Scoring Scales. Oxford: Pergamon.

Graesser, A., Golding, J. and Long, D. (1991) Narrative Representation and Comprehension, in R. Barr et al (Eds) Handbook of Reading Research, Vol II. New York, Longman.
Guthrie, J.T. and Kirsch, I.S. (1984). The Emergent Perspective on Literacy. Phi Delta Kappan. Jan., 351-5.

Graff, H.J. (Ed.) (1981). Literacy and Social Development in the West: A Reader. Cambridge. UK: Cambridge University Press.

Guthrie, J.T. and Kirsch, I.S. (1987). Distinctions Between Reading Comprehension and Locating Information in Text. Journal of Educational Psychology, 79, 220-8.

Guthrie, J. (1988) Locating information in documents: Examination of a cognitive model. Reading Research Quarterly, 23, 2, 178-199.

Haertel, E. (1980). Determining What is Measured by Multiple-Choice Tests of Reading Comprehension. Unpublished Ph.D. Dissertation, University of Chicago.

Hillocks, G. and Ludlow, L. (1984). A Taxonomy of Skills in Reading and Interpreting Fiction. American Educational Research Journal, 21, 7-24.

Hulin, C.L. (1987). A Psychometric Theory of Evaluations of Item and Scale Translations. Journal of Cross-Cultural Psychology, 18, 2, 115-42.

Kintsch, W. (1977). On Comprehending Stories. In M.A. Just and P.A. Carpenter (Eds) Cognitive Processes in Comprehension. Hillsdale, N.J. Erlbaum.

Kirsch, I.S. and Jungeblut, A. (1986). Literacy: Profiles of America's Young Adults. Princeton, N.J. Educational Testing Service.

Kirsch, I. and Mosenthal, P. (1990) Exploring document literacy: Variables underlying the performance of young adults. Reading Research Quarterly. 25, 5-31.

Levine, K. (1986). The Social Context of Literacy. London: Routledge and Kegan Paul.

Lunzer, E. and Gardner, K. (Eds) (1979) The Effective Use of Reading. London, Heinemann.

Mandler, J.M., Johnson, N.S. (1977). Remembrance of Things Parsed: Story Structure and Recall. Cognitive Psychology, 1977, 9, 111-51.

McCauley, D. and Colberg, M. (1983) Transportability of deductive measurement across cultures. Journal of Educational Measurement. 20, 1, 81-92.

McLean, L.D. and Goldstein, H. (1988). The U.S. National Assessments in Reading: Reading Too Much into the Findings. Phi Delta Kappan. Jan. 369-72.

Modu, C.C. and May, G. (1977). Exploring the Use of Shorter Reading Comprehension Passages in the SAT-V. Princeton, N.J. Educational Testing Service, Report SR-77-33.

Reid, N.A., Croft, A.C. and Jackson, P.F. (1978) Progressive Achievement Tests of Study Skills: Teachers' Manual. Wellington, New Zealand Council for Educational Research.

Spearritt, D. (1972). Identification of Subskills in Reading Comprehension by Maximum Likelihood Factor Analysis. Reading Research Quarterly, 8, 92-111.

Steffenson, M.S., Joag-dev, C. and Anderson, R.C. (1979). A cross-cultural perspective on reading comprehension. Reading Research Quarterly, 15, 10-29.

Stenner, A., Horrabin, I., Smith, D. and Smith, M. (1988) Most comprehension tests do measure reading comprehension. Phi Delta Kappan. June, 765-7.

Thorndike, R.L. (1973). Reading as Reasoning. Reading Research Quarterly, 9, 135-47.

Thorndike, R.L. (1973). Reading Comprehension Education in 15 Countries. Uppsala, Almquist and Wiksell.

Thorndyke, P.W. (1977). Cognitive Structures in Comprehension and Memory of Narrative Discourse. Cognitive Psychology, 9, 77-110.

Tinker, M.A. (1966). Experimental Studies on the Legibility of Print: An Annotated Bibliography. Reading Research Quarterly, 1, 67-118.

Traub, R.E. and Fisher, C.W. (1977). On the Equivalence of Constructed Response and Multiple-Choice Tests. Applied Psychological Measurement, 1, 355-369.

Van den Bergh, H. (1988). Examens Geexamineerd. Gravenhage, Netherlands. S.V.O.

Vernon, P.E. (1962). The Determinants of Reading Comprehension. Educational and Psychological Measurement, 22, 269-286.

Wagner, D.A. (1990). Literacy Assessment in the Third World. Comparative Education Review, 34, 1, 112-138.

Watts, L. and Nisbett, J. (1974). Legibility of Children's Books, Sussex, UK. NFER Publishing Co.

Weaver, C. and Kintsch, W. (1991) Expository Text, in R. Barr, M. Kamil, P. Mosenthal and D. Pearson (Eds) Handbook of Reading Research, Vol II. New York, Longman.

Wolf, R.M. (1969) The Analysis of Item Data for the IEA Mathematics Test. International Review of Education. 15, 2, 237-241.

Wright, P. (1988) Writing technical information, in E.Z. Rothkopf (Ed) Review of Research in Education. 14, 327-385.

Zachrisson, B. (1965). Studies in the Legibility of Printed Text. Stockholm: Almquist and Wiksell.

Zwick, R. (1987). Assessing the Dimensionality of NAEP Reading Data. Journal of Educational Measurement, 24, 4, 293-308.

Chapter 4

Sample Design Procedures for the International Study of Reading Literacy

Kenneth N. Ross

DEFINITION OF TARGET POPULATIONS

The desired target population definitions for the IEA Reading Literacy Study were prepared by the International Steering Committee following consultation with IEA member countries concerning the research objectives of the study. It was the responsibility of National Research Centres to employ these definitions as guidelines for the preparation of their own defined target populations. The defined national target population represented a subset of the students described by the desired international target population. The difference between these two populations was referred to as the excluded population.

In educational research studies conducted by the IEA, the desired international target population has always focussed on schools that operate within mainstream national education systems. That is, the word "school" in the definition of this population has referred to those schools that, because of their curriculum, medium of instruction, and/or teaching methods, are deemed to provide educational environments that are quite typical of most schools in the system. This decision has resulted in the allocation of "different" schools (for example, schools for the handicapped, correspondence schools, and radio education schools) to the excluded population.

The Desired Target Populations

There were two desired target populations for this study: Population A and Population B.

Population A: All students attending school on a full-time basis at the Grade level in which most students aged 9:00-9:11 years (during the first week of the eighth month of the school year) are enrolled.

Population B: All students attending school on a full-time basis at the Grade level in which most students aged 14:00-14:11 years (during the first week of the eighth month of the school year) are enrolled.

In order to identify the correct grade level for each population, the National Research Coordinators (NRC's) for each country were required to prepare "Age-by-Grade" tables. An example of an Age-by-Grade table for Population A in Hong Kong has been reproduced in **Table 1**.

Table 1: Age-by-Grade Table for Population A in Hong Kong

Age	Grade					Total (%)
	Grade 2	Grade 3	Grade 4	Grade 5	Other	
8:00-8:11	5.65	72.57	21.41	0.00	0.37	100.00
9:00-9:11	0.87	7.97	72.45	18.65	0.06	100.00
10:00-10:11	0.17	1.46	10.99	71.12	16.26	100.00

From Table 1 it can be seen that 72.45 percent of Hong Kong students aged 9:00-9:11 (during the first week of the eighth month of the school year) were in Grade 4, a further 18.65 percent were in Grade 5, and the remaining 8.90 percent were located in other grades. From this distribution it was clear that Grade 4 was the appropriate grade level for the Population A defined target population in Hong Kong.

Some Key Decisions About the Specification of Target Populations

The International Steering Committee made two important decisions concerning the specification of the desired target populations. These decisions had implications for the conduct of the sampling and also for the interpretation of research results.

First, it was decided that the desired target populations should be focussed on the first week of the eighth month of the school year. This decision was taken because the Steering Committee wanted to have the testing date for the study fixed as close as possible to a period of the school year when schools had settled into a regular pattern of work. The eighth month of the school year was considered to be a time when schools were in "full flight" - yet it was considered to be not so close to the end of the school year as to create problems because of the proximity of examinations and annual holidays.

Second, it was decided that the desired target populations should be grade-based rather than age-based, and that one (or sometimes two) intact class(es) would be sampled in each selected school. This second decision was taken in order to simplify administrative procedures by ensuring (a) that the sample members for any one population were associated with only one year level within each country, and (b) that one (or at most two) teacher(s) would be involved in the study for each selected school. The second decision also implied that no subsampling of students would be conducted across classes in selected schools. This improved the possibility of having sufficiently large numbers of sample members per class to be able to employ multi-level data analysis procedures.

It is important to note that the decision to employ a grade-based approach automatically resulted in the need for great caution in the interpretation of differences between national mean scores. This occurred because there were differences between countries with respect to the distribution of age groups across the various grade levels. These differences arose from varying national policies concerning the general structure of primary schooling — particularly those policies concerned with school commencement age and grade promotion policies. In Table 2 the mean ages of the sample members and the grade level tested have been presented for each country and for Populations A and B.

Defined National Target Populations

National Research Coordinators (NRC's) were required to have their defined national target population descriptions approved by the IEA Sampling Co-ordinator. If, in a particular country, there were large differences between the desired international target population and the defined national target population, then the International Co-ordinator had the power to recommend exclusion from the international reports.

The defined national target population descriptions for each country were required to include the following pieces of information:

1. A description of the defined national target population.

2. A description of the contents of the excluded population.

3. A tabular description of the numbers of schools and students in the desired, defined, and excluded populations - listed according to the strata that were to be used for the sample design.

The tabular descriptions of the desired , defined, and excluded populations have been reproduced for each country in the Appendices of this technical report.

Table 2: Mean Age of Sample Members and Grade Level Tested

COUNTRY	Population A		Population B	
	Grade Tested	Mean Age	Grade Tested	Mean Age
Belgium (F)	4	9.8	8	14.3
Botswana	-	-	9	14.7
Canada (BC)	3	8.9	8	13.9
Cyprus	4	9.8	9	14.8
Denmark	3	9.8	8	14.8
Finland	3	9.7	8	14.7
France	4	10.1	9	15.4
Germany (E)	3	9.5	8	14.4
Germany (W)	3	9.4	8	14.6
Greece	4	9.3	9	14.5
Hong Kong	4	10.0	9	15.2
Hungary	3	9.3	8	14.1
Iceland	3	9.8	8	14.8
Indonesia	4	10.8	-	-
Ireland	4	9.3	9	14.5
Italy	4	9.9	8	14.1
Netherlands	3	9.2	8	14.3
New Zealand	5	10.0	10	15.0
Nigeria	-	-	9	15.3
Norway	3	9.8	8	14.8
Philippines	-	-	8	14.5
Portugal	4	10.4	9	15.6
Singapore	3	9.3	8	14.4
Slovenia	3	9.7	8	14.7
Spain	4	10.0	8	14.2
Sweden	3	9.8	8	14.8
Switzerland	3	9.7	8	14.9
Thailand	-	-	9	15.2
Trinidad	4	9.6	9	14.5
U.S.A.	4	10.0	9	15.0
Venezuela	4	10.1	9	15.5
Zimbabwe	-	-	9	15.5

SAMPLING FRAMES

National Research Coordinators were required to prepare a sampling frame for each of their defined target populations. The sampling frame provided a means of "listing" the students in the defined target population without actually creating a physical list consisting of an entry for each and every student.

Sampling frames that (a) permitted schools to be selected as the first stage of sampling, and (b) consisted of accurate information concerning the numbers, locations, and the defined target population enrolments of schools for each stratum, were developed by all but four countries involves in the study. These types of sampling frames were referred to as single-level sampling frames.

In situations where a comprehensive national list of schools was not available, a double-level sampling frame was required. This was the case in the USA (where Census Counties were used as a first stage of sampling), Switzerland (where Cantons were used), Indonesia (where Education Regions were used), and Thailand (where Education Districts were used).

In operational terms, the single level sampling frames consisted of a separate list of schools for each stratum of the defined target population. Each school was represented by one entry on one of these lists. Each list consisted of entries that contained the following information: (a) school identification name or number, and (b) a measure of school size that was equal to the total number of students in the defined target population.

The first level of the double level sampling frames employed by the USA, Switzerland, Indonesia, and Thailand consisted of a separate list of geographical areas (GA's) for each stratum of the defined target population. Each GA was represented by one entry on one list. Each list consisted of entries that contained the following information: (a) GA identification name or number, and (b) a measure of GA size assessed by the total number of students in the defined target population.

The second level of the double level sampling frames were composed of a list of schools for each GA selected at the first stage of sampling. Each school in the selected GA's was represented by one entry on each list. These lists consisted of entries that contained the following information: (a) school identification name or number, and (b) a measure of school size assessed by the total number of students in the defined target population.

SELECTION AND APPLICATION OF STRATIFICATION VARIABLES

The technique of stratification is usually employed in the preparation of sample designs because it provides increased accuracy in sample estimates for important criterion variables without leading to substantial increases in costs. Stratification does not imply any departure from probability sampling - it simply requires a prior division of the population into subpopulations called strata followed by probability sampling conducted independently within each stratum. The sample estimates of population characteristics are then obtained by combining the information from each stratum. The gains in sampling accuracy that follow from stratification depend upon the extent to which variation in the variable under investigation is less within strata than it is for the population as a whole. The stratification variables used in this study were mostly selected from the following list of six variables.

1. Geographical Location (region, province, or state).
2. Urbanization (urban/rural).
3. Type of School (government/non-government).
4. School Size (large/medium/small).
5. School Program (academic/vocational).
6. Gender Composition (male/female/co-educational).

The Number of Strata

Stratification provides an effective means of increasing sampling precision and therefore the use of too few strata was discouraged. The maximum number of strata was governed by the need to have two primary sampling units per stratum in order to permit the calculation of sampling errors. Since most NRC's were selecting schools at the first stage of sampling, it was recommended to NRC's that the maximum number of strata should not exceed half the number of schools.

Selection of Stratification Variables

In Table 3 the selection of stratification variables has been summarized for all countries. The most popular variable for creating strata was "Geographical Location". The subclasses of this variable were usually states, provinces, or administrative regions. The second most popular variable was "Urbanization" —which was mostly applied as a simple Urban/Rural dichotomy. Other popular stratification variables included "Type of School" (mostly applied as Government/Private), "School Size" (mostly applied as several broad size classifications like Small/Medium/Large), School Program (Vocational/Academic), and Gender Composition of School (Male, Female, Co-educational).

Table 3: Summary of Stratification Variables

Country	Pop.	No. of Strata	Stratification Variable						
			Geographical Location	Urbanisation	Type of School	School Size	School Program	Gender	Other
Belgium	Both	15	X		X				
Botswana (BC)	B	14	X						
Canada	Both	18	X		X				Achievement
Cyprus	Both	16	X	X		X			
Denmark	Both	14	X						
Finland	Both	9	X	X					
France	A	17	X	X		X			
	B	16	X	X		X			
Germany (E)	Both	15	X						
Germany (W)	A	10	X						
	B	40					X		
Greece	Both	15	X	X	X				
Hong Kong	Both	27	X		X				Achievement
Hungary	Both	13	X						
Iceland	Both	24	X	X					
Indonesia	A	18	X	X					
Ireland	A	7				X		X	
	B	7							
Italy	Both	20	X	X					
Netherlands	A	30				X			Socio-Economic
	B	25				X	X		
New Zealand	A	24	X				X		
	B	16		X	X			X	
Nigeria	B	18	X	X					
Norway	Both	18		X		X			
Philippines	Both	26	X		X				
Portugal	Both	17	X	X	X				

Table 3 (Continued): Summary of Stratification Variables

Country	Pop.	No. of Strata	Stratification Variable						
			Geographical location	Urbanisation	Type of School	School Size	School Program	Gender	Other
Singapore	A	9			X	X		X	
	B	11			X	X		X	
Slovenia	Both	70							Zones
Spain	Both	42	X	X	X	X			
Sweden	Both	24	X						
Swizerland	A	48	X	X					
	B	21	X				X		
Thailand	Both	26	X		X				
Trinidad / Tobago	Both	24	X		X				
USA	A	26	X	X					Socio-economic
	B	21	X	X					Socio-economic
Venezuela	A	27	X	X	X				
	B	18	X	X	X				
Zimbabwe	B	18	X				X		

STAGES OF SAMPLING

The general sample design framework for this study consisted of a stratified two-stage cluster sample design in which schools were selected within strata at the first stage with probability proportional to size, followed by the selection within each sample school of one (or two) intact class(es) of students with probability proportional to size.

In **Table 4** the various stages of sampling have been described for each country. For example, in Belgium there were two stages of sampling. At the first stage a probability proportional to size (PPS) selection of schools was made. At the second stage a probability proportional to size (PPS) selection of one intact class was made for each selected school. The sampling approach adopted in Belgium was used by most of the countries involved in the study. Three countries (Thailand, Trinidad/Tobago, and Zimbabwe) used simple random sampling (SRS) to sample students in selected schools.

From **Table 4** it can be seen that all countries except Indonesia, Switzerland, Thailand, and the USA employed single level sampling frames in which schools (and in West Germany classes) were selected at the first stage of sampling. In Iceland, total populations were tested and therefore no sampling stages were required. In Cyprus (Population B), Trinidad/Tobago (Population B), and Singapore (Population A and B), all schools in the population were selected at the first stage. The majority of countries used PPS sampling to select intact classes at the final stage of sampling — the only exceptions were France (Population B), Thailand (Population A and B), Trinidad/Tobago (Population A and B), and Zimbabwe (Population A and B).

Table 4: Summary of Sampling Stages

Country	Pop.	Levels in Sampling Frame	Sampling Stages			
			Geographical Areas	Schools	Classes	Students
Belgium	Both	Single	-	PPS	PPS	Intact Class
Botswana	B	Single	-	PPS	PPS	Intact Class
Canada(E)	A	Single	-	PPS	PPS	Intact Class
	B	Single	-	PPS	PPS	Intact Class
Cyprus	A	Single	-	PPS	PPS	Intact Class
	B	Single	-	All Schools	-	SRS
Denmark	Both	Single	-	PPS	PPS	Intact Class
Finland	Both	Single	-	PPS	PPS	Intact Class
France	A	Single	-	PPS	PPS	Intact Class
	B	Single	-	PPS	-	SRS
Germany(E)	Both	Single	-	SRS	SRS	Intact Class
Germany(W)	Both	Single	-	-	SRS	Intact Class
Greece	Both	Single	-	PPS	PPS	Intact Class
Hong Kong	Both	Single	-	PPS	PPS	Intact Class
Hungary	A	Single	-	PPS	PPS	Intact Class
	B	Single	-	PPS	PPS	Intact Class
Iceland	Both	-	-	-	-	-
Indonesia	A	Double	PPS	PPS	PPS	Intact Class

Table 4 (Continued): Summary of Sampling Stages

Country	Pop.	Levels in Sampling Frame	Sampling Stages			
			Geographical Areas	Schools	Classes	Students
Ireland	A	Single	-	PPS	PPS	Intact Class
	B	Single	-	PPS	PPS	Intact Class
Italy	Both	Single	-	PPS	PPS	Intact Class
Netherlands	Both	Single	-	PPS	PPS	Intact Class
New Zealand	Both	Single	-	PPS	PPS	Intact Class
Nigeria	B	Single	-	PPS	PPS	Intact Class
Norway	Both	Single	-	PPS	PPS	Intact Class
Philippines	Both	Single	-	PPS	PPS	Intact Class
Portugal	Both	Single	-	PPS	PPS	Intact Class
Singapore	Both	Single	-	(All Schools)	PPS	Intact Class
Slovenia	Both	Single	-	PPS	PPS	Intact Class
Spain	Both	Single	-	PPS	PPS	Intact Class
Sweden	Both	Single	-	PPS	PPS	Intact Class
Switzerland	Both	Double	PPS	-	PPS	Intact Class
Thailand	Both	Double	-	PPS	-	SRS

Table 4 (Continued): Summary of Sampling Stages

Country	Pop.	Levels in Sampling Frame	Sampling Stages			
			Geographical Areas	Schools	Classes	Students
Trinidad / Tobago	A	Single	-	PPS	PPS	SRS
	B	Single	-	(All Schools)	-	SRS
USA	Both	Double	PPS	PPS	PPS	Intact Class
Venezuela	Both	Single	-	PPS	PPS	Intact Class
Zimbabwe	Both	Single	-	PPS	-	SRS

RESPONSE RATES

In large-scale educational research studies that require the co-operation of many schools, it is inevitable that some schools in the designed sample will be unable to participate. For this reason several countries prepared two extra "replacement samples" which were selected in parallel with the original designed sample of schools. The use of replacement samples was required only in a few countries because of the excellent level of participation received from schools in the initial sample.

The number of schools in the designed sample was known prior to the data collection because this number was established during the preparation of the sample design. The number of schools in the achieved sample was established after the data collection had been completed because this number referred to schools that participated in the study and also returned appropriate data to the National Centres. The response rates for schools have been set out for each country and each population in **Table 5**.

The interpretation of the response rates in **Table 5** requires tabulation of the numbers of schools selected in the first draw, the first replacement sample, and the second replacement sample. These figures have been presented for Population A and Population B in **Table 6** and **Table 7**. It can be seen that the overall response rates were excellent across most countries. However it can alsoâ be seen that the high overall response rates for four countries: Belgium (FR), Canada (BC), Italy, and the Netherlands required the use of quite large numbers of replacement schools. The number of replacement schools was also quite high in West Germany for Population A.

Table 5: Overall Response Rates for Schools

Country	Population A			Population B		
	Designed	Achieved	%	Designed	Achieved	%
Belgium	150	149	99	153	144	94
Botswana	---	---	---	137	137	100
Canada (BC)	157	157	100	199	168	84
Cyprus	182	181	99	52	52	100
Denmark	166	164	99	140	139	99
Finland	71	71	100	71	71	100
France	140	136	97	140	136	97
Germany (E)	100	100	100	100	100	100
Germany (W)	150	150	100	200	196	98
Greece	176	175	99	148	147	99
Hong Kong	167	167	100	158	158	100
Hungary	144	144	100	144	144	100
Iceland (Total Population)	129	129	100	114	114	100
Indonesia	176	174	99	---	---	---
Ireland	134	122	91	162	151	93
Italy	177	154	87	175	173	99
Netherlands	91	91	100	174	162	93
New Zealand	177	176	99	125	124	99
Nigeria	---	---	---	136	80	59
Norway	186	186	100	142	138	97
Philippines	---	---	---	245	244	99
Portugal	165	145	88	136	130	96
Singapore	206	206	100	142	142	100
Slovenia	140	140	100	139	139	100
Spain	296	296	100	308	308	100
Sweden	124	123	99	149	149	100
Switzerland	229	225	98	323	322	99
Thailand	---	---	---	217	139	64
Trinidad & Tobago	184	182	99	93	93	100
United States	191	165	86	192	165	86
Venezuela	181	161	89	178	162	91
Zimbabwe	---	---	---	192	143	75

Table 6: The Structure of the Achieved Sample in Terms of First Draw, First Replacement and Second Replacement Schools (Population A).

Country	Designed Sample	First Draw	First Replacement	Second Replacement	Achieved Sample
Belgium	150	109	35	5	149
Canada/BC	157	125	26	6	157
Cyprus	182	181	--	--	181
Denmark	166	164	--	--	164
Finland	71	71	--	--	71
France	140	136	--	--	136
Germany/E	100	90	9	1	100
Germany/W	150	133	16	1	150
Greece	176	175	--	--	175
Hong Kong	167	167	--	--	167
Hungary	144	144	--	--	144
Iceland	129	129	--	--	129
Indonesia	176	174	--	--	174
Ireland	134	122	--	--	122
Italy	177	126	28	--	154
Netherlands	91	59	25	7	91
New Zealand	177	164	11	1	176

Table 6 (Continued): The Structure of the Achieved Sample in Terms of First Draw, First Replacement and Second Replacement Schools (Population A).

Country	Designed Sample	First Draw	First Replacement	Second Replacement	Achieved Sample
Norway	186	186	--	--	186
Portugal	165	145	--	--	145
Singapore	206	206	--	--	206
Slovenia	140	140	--	--	140
Spain	296	296	--	--	296
Sweden	124	123	--	--	123
Switzerland	229	223	2	--	225
Trinidad & Tobago	184	182	--	--	182
United States	191	165	--	--	165
Venezuela	181	158	3	--	161

Table 7: The Structure of the Achieved Sample in Terms of First Draw, First Replacement and Second Replacement Schools (Population B).

Country	Designed Sample	First Draw	First Replacement	Second Replacement	Achieved Sample
Belgium/FR	153	80	39	25	144
Botswana	137	137	--	--	137
Canada/BC	199	167	1	--	168
Cyprus	52	52	--	--	52
Denmark	140	139	--	--	139
Finland	71	71	--	--	71
France	140	136	--	--	136
Germany/E	100	79	19	2	100
Germany/W	200	160	36	--	196
Greece	148	147	--	--	147
Hong Kong	158	158	--	--	158
Hungary	144	144	--	--	144
Iceland	114	114	--	--	114
Ireland	162	151	--	--	151
Italy	175	147	20	6	173
Netherlands	174	92	53	17	162
New Zealand	125	118	5	1	124

Table 7 (Continued): The Structure of the Achieved Sample in Terms of First Draw, First Replacement and Second Replacement Schools (Population B).

Country	Designed Sample	First Draw	First Replacement	Second Replacement	Achieved Sample
Nigeria	136	80	-	-	80
Norway	142	138	-	-	138
Philippines	245	244	-	-	244
Portugal	136	130	0	0	130
Singapore	142	142	-	-	142
Slovenia	139	139	-	-	139
Spain	308	308	-	-	308
Sweden	149	148	1	-	149
Switzerland	323	300	16	6	322
Thailand	217	139	-	-	139
Trinidad & Tobago	93	93	-	-	93
United States	192	165	-	-	165
Venezuela	178	162	-	-	162
Zimbabwe	192	143	-	-	143

SAMPLING WEIGHTS

The sample designs used by most countries participating in the study were based on the single level sampling frames described earlier in this chapter. These sampling frames permitted the application of stratified two-stage probability proportional to size sampling. That is, schools were selected within strata with probability proportional to the number of students in each school who were members of the defined target population, and then intact classes were selected with probability proportional to the number of students within each class who were members of the defined target population.

The application of this kind of multi-stage sampling resulted in unequal probabilities of selection and therefore sampling weights were added to the student data files prior to the estimation of population characteristics. The calculation of sampling weights followed the classical (Horvitz-Thompson) procedure of assigning each student a weight that was proportional to the reciprocal of the probability of including a student in the sample.

The reciprocals of these probabilities are sometimes referred to as "raising factors" because they refer to the number of elements in the population that are "represented" by the various sample elements. These raising factors are often multiplied by a constant so that the "weighted sample size" is equal to the actual sample size. In this case the constant was n/N and the sampling weights were as follows.

$$wt = (N_h \times n) / (a_h \times b_{hi} \times N_{hij} \times N)$$

Where
- N_h = stratum size
- n = sample size
- a_h = no. schools selected in stratum h
- b_{hi} = no. classes selected in school i
- N_{hij} = no. students in selected class
- N = population size.

The sampling weights described above were required for undertaking analyses at the between-student level. For analyses conducted at other levels of data aggregation it was necessary to calculate appropriate weights. For this study, it was decided that the weight for a school would be proportional to the sum of its student's weights, and the weight for a teacher would be proportional to the sum of his/her student's weights.

One of the consequences of this approach to weighting was that the weighted mean score for a school variable referred to a school characteristic experienced by "the average student" — but not necessarily to a characteristic of "the average school". Similarly, a weighted mean score for a teacher variable referred to a teacher characteristic experienced by "the average student" — but not necessarily a characteristic of "the average teacher".

The student sampling weights for the countries that employed double-level sampling frames were calculated on a country by country basis using the same general principle of making the weights proportional to the inverse of the probability of selection.

SAMPLING ERRORS

The allocation of the sample across the strata of the defined target population required the NRC's to take decisions concerning the numbers of schools and students that were to be selected. The IEA standard for sampling accuracy required that all samples should aim to have an effective sample size for the main criterion variables of at least 400 students. That is, all samples were designed in an attempt to obtain sampling errors that were equal to, or smaller than, a simple random sample of 400 students.

An effective sample size of 400 students provided the following approximate 95 percent confidence limits for sample estimates of population values for means, percentages, and correlation coefficients (Ross, 1985).

1. Means: $m \pm 0.1s$ (where m is a sample mean, and s is the sample standard deviation for students).

2. Percentages: $p \pm 5.0\%$ (where p is a sample percentage).

3. Correlations: $r \pm 0.1$ (where r is a sample correlation).

The sample designs employed in this study departed markedly from the usual "textbook model" of simple random sampling. This departure demanded that special steps be taken in order to calculate measures of the stability of parameter estimates derived from the data. The Steering Committee decided that the most appropriate approach would be to use the "Jackknife" procedure (Tukey, 1958) to make the necessary calculations. The Jackknife procedure was applied to the data from each country after all data had been cleaned and merged at the International Co-ordinating Centre.

These calculations also permitted the calculation of the design effects (the ratio of the variance of the mean calculated from the complex sample designs used in this study to the variance of the mean that would be estimated if simple random sampling had been used).

The Jackknife Procedure

The development of the Jackknife procedure may be traced back to a method used by Quenouille (1956) to reduce the bias of estimates. Further refinement of the method (Mosteller and Tukey, 1968) led to its application in a range of social science situations where formulae are not readily available for the calculation of sampling errors.

The Jackknife procedure requires that an initial estimate, y_{all}, of a statistic, y, be made on the total sample of data. The total sample is then divided into k subgroups and y_i is computed, using the same functional form as y_{all}, but based on the reduced sample of data obtained by omitting the ith subgroup. Then k "pseudovalues" y_i^* (i=1,2,...,k) can be defined - based on the k reduced samples:

$$y_i^* = k y_{all} - (k - 1) y_i$$

Quenouille's estimator (also called the "Jackknife value") is the mean of the k pseudovalues:

$$y^* = 1/k \sum_{i=1}^{k} y_i^*$$

Quenouille's contribution was to show that, while y_{all} may have bias of order 1/n as an estimate of y, the Jackknife value, y^*, has bias of order $1/n^2$.

The variance of y^* may be estimated from the pseudovalues.

$$var(y^*) = var(y_i^*)/k$$

$$= 1/k\,(k-1) \sum_{i=1}^{k} (y_i^* - y^*)^2$$

Tukey (1958) set forward the proposal that the pseudovalues could be treated as if they were approximately independent observations and that Student's t distribution could be applied to these estimates in order to construct confidence intervals for y^*. Later empirical work conducted by Frankel (1971) provided support for these proposals when the Jackknife technique was applied to complex sample designs and a variety of simple and complex statistics.

Substituting for y_i^* in the expression for $var(y^*)$ permits the variance of y^* to be estimated from the k subsample estimates, y_i, and their mean - without the need to calculate pseudovalues.

$$var(y^*) = 1/k\,(k-1) \sum_{i=1}^{k} (y_i - \bar{y}_i)^2$$

Wolter (1985, p. 156) has shown that replacing y^* by y_{all} in the right hand side of the first expression for $var(y^*)$ given above provides a conservative estimate of $var(y^*)$ - the overestimate being equal to $(y_{all} - y^*)^2/(k-1)$.

Wolter (1985) and Rust (1985) have presented an extension of these formulae for complex stratified sample designs in which there are k_h primary sampling units in the hth stratum (where h = 1,2,...,H). In this case, the formula for the variance of y_{all} employs y_{hi} to denote the estimator derived from the same functional form as y_{all} - calculated after deleting the ith primary sampling unit from the hth stratum.

$$var(y_{all}) = \sum_{h=1}^{H} (k_h - 1)/k_h \sum_{i=1}^{k_h} (y_{hi} - y_{all})^2$$

Where $K = \sum k_h$ is the total number of samples that are formed.

The majority of sample designs used in this study employed schools as the primary sampling units. For these sample designs, the K estimates of y_{hi} were obtained by using data based on the removal of one school at a time from the total sample. Where countries used geographical areas as the primary sampling unit, followed by the selection of more than one school per area, the estimates of y_{hi} were based on reduced samples formed by omitting one geographical area at a time from the total sample.

In Tables 8 and 9 the total sample size and mean score on the reading test for Populations A and B have been presented for each country in association with the standard deviation of the reading scores, the sampling error of the means, the design effect, and the effective sample size.

The sample designs for Population B in Nigeria, Thailand, and Zimbabwe did not meet the sampling standards required for the study. Therefore it was not possible to calculate the sampling errors for these three samples. This is indicated in Table 9 by a double asterisk (**).

In Table 8 two countries, Iceland and Singapore, have been allocated a single asterisk (*) for their sampling error, design effects, and effective sample sizes. At Population A level Iceland tested the whole population of students and Singapore selected the whole population of schools. This resulted in sampling errors of either zero (Iceland) or near to zero (Singapore). That is, the sampling errors for Iceland and Singapore at the population A level were far superior to the standard that had been set down for the study. For the same reasons, it can be seen for Population B in Table 9 that Cyprus, Iceland, Singapore, and Trinidad-Tobago had sampling errors that were zero or close to zero.

The effective sample sizes for each country have been presented in the final column of Tables 8 and 9. At the population A level it can be seen all countries except France exceeded the standard of an effective sample size of 400 students. The effective sample size for France was slightly below this level – but not so much as to introduce a major instability into the estimate of the mean score on the reading test.

At the Population B level several countries did not meet the sampling standard of an effective sample size of 400 students. The main reason for this was that a key factor (the coefficient of intraclass correlation) in the calculation of required sample size prior to the data collection was underestimated. For these countries an estimated value of 0.3 was used for this coefficient. However, after the data were collected it was shown that the values that should have been used were 0.35 for France, 0.48 for Ireland, 0.50 for the Netherlands, 0.41 for New Zealand, and 0.61 for the Philippines. Fortunately, for all of these countries, the effective sample size was between around 300 and 400 which ensured that 95 percent confidence limits for the mean total reading score were only slightly larger (at a maximum of ± 0.12 student standard deviation units) compared with the sampling standard for the study (at ± 0.10 student standard deviation units).

Table 8: The Sampling Errors, Design Effects, and Effective Sample Size for Total School on the Reading Test (Population A)

COUNTRY	Sample Size	Mean	Standard Deviation	Sampling Error	Design Effect	Effective Sample Size
BELGIUM	2698	508	86	3.35	4.04	668
BOSTWANA	2662	499	91	3.39	3.81	699
CYPRUS	1494	480	87	3.28	1.00	1494
DENMARK	3375	472	124	3.93	3.56	948
FINLAND	1552	577	80	3.35	2.70	575
FRANCE	1877	534	84	4.26	4.52	389
GERMANY(E)	2871	503	93	3.21	4.98	577
GERMANY(W	1863	498	93	4.67	3.68	506
GREECE	3499	506	85	3.79	6.83	512
HONG KONG	3313	516	79	4.14	7.30	454
HUNGARY	3010	500	87	3.72	5.53	544
ICELAND	3961	521	95	*	*	*
INDONESIA	3169	381	66	3.22	5.90	537
IRELAND	2710	510	88	3.99	5.56	487
ITALY	2218	534	91	4.51	5.47	405
NETHER-LANDS	1700	483	81	3.86	3.88	438
NEW ZEALAND	3015	532	97	3.53	3.98	758
NORWAY	2449	528	104	2.89	1.76	1391
PORTUGAL	2793	475	81	4.10	6.52	428
SINGAPORE	7331	517	82	*	*	*
SLOVENIA	3297	498	87	2.74	3.28	1005
SPAIN	8230	504	87	2.92	7.47	1102
SWEDEN	4301	545	108	3.34	4.20	1024
SWITZER-LAND	3430	512	92	2.99	3.33	1030
TRINIDAD	3684	445	89	3.84	6.85	538
UNITED STATES	6571	553	85	2.98	7.22	910
VENEZUELA	4704	369	84	3.95	8.86	531

Table 9: The Sampling Errors, Design Effects, and Effective Sample Size for Total School on the Reading Test (Population B)

COUNTRY	Sample Size	Mean	Standard Deviation	Sampling Error	Design Effect	Effective Sample Size
BELGIUM	2733	480	86	4.26	6.76	404
BOSTWANA	4763	313	48	1.79	6.61	721
CANADA	4868	522	94	3.13	5.39	903
CYPRUS	1412	497	80	*	*	*
DENMARK	3854	527	85	2.49	3.34	1153
FINLAND	1377	565	72	2.31	1.39	991
FRANCE	2618	544	76	4.43	8.82	297
GERMANY(E)	4210	525	85	2.52	3.49	1206
GERMANY(W	1852	527	82	3.47	3.91	474
GREECE	3922	511	71	2.21	3.80	1032
HONG KONG	3160	538	69	3.90	7.71	410
HUNGARY	3374	540	81	3.44	5.04	669
ICELAND	3846	539	86	*	*	*
IRELAND	3639	513	90	5.11	10.69	340
ITALY	3064	517	80	3.26	5.05	607
NETHER-LANDS	3737	515	83	4.62	11.56	323
NEW ZEALAND	3123	551	103	5.96	9.29	336
NIGERIA	2374	389	78	* *	* *	* *
NORWAY	2291	518	80	2.56	1.71	1340
PHILLIPINES	9708	423	72	4.12	26.04	372
PORTUGAL	3420	526	64	3.55	7.87	435
SINGAPORE	4805	539	73	*	*	*
SLOVENIA	3230	536	69	2.29	3.59	900
SPAIN	8485	489	70	2.97	8.87	957
SWEDEN	3627	551	89	2.43	2.72	1333
SWITZER-LAND	6529	540	81	3.25	8.63	757
THAILAND	2749	475	90	* *	* *	* *
TRINIDAD	3027	478	97	*	*	*
UNITED STATES	3417	539	94	5.15	8.51	402
VENEZUELA	4316	409	66	3.05	7.77	555
ZIMBABWE	2768	359	67	* *	* *	* *

CONCLUSION

The IEA International Study of Reading Literacy was a large and complex study whose validity depended to a great extent on the quality of the sample designs applied in each participating country. An examination of Tables 8 and 9 shows that, for most countries, the National Research Coordinators prepared and executed sampling plans that went far beyond the accuracy standards set for the study. Only in three countries, and only at the population B level, were sample designs prepared that prevented the estimation of sampling errors.

The finding that several countries at the Population B level had effective sample sizes below 400 was due to the fact that the co-efficient of intraclass correlation was generally larger than expected. However, the loss in accuracy for these countries was not at a level that threatened the generalizability of the sample estimates.

The generally higher values of the coefficient of intraclass correlation that were obtained for the Population B samples provide a strong message to the designers of future IEA studies to be cautious about making prior estimates of sample size. The results from this study suggest that when using reading achievement as a criterion measure a general prior estimate of 0.3 for the value of the coefficient of intraclass correlation when intact classes are sampled in primary schools (Population A) will give quite good estimates of the required sample size. However, for the secondary school level (Population B) a value in the range of 0.4 to 0.5 may be required for many countries.

REFERENCES

Frankel, M.R. (1971). Inference from survey samples. Ann Arbor, Michigan: Institute for Social Research.

Kish, L. (1965). Survey sampling. New York: Wiley.

Mosteller, F. and Tukey, J.W. (1968). Data analysis and regression: A second course in statistics. Reading, Mass.: Addison-Wesley.

Quenouille, M.J. (1956). Notes on bias in estimation. Biometrika, 43, 353-360.

Ross, K.N. (1985). Survey sampling. In: T. Husen and T.N. Postlethwaite (Eds.), The International Encyclopedia of Education (pp. 4370-4381). New York: Pergamon.

Rust, K. (1985). Variance estimation for complex estimators in sample surveys. Journal of Official Statistics, 1 (4), 381-397.

Tukey, J.W. (1958). Bias and confidence in not-quiet large samples (abstract). Annals of Mathematical Statistics, 29, 614.

Wolter, K.M. (1985). Introduction to variance estimation. New York: Springer-Verlag.

Chapter 5

Questionnaire Development

Hans Wagemaker

QUESTIONNAIRE DEVELOPMENT

Elley (1994:3)[1] notes that one of the key aims of the IEA Reading Literacy study was to produce 'valid international tests and questionnaires which could be used to investigate reading literacy development in other countries.' This chapter describes briefly the procedures and activities which were aimed at producing instruments which were reliable across many countries and which captured the information necessary to provide information about those explanatory variables that were believed to be implicated in the process of reading acquisition.

The Conceptual Framework

In some ways, unlike the process of developing the reading tests and deciding on their content, the kinds of questions which can be asked of students, teachers and school principals about the factors which might be implicated in the process of learning to read are almost limitless. These questions are shaped by the different theories of learning participants might hold together with the different country circumstances which need to be reflected and captured in any international study. In order to provide a structure or frame of reference for the preparation of background questionnaires a conceptual model was developed in the initial stages of proposal development. Figure 1 below provides an outline of the conceptual or theoretical model which guided the development of the questionnaires. While the model was modified slightly in the course of the study's progress the four sets of constructs delimited the nature of the questions that were to be included in the questionnaires and the analysis that was to follow. It should be noted that the model was designed as a basis for the proposed multi-variate analyses the were planed for the Population A data. While no such analyses were planned for Population B, the model served as a basic framework for questionnaire construction for Population B.

[1] Some of the material contained inthis chapter has previously been prepared for; Introdution Elley, W.B., Schleicher, A. and Wagemaker H., in Elley, W.B., (Ed.) *The IEA Study of Reading*

Figure 1: Overall theoretical model used for IEA reading literacy study, Population A

*Note: Included in the model by some countires (viz.New Zealand)

The constructs labelled Background Variables included those home variables for which information can be collected from students aged 9 to 14 and which were believed to be implicated in achievement in reading literacy. Socio-economic status for example, as measured by such variables as parental education and resources in the home, is implicated in explaining achievement in most countries. Children whose home language differs from the language of instruction, too, may find themselves to be disadvantaged in the school setting particularly when it comes to reading. Furthermore, those activities such as television viewing which compete with reading for a student's discretionary time may also be related to performance in reading. Questions and hypothesis such as these influenced the construction of this section of the questionnaire.

As with the factors that can be found in the home which are related to school outcomes, there are a number of factors within the school that may be hypothesised to impact on achievement. School resources such as the availability of books in the classroom, the availability of classroom and school libraries, class size may all be presumed to have the potential to impact on reading outcomes. Furthermore, on the basis of research in the area of school effectiveness, the role of the principal and the relationship between school and community may also be considered to be important in accounting for differences in reading outcomes.

A key element in any model of classroom learning are the teachers, the behaviours they exhibit and the skills and experience that they posses. These assumptions were to be tested by the development and inclusion of a set of questions which focussed on the levels of education and training teachers had received, their teaching experience, their classroom practices and pedagogical beliefs with respect to the reading process. School polices and classroom practices such as those related to library usage and borrowing, homework and assessment which interact with teaching behaviour and experience were also included and were to be examined when controlling for other factors such as home background.

The right hand side of the conceptual model includes the major dependent variables in the study. In addition to the three test domains, Narrative, Expository and Document achievement scores, Voluntary Reading Activities is identified. The latter generically describes those student based reading behaviours (type and frequency of reading different types of materials) that students claim characterises their reading habits either in school or at home . The information for this variable was to be collected by way of the student questionnaire and it was proposed to study the impact of home and school factor in relation to these activities.

Finally, although not described in the conceptual model, the questionnaire development process included the preparation of a National Case Study questionnaire which was intended to capture information on the wider social, economic and educational policy context in each country. The degree of system centralisation, the relative wealth of a nation, the apparent emphasis on a literate culture in a nation (as measured by newspaper readership for example) are all factors which might be taken into account when attempting to explain differences in reading achievement among

countries. The information from the latter instrument was used to construct a Composite Development Index[2] which was found to correlate highly with mean reading achievement.

Questionnaire Preparation

The development of the questionnaires occurred largely in the working meetings of a sub- group of NRC's at the time of the scheduled NRC meetings. These meeting were augmented from time to time by smaller meetings of a sub-group of NRC's who met with members of the steering committee to review both questions and aspects of the model. In addition, the sometimes intense debate about how best to capture information internationally for some of the key variables and constructs and their development was influenced by the debates in the plenary sessions which were held to refine the aims of the study and the model of reading achievement that was to be explored.

As implied by the conceptual framework, parallel development of three key questionnaires for Students, Teachers and the School principal was progressed at each of the NRC meetings. All NRC's were invited to submit to the subcommittee models of those questions which were imbedded in the conceptual framework in a format which would be suitable for all countries. The procedure adopted included identifying an indicator which could be operationalised, (either as a single variable or as a construct which required the input of several variables), the identification of the most appropriate source of information for this indicator (student, teacher or school principal) and finally how to accommodate national differences to ensure data were captured in a comparable way internationally.

Trial and Review

Student, Teacher and School questionnaires were trailed in 28 countries for Population A and 25 Countries for Population B. Feedback on the National Case Study Questionnaire was received from 25 countries.

For review purposes by country tables were prepared for each item/variable which categorised each item according to the four types of scales used (nominal, ordinal, interval, ratio) , the coding scheme for the variable, and whether or not the item had been recoded. Information as to population, filetype and instrument in which the item was to be found was also provided.

Statistics for each item included the number of valid observations (N), means and standard deviations as well as percentages for options where appropriate. The task for reviewers (a sub-group of NRC's) and NRC's was to identify those items in respective countries with implausible ranges and implausible means and to examine the results with a view to determining where compliance effects may have occurred. A small committee reviewed variable by variable all variables in nearly all countries. A list which covered all variables in the Student, teacher and School questionnaire was drawn up

[2]This was a numerical index based on six indicators combined with equal weight and was designed to show the extent of advantage enjoyed by studeNts brought up in more literate and wealthier countries. A full description of the CDI can be found in Elley (1994).

and was used to identify those variables in each country which appeared to be problematic. For each country variables were flagged where:

Missing data ranged between 10-19%
Missing data was greater than 20%
Data contained unlikely means and ranges
Instructions or variables had apparently been mistranslated
Variables had been left out
The data reflected an apparent compliance effect.

National Case Study Questionnaires were reviewed individually with NRC at the meeting in Denmark to discuss anomalies which had arisen in the collection of that data.

In addition to simple frequencies principal component analyses were conducted for the Quality of School Life variables (student questionnaire B only.) and for the preliminary work on the constructs.

The term construct here refers to a measure or summary to account for the regularity that we can observe among a number of conceptually related variables from the questionnaires. It represents the indirect observation of that conceptual relationship.

For the Quality of School Life variable a principal axis factoring with iterated communality estimation was performed as recommended by Williams and Batten(1981). The number of factors was determined using the mineigen criterion. Only factors with an eigen value >=1 were retained. For the remaining constructs interest lay only in how well the variables could be explained by a single construct so the number of factors to be retained was always set to one.

Tables prepared for the review a of quality of school life showed the factor loadings and eigenvalues for each country where as for the other constructs, tables showed eigenvalues and factor patterns with a second table showing final communality estimates.

This process of condensation reduced the some 600 variable to a set of some 41 constructs (or single variables) which were subject to further review by small technical committee at a meeting in Hamburg and included contributions from Prof Warwick Elley and Dr Trevor Williams of Westat.

Following the main data collection the final specifications for the constructs were generated under the leadership of Ingvar Lundberg working closely with Andreas Schleicher, Nadir Atash and Dirk Hastedt.

For the final analyses this group developed a general model with 12 (Figure 2) constructs onto which the 41 or more constructs could be mapped. The two main paths followed by the group concerned the reduction of the general framework to a manageable and useful size and the other concerned the validation in a technical sense of the constructs. In producing the constructs described in Figure 2 and the relationships between them in Figure 3, the group was conscious to temper the statistical and technical specifications by the theoretical issues and concerns related to the development of a model of reading.

The focus was to be on extracting the essence rather than trying to capture everything in the questionnaires. The group work followed two main

paths. One concerned the reduction of the general framework to a manageable and useful size. The other concerned more technical issues in the factor analytic efforts to specify and validate some of the constructs in A.

1. Higher-order constructs

The group developed a general model with only 12 constructs onto which the 41 (or more) constructs could be mapped. This general model is intended to be used as an organising principle of higher order. More specific and more substantive issues can then be mapped onto this model, eg., the influence of teaching strategies on achievement. The high-order constructs (except for the outcomes) are listed in Fig. 2 together with a specification of their lower order basis. In Fig. 3 the constructs have been organised into a structural model.

1.	Community resources	6, 6 c
2.	Home resources	1, 1 a, 2, 4
3.	Student characteristics	5, 6 b
4.	Teacher characteristics	7, 8, 9, 10, 11, 15, 7 a
5.	Principal characteristics	22, 23, 37
6.	Classroom characteristics	16, 19, 24, 32
7.	School characteristics	17, 18, 20, 21
8.	Student activities	3, 6 a, 4 a, 30, 34
9.	Teaching time	12, 13, 14
10.	Instructional approaches	25, 26, 27, 28, 29, 31, 33

Fig 2. Higher-order constructs

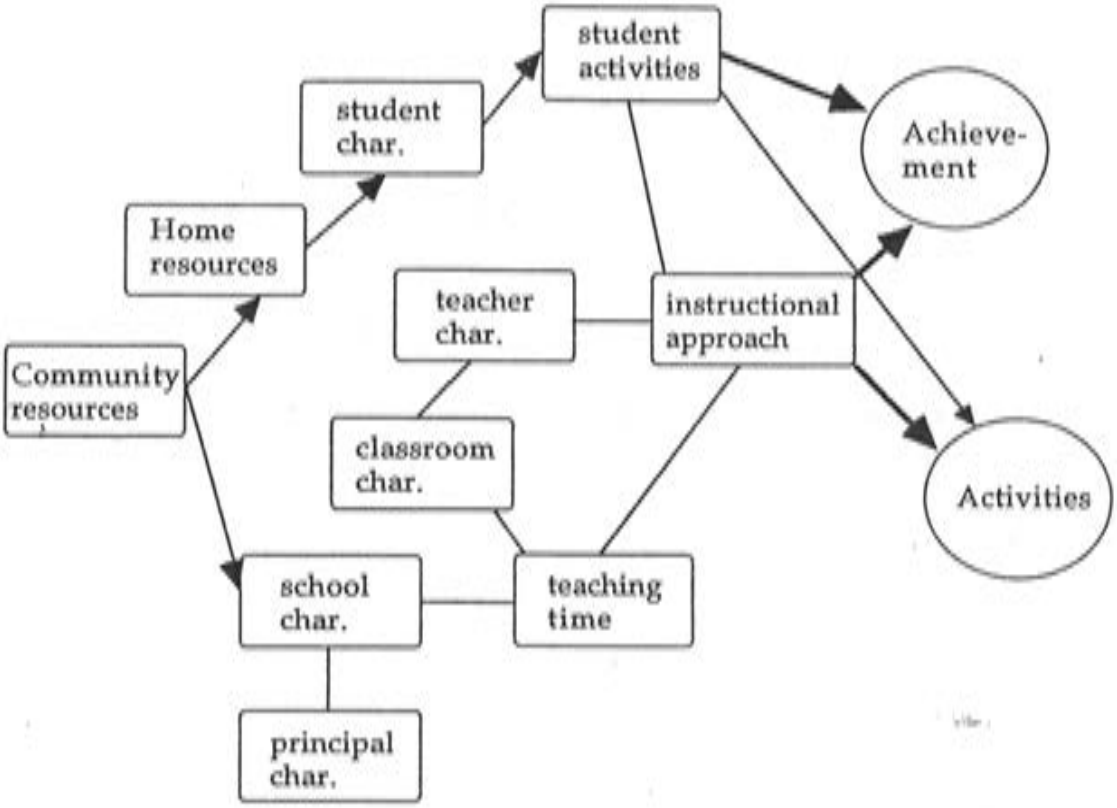


Fig 3. Structural relationships between higher-order constructs

Student Questionnaires

Separate questionnaires were developed for each population to be tested, each consisting of two parts. The first section o f each questionnaire focused on the students home and school conditions where as the second section contained questions which attempted to capture information related to students voluntary reading activities. The challenge in each case was to develop questions which intended to asses sometimes complex relationships, but which were simple and direct so as to accommodate the relative young age of respondents particularly for the students at the 9 year old level. Many of the questions that were asked of students including such factors as age, sex and home language and in some countries ethnicity were considered to be comparatively straight forward.

However previous experience in IEA studies had identified problems is securing information in a comparable fashion cross nationally for some types of information including, for example, socioeconomic status of the home particularly with the younger aged students. Not only was there a problem with reliability of the information, but in some countries it was also considered to be culturally unacceptable or in fact was prohibited to ask for this type of information.

An attempt to capture information on the socioeconomic status of the home was made by constructing a scale of student's home possessions and this was intended to be augmented by information about the frequency of meals and parents educational level. Each NRC was asked to provide a list of home possessions (Telephone, radio, video television etc.) which was judged would be found in the more affluent households but not all households in the country. It was obvious given the range of countries participating in the study, that a set of items common to all countries which would distinguish between more and less affluent households could not be produced. As reported by Elley (1994), despite the differences in the lists used in the various countries, both family and personal possessions predicted mean national achievement levels with correlations of .48 and . 50 respectively and when combined the correlation rose to .57 Following the pilot study, the attempt to capture information about parental educational level was dropped for the 9 year old population because of the high level s of missing data and questions about that data internationally. The pilot test also resulted in the modification of the home possessions scale with a reduction in the number of items to ten and the conclusion of a further eight items which represented students personal possessions. Furthermore, what was readily apparent following the pilot testing experience, was the need to rewrite questions which required students to provide information in a conditional situation typified by the "If yes, then: ...' structure. The final form of the questionnaire for students was constructed in such a way as to avoid these conditional situations where possible. This structure however had implications for the data cleaning programs which consequently had to include logical error checks,

To ensure reliability in the capture of certain types of information such as age, guidelines which were provided in accompanying notes for NRC's and in the manual for test administrators which included suggestions that

students be given prompts to assist them in supplying the information requested.

The second part of the student questionnaire was intended to capture information about students voluntary reading activities. Previous studies by Thordike (1973) who found consistently positive but low correlations between reading attitudes and achievement and Purves (1973) who found high correlations between professed interest in reading and achievement, together with other studies which focussed on the purposes and student perceptions of reading (Neumann 1990) suggested that was a potentially important area of investigation in this study. The pilot study attempted to capture information on both the frequency and duration with which students read a wide variety of materials. Trialing revealed several difficulties with the initial format of the questions. While the value of having test administrators read the questions aloud together with providing ample time for providing an answer was established, the complexity of the task presented considerable difficulties for the younger respondents.

While the questions for the main survey were radically shortened and in the case of the younger students the complexity of the task being reduced to a more simple yes/no response to a question about whether or not they had recently read different types of material, the difficulties experienced in the trial persisted. Implausible values suggested that students continued to find it difficult to estimate reliably how frequently they read different types of materials and in some countries there was evidence of an apparent compliance effect. In a separate study of this aspect of the questionnaires Otter (1993) reported that while the estimates for reading time reported for the trail version of the instruments while stable over two separate administrations, appeared to overestimate reading activity and deviated significantly from similar research studies in the Netherlands.

In the case of the fourteen year old population an International option to include a series of questions related to the quality of school life was also provided. These questions were based on the work of Williams and Batten (1981) which attempted to identify a series of factors related to school climate which might be implicated in understanding achievement outcomes. Trialing of this material revealed that the factor structure found in Australia was replicated in most countries that decided to include these questions as part of the student questionnaire. Where there were differences these were considered to be minor with the last factor in one or two cases splitting into two.

Teacher Questionnaire

The intent of the study design was to be able to relate teacher variables to student outcomes both within and between countries. Some forty six questions were asked of teachers which related to their training, experience, instructional practices, beliefs related to the teaching of reading, classrooms and identity.

The design of the separate questionnaires for the two populations was guided by the belief that in most countries formal classroom based reading instruction ceased by age 14. This immediately posed some difficulties when the circumstances of some of the less developed countries was considered. It was argued in those counties, the practices that characterised reading

practices in the junior classes in developed countries would also be found in the classes for the older age pupils. An initial attempt to capture more detailed information about teaching practices and beliefs from teachers at the secondary level was made by designing a two part questionnaire which afforded the opportunity for those teachers whose practices could be characterised as those which would normally be found in the junior school, to self select into that part of the questionnaire. This part of the questionnaire was targeted particularly at teachers in the less developed countries. In the event, the pilot study showed that the this was not an effective strategy. The questionnaire was redesigned to allow all teachers to complete all parts of the questionnaire..

Demographic Characteristics.

Teachers were asked to provide information on their gender, home language and years of teaching experience. Information was also requested about years of schooling , pre-service training and post-secondary education. Finding a common metric and language which would at the same time capture the considerable between country differences proved to be a considerable challenge. The as with several other questions was resolved by permitting countries to ask the questions in a culturally appropriate way with the proviso that the intention of the question was not lost and that the information could be re-coded into an international format.

While some questions, including those related to teachers readership, classroom conditions posed few design difficulties, questions related to the total time teachers spent on teaching reading were less successful as teachers had difficulty separating their reading instruction from their general language instruction.

School (Principal) questionnaire.

The context in which reading instruction occurs was the focus of the questionnaire that was designed to be completed by the school principal. Twenty four questions which focussed on school and community resources included details of school staffing, resources available in the community, instructional days, extent of absenteeism, size of the school library and the presence or absence of classroom libraries, special reading programs and the way in which the principal perceived his or her role. Again is the absence of a meaningful common definition for such things as whether or not the school was located in a rural or urban community, meant that information had to be captured in a way which was appropriate locally and then recoded into the international categories.

National Case Study Questionnaire.

In order to understand the way in which polices at the national level impact on reading instructional practices and the ways in which the broader social context might yield new insight into the way in which the teaching of reading was carried out across countries, a National Case Study Questionnaire (NCSQ) was designed.

The NCSQ attempted to gather information on educational , social, economic, political and cultural factors which might help explain variance in reading achievement across countries.

Seven clusters of questions which included ; the structure of the school system, curriculum in reading, instructional time, teachers, national language and culture, and general socioeconomic conditions were included in the questionnaire. In addition, the final section allowed national research co-ordinators to identify any factors which they considered would assist with interpreting the results for their respective countries. The data from the NCSQ was used to contribute to the development o f a composite development index (CDI)[3]

Following a trial of this instrument it became clear that where the information that was reported could not be sourced to published documents our which required a degree of estimation, there was a need for some associated estimate of reliability of that information. This information was supplied by NRC's on the final format of the questionnaire. The final form of the questionnaire was structured in the way described below.

Structure of the School System.

Countries my be described or characterised differently in terms of factors such as the provision and nature of pre-compulsory education, age of entry and length of compulsory schooling, and the provision of vocational and academic training and education. Questions may be asked as to how these structural characteristics impact on reading outcomes. However aside from considering whether or not differences in age of entry to compulsory schooling might have an impact on reading at ages 9 or 14, data asked in this section of the questionnaire was intended to provide information which would allow the results of the reading tests to be place in the broader educational policy context.

Reading Curriculum.

The variation that may be observed it the structures and administrative e policies that characterises the 32 countries that participated in the Reading Literacy study are also likely to be reflected in the curriculum as it relates to reading. Questions about the extent of centralisation of the curriculum and the extent to which teachers had discretion as to how the curriculum was implemented was considered to be an important set of variables when explanations for differences in were sought. Where the reading curriculum was prescribed , NRC's were asked to provide the International Co-ordianting centre with a copy.

Teachers.

In addition to those questions asked directly of teachers in the teacher questionnaire, NRC's were asked to provide d information on the levels of education for both primary and secondary teachers, , the percentage of teachers who were female. , current developments in teacher training, the status of teaching as a profession (comparative remuneration) and any information on current developments in the area of teacher training. Capturing this information was premised on the belief teacher quality as

measured by training and education was positively associated with learning outcomes.

Instructional Time

Instructional time as measured by instructional days, length of school day and proportion of time devoted to reading instruction represents the opportunity students have to learn which is directly related to learning outcomes. This information was collected at the policy level for grades K to 9 and was intended to complement the information collected from teachers and schools. Again this information was intended to provide a broader context into which the reading outcomes for the various countries could be placed. Variation both within and between countries could provide significant clues as to how to understand differences in performance.

Social and Economic Conditions.

The economic resources a country can bring to bear to support education and the broader social context in which education occurs were also considered to be potentially significant for examining and explaining reading outcomes. Of particular interest was educational expenditure (both private and public), live birth rate (an estimate of health status), percentage of women in full time/part time employment, unemployment rates, GDP, average disposable income, infant mortality, ratio of hospital beds per 1000, and the percentage of households with telephones. These variables in particular were used to construct an index of development.

National Language and Culture.

Countries which were culturally and linguistically heterogenous might be expected to have overall lower levels of reading performance where reading was assessed only in the majority language(s). Information was gathered on the major mother tongue languages spoken in the country, the nature of the official languages, the proportion of children whose home language differed from the language of the school, the presence or absence of any political debate on language policy and the extent to which minority cultures were represented in the content of the curriculum.
In addition information was gathered on the existence and nature of the infrastructures that existed which might be considered to provide support for or a culture of literacy. Information collected included daily newspaper circulation, the number of books published, library circulation, and book sales at the national level.

Finally NRC's in the final section of the questionnaire were provided with an opportunity to identify any particular features of their system or any circumstances peculiar to their country which, in their view, might explain or assist with the interpretation of reading performance in their country.

Data management for this questionnaire was managed at the New Zealand national centre by Dr. Hans Wagemaker. Information was finally received from 31 out of the 32 countries that participated in the study and despite considerable simplification, it was clear that for some of the less

developed countries in particular, requested information was difficult to obtain in a reliable way.

Summary

The construction of questionnaires which were capable of capturing information from students at two age levels, teachers, principals and policy makers across some 32 countries which represented some 23 different languages, and several different orthographies and many different research traditions, proved to be a considerable challenge. The experience has clearly identified some of the limits and possibilities of the use of questionnaires as a means of obtaining reliable estimates for some types of information from the younger aged students. However it has also shown that large scale surveys can be conducted in a way in which transcends the restrictions imposed by language, culture and tradition. Copies of the final questionnaires are presented in Appendices A and B,

References

Elley, W. B., (1993) *How in the world do Students Read*, The International Association for the Evaluation of Edcuational Acheivement, Hamburg

Elley, W.B., (1994) *The IEA Study of Reading Literacy: Acivement and Instruction in Thirty -Two School Systems*, Pergamon

Lundberg I., Linnakyla P., (1993), *Teaching Reading aaournd the World*, The International Association for the Evaluation of Educational Achievement, Hamburg.

Neumann, S.B., and Kostinen, P. (1992) Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27, 1, 94-106

Otter, M. E., (1993) The quality of the IEA-Reading Activity Inventory pilot version, *Onderwijs Research*, Jaargang 18-Nr. 5, 261-271

Purves, A.C., (1973) *Literature education in ten counties*, Stockholm: Almquist and Wiksell

Thorndike, R. L., (1973). *Reading comprehension education in fifteen countries. International studies in evaluation III*. Stockholm: Almquist and Wiksell.

Williams, T., & Batten, M., (1981). The Quality of School Life. *Research Monograph, No. 12*. Melbourne, Australia: Australian Council for Educational Research.

Chapter 6

The Study Archive of the IEA Reading Literacy Study

Andreas Schleicher and Jean Yip

1. INTRODUCTION

1.1 THE STUDY ARCHIVE OF THE IEA READING LITERACY STUDY

2. STUDY DOCUMENTATION

2.1 OVERVIEW

2.2 NATIONAL DOCUMENTATION (FOLDERS 1-105)
2.2.1 Conceptual framework and methodology
2.2.2 National assessment material
2.2.3 Sampling
2.2.4 National survey operations
2.2.5 National reports and publications
2.2.6 Correspondence
2.2.7 National data documentation and datafiles
2.2.8 List of national folders
2.3 INTERNATIONAL FOLDERS (FOLDERS 106-122)
2.3.1 International Assessment instruments
2.3.2 International survey operations
2.3.3 International codebook
2.3.4 Panel rating
2.3.5 International meetings
2.3.6 Conceptual framework and test specification
2.3.7 List of international folders

3. INTERNATIONAL ELECTRONIC DATABASE

3.1 INTRODUCTION
3.1.1 Structural and Naming Convention of Datafiles
3.1.2 International Codebook
3.1.3 File Naming Conventions
3.1.4 List of datafiles (Population A)
3.1.5 List of datafiles (Population B)

4. DATA VERIFICATION RULES AND DATA QUALITY STANDARDS

4.1 INTRODUCTION
4.1.1 Detecting data verification problems
4.1.2 Resolving data verification problems
4.1.3 Checks undertaken in the countries before coding and entry of data
4.1.4 Handling of missing data
4.1.5 Control variables and participation information
4.1.6 National Recodings
4.2 DATA VERIFICATION STEPS APPLIED AT THE ICC
4.2.1 Verification of file integrity
4.2.2 Verification of control variables
4.2.3 Verification of identification variables
4.2.4 Merging of student, teacher and school data
4.2.5 Value validation
4.2.6 Consistency checks
4.2.7 Specific data verification rules applied to resolve inconsistencies

1.1 THE STUDY ARCHIVE OF THE IEA READING LITERACY STUDY

This chapter describes the international study archive of the IEA Reading Literacy Study which was conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA) between 1989 and 1992.

The study archive comprises both an electronic database and extensive documentation describing the national and international assessment instruments, the international survey design and the national sampling plans, the international and national survey operations, and the data processing.

The study archive involves data from some 10,000 teachers and 210,000 students from 31 nations, located on 5 continents, and representing more than 20 different languages. Twenty-seven education systems participated in the final survey for Population A which focused on 9-year old students, and 32 education systems participated in Population B which focused on 14-year old students. An educational system represents either a national system of education or a territorial part of a national system (in which case this territorial part is always indicated in brackets following the name of the national system). The former territories of the Federal Republic of Germany and the former territories of the German Democratic Republic were treated as two separate education systems. In accordance with general IEA terminology, in the following participating education systems are often referred to as "countries".

Participating systems were: Belgium (French community), Canada (British Columbia), Cyprus, Denmark, Finland, France, Germany (former territories of the Federal Republic of Germany: *fIFRG*), Germany (former territories of the German Democratic Republic: *fIGDR*), Greece, Hong Kong, Hungary, Iceland, Indonesia, Ireland, Italy, Netherlands, New Zealand, Nigeria, Norway, Philippines, Portugal, Singapore, Slovenia, Spain, Sweden, Switzerland, Thailand, Trinidad and Tobago, United States, Venezuela, Zimbabwe.

This chapter has three parts:

- the first part describes the national and international components of the study documentation;
- the second part describes the electronic database and presents naming and structural conventions that will be of help for the user of the data; and
- the third part describes the data verification procedures that were applied and defines the data quality standard of the electronic database.

2. STUDY DOCUMENTATION

2.1 OVERVIEW

The study documentation of the IEA Reading Literacy Study consists of 122 color-coded folders which are located at the IEA Secretariat.

Folders 1-105 document the implementation of the IEA Reading Literacy Study in each of the participating education systems. They provide materials on the conceptual framework used by the participating education systems, a copy of the national assessment materials, the national sampling plans and documentation on the sample execution, national survey operation materials, and copies of the national reports and publications.

Folders 106-122 document the development of the international assessment instruments from the pilot stage to the main study and also contain the international survey operation manuals, the international codebook for the electronic database, and other important information on the conduct of the study.

The contents of these folders is described below.

2.2 NATIONAL DOCUMENTATION (FOLDERS 1-105)

The national documentation has been filed for each participating education system in a separate set of folders. Each of these sets contains the following sections:

2.2.1 CONCEPTUAL FRAMEWORK AND METHODOLOGY

The international assessment materials were developed collaboratively by the participating countries under the direction of the Steering Committee of the IEA Reading Literacy Study. The section "Conceptual Framework and Methodology" in the national documentation folders contains various documents on the national curricula, reading assessment activities, and relevant publications in the countries which countries contributed to the development of the international assessment materials and to the methodology of the study. However, note that the many test items that were contributed by the countries are filed in the international folders 107-110 of the study archive.

2.2.2 NATIONAL ASSESSMENT MATERIAL

This section of the National Documentation folders contains the complete set of translated and/or nationally adapted questionnaires and reading tests as used in the pilot and main study by the participating countries (see also section "International Assessment Instruments" in chapter "Study Documentation").

2.2.3 SAMPLING

This section of the National Documentation folders documents the national sampling plans and the national sample execution.

In most countries a two-stage stratified cluster sample design was used in which the first-stage sampling units were individual schools, but sometimes also geographical areas (in which case a third sampling stage was introduced). The schools in each national defined target population were selected from an explicitly stratified national sampling frame of schools, with a probability proportional to the estimated number of students in the school who were in the respective national defined target populations (except in Iceland, where the whole student target population was tested). In some cases, an estimate of the total enrollment of the school was substituted for the estimated number of students in the target grades. Intact classes were then randomly selected in each selected school. The application of this design resulted in unequal selection probabilities which were compensated for by sampling weights. The sum of weights was then adjusted so that it was equal to the total number of students in the countries. For students who were excluded by within-school exclusion procedures, the weight was reset to zero afterwards. For a detailed documentation of the sampling design refer to the corresponding chapter in this Technical Report prepared by Ross.

2.2.3.1 SAMPLE DESIGN TABLES

The national target populations and their spatial and hierarchical structure, as well as the design and allocation of the samples, are documented in a series of seven sample design tables which are enclosed in this section of the National Documentation folders.

Table	*Sample design tables*
Table 1	Description of the National Target Population
Table 2	Tabular Description of the National Desired, Defined, and Excluded Population
Table 3	Stratification
Table 4	Sample Allocation
Table 5	Number of Schools in the Designed and Achieved Samples
Table 6	Number of Teachers in the Designed and Achieved Samples
Table 7	Number of Students in the Designed and Achieved Samples

2.2.3.2 MARKER VARIABLES

In order to check the quality of the sample data, comparisons of known characteristics of the defined target populations (referred to as marker variables)

were made with the sample data. National Centers were therefore requested to provide a report on comparison of the frequency distribution of selected marker variables with the corresponding frequency distribution associated with the defined target population (see also section "International Survey Operations" in chapter "Study Documentation"). The report has been included in this section of the National Documentation folders.

It was proposed by the IEA Technical Committee (IEA, 1986) that National Centers should consider the following list of commonly employed marker variables during their selection of marker variables.

- Sex of student
- Age of students in grade level
- Father's occupation within an appropriate age band
- State or regional location of school
- Residence in metropolitan or non-metropolitan area
- Ethnic background of students
- Teacher characteristics (e.g. sex, age, and qualifications)
- School characteristics (e.g. size, type and location)

2.2.4 NATIONAL SURVEY OPERATIONS

This section of the National Documentation folders documents important aspects of the national survey operations. In particular, it contains the following materials:

2.2.4.1 FIELD OPERATION MANUALS

It contains the (translated and nationally adapted) Test Administrator Manuals and School Coordinator Manuals.

2.2.4.2 SURVEY TRACKING FORMS

It further contains the Sampling Master List, the School Tracking Forms, and the Student Tracking Forms.

2.2.4.2.1 *Sampling Master List*

The Sampling Master List describes the achieved sample of schools and students. For each explicit stratum, it lists the number of sampled schools and the corresponding number of schools in the nationally defined target population in that stratum. It further lists for each school the number of sampled students in the achieved sample, and the number of students in the national defined target population in the corresponding school. Depending on which type of sampling frame the country used (type A, B, C, or F), three different types of the Sampling Master List were used.

2.2.4.2.2 *School Tracking Forms*

The School Tracking Forms provide information on the selection of classes within schools.

2.2.4.2.3 *Student Name Forms*
The Student Name Forms were used to track students during the various stages of the study, ranging from the identification and assignments of test instruments to the students, through the recording of the participation status in the different testing sessions, up to the data entry and verification procedures.

2.2.4.3 DATA PROCESSING
The section contains the documentation on the data processing as undertaken at the ICC and a list of changes, deviations, and problems that were encountered together with a documentation of the solutions undertaken (see also section "National data documentation and datafiles" in chapter "Study Documentation").

2.2.4.3.1 *Progress Documentation*
The progress documentation documents the materials received from National Centers and the history of the processing of these materials.

2.2.4.3.2 *Change Report and Deviation Report*
All National Centers were meant to report all deviations in the coding from the standards specified in the International Codebook. These deviations were reported in the Change Report and the Deviation Report. The Change report documents the implementation of questions where countries were asked to define their own options or where countries were asked to make certain adaptations. The Deviation Report is the National Center's report of any modifications to the instrument that would affect the coding of the data or the interpretation of the codes.

2.2.4.3.3 *Valid Range Forms*
For open-ended questions in the background questionnaires, National Centers were asked to provide valid ranges and other criteria for the verification of such variables in the Valid Range Forms.

2.2.4.3.4 *Data verification Report*
Data verification and data processing was undertaken at the ICC. This section contains the following:

- listing of problems encountered
- correspondence between the ICC and the National Centers regarding problems encountered
- documentation of any manual data verification undertaken

2.2.5 NATIONAL REPORTS AND PUBLICATIONS

This section of the National Documentation folders contains national reports and publications that were prepared by the National Centers.

2.2.6 CORRESPONDENCE

This section of the National Documentation folders contains the correspondence between the ICC and the National Centers during the conduct of the study.

2.2.7 NATIONAL DATA DOCUMENTATION AND DATAFILES

This section of the National Documentation folders contains the data documentation and a copy of the cleaned and weighted data of the main study together with important documentation on the calculation of sampling weights and the summary statistics on the background questions and test items.

Since a copy of these sections was returned to the National Centers, some important materials from previous sections are repeated in this section of the archive in a summary form.

It should be noted that for many practical purposes, this section of the National Documentation folders will contain sufficient information for the use of the data.

2.2.7.1 DATA VERIFICATION RULES

All data returned to the ICC were checked by a sequence of special data verification procedures. This section of the National Documentation folders summarizes the data verification rules that were applied. For a detailed description of the data verification procedures, see part three of this chapter.

2.2.7.2 REPORT ON THE DATA VERIFICATION

This part provides a summary of the deviations, coding problems, and inconsistencies that were encountered during the data verification process. The type of problems which were identified during the application of the data verification steps are indicated. It also provides a listing of the valid ranges applied by the ICC, and the manual corrections and national recodings that were undertaken at the ICC.

2.2.7.3 WEIGHING REPORT

This part documents the calculation of sampling weights in the respective countries. It contains in particular: (a) school level sampling information for the countries, (b) a list of excluded students, (c) a list of merging problems that affected the calculation of the sampling weights, (d) a list of descriptor variables that were used for weighting purposes, (e) a frequency distribution of the sampling weights, and (f) the weighting formulas.

2.2.7.3.1 *School level sampling information*

This comprises the basic sampling information at the school level, such as the identification codes of schools, the number of students in the defined target population in the explicit stratum from which the school was selected, the number of students in the defined target population in the school, and the total enrollment rates in the school and the grade enrollment rates in the school as collected from the school principal for boys and girls separately.

2.2.7.3.2 *List of excluded students*

This comprises of a list of those students that had been marked for exclusion in the student name forms either by the school coordinator or by the test administrator.

2.2.7.3.3 *List of merging problems that affected the calculation of sampling weights*

This comprises of a list of problems found when student, teacher and school files were merged. Student, teacher and school files were merged after data verification by Stratum ID, School ID and Class ID. The list also indicates duplicate Student IDs or

Teacher IDs, students with no teacher matches, students with no school match, teachers with no student match, teachers with no school match, schools with no student match, and schools with no teacher match.

2.2.7.3.4 *List of descriptor variables for weighting purposes*
This list provides the following information *for each class* in the datafile:

- the Country ID (ID_CNTRY),
- the Stratum ID (ID_STR),
- the School ID (ID_SCH),
- the Class ID (ID_CLS),
- the Teacher ID (ID_TCH),
- the total class size of the class being tested (A/BTCLSIZ),
- the number of grade level students in the class as collected from the Teacher Questionnaire (A/BTGRSIZ),
- a check on whether the A/BTGRSIZ values were greater than A/BCENRGB+A/BCENRGG, the number of students in the sample which is the number of students in the merged datafile (NR_STD)),
- the number of students in the defined target population in the stratum as collected from the Sampling Master List (STRSAMP),
- the number of students in the defined target population in the school as collected from the Sampling Master List (SCHSAMP),
- the number of students in the defined target population (which is the sum of STRSAMP over all strata) (AS_N),
- the number of schools in the stratum as counted in the datafile (DS_SCH),
- the number of students in the class as counted in the datafile (TS_STD),
- the percentage of students in the school in the target population based on the number of students in the stratum (target population) (STD_PER),
- the number of students in the school as counted in the datafile (CS_STD),
- the number of schools in the file as counted in the datafile (AS_SCH),
- the number of teachers as counted in the datafile (AS_TCH),
- the number of strata as counted in the datafile (AS_STR),
- the number of classes as counted in the datafile (AS_CLS),
- the number of students as counted in the datafile (AS_STD),

- the number of classes in the stratum as counted in the datafile (DS_CLS),
- the number of students in the stratum as counted in the datafile (DS_STD),
- the counted number of classes in the school as counted in the datafile (CS_CLS),
- the number of students in school as counted in the datafile (CS_STD),
- the number of students in the school, including students with EXCLNRD=1,2,3,4 as counted in the datafile (CSXX_STD),
- the number of students in the class as counted in the datafile (TS_STD),
- the number of students in class in the datafile, including students with EXCLNRD=1,2,3,4 as counted in the datafile (TSXX_STD),
- the student base weight with students included for which the variable EXCLRND was coded to 2, 3, or 4 included (STDWGTX),
- the student base weight with students included for which the variable EXCLNRD was coded to 2, 3, or 4 (STDWGT),
- the rounded student weight (STDWGTR),
- the student weight control flag (1 indicates that the student weight was imputed because it could not be calculated for that student (WGTERR)),
- the teacher weight (TCHWGT) (Note that the sum of student weights including students with EXCLNRD=2,3,4 but excluding students with EXCLNRD=1 linked to a teacher/school was taken as the corresponding teacher/school weight),
- the teacher weight control flag (1 indicates that the teacher weight was imputed because it could not be calculated directly for that teacher (TWGTERR)),
- the school weight calculated as the sum of student weights within school adjusted to total number of schools in file (SCHWGT) (the school weight calculated as the inverse of the selection probability of the school adjusted to total number of schools in file (SCHWGT1)),
- the school weight control flag for SCHWGT (1 indicates that school weight could not be calculated directly and had to be imputed (CWGTERR)),
- the sum of school weights SCHWGT1 (AW_SCH),
- the school weight control flag for SCHWGT1 (1 indicates that school weight could not be calculated and was imputed (CWGTERR1)),
- the sum of student weights in the class (TW_STD),
- the sum of student weights in the stratum (DW_STD),
- the sum of student weights in the sample (AW_STD),

- the weighting indicator variable (CK_STD) (1 indicates that the weighted sample size equals (or nearly equals) the total sample size),
- the weighting indicator variable (CK_STDH) (1 indicates that the weighted number of students in each stratum is in proportion to the sampling plan),
- the weighting indicator variable (CK_STDST) (1 indicates that the weighted number of students in the class is the same as the weighted number of students per stratum divided by the number of schools in the stratum)

2.2.7.3.5 *Weighting formulas*

This provides a brief description of the formulas used for the calculation of student weights. The weighting formulas and the weighting procedures were specified by the Sampling Referee of the IEA Reading Literacy Study, Dr. Kenneth Ross. They are described in detail in a separate chapter of this Technical Report.

2.2.7.4 SUMMARY STATISTICS ON BACKGROUND QUESTIONNAIRES

To assist National Centers in the preparation of National Reports, a set of tables with summary statistics for the student, teacher, and school background questionnaires was prepared. These are enclosed in this part of the folders. For each countries, there are three tables for each population, one for student data, one for teacher data and one for school data. Each table contains the variables analyzed as rows and the statistics as columns. The variables are sorted alphabetically by their variable name. The header of the table indicates the population and filetype as well as the respective statistics. The labels used to indicate the statistics are explained in the following:

N OBS	represents the unweighted total number of observations for a variable in the datafile.
N VLD	represents the unweighted number of observations with "not administered", "invalid" and "logical not applicable" values excluded, but "missing" included. The calculation of percentages is based on this value.
N	represents the unweighted number of observations on which the calculation of statistics, except percentages, is based, i.e. the number of observations with "not administered", "invalid", "logical not applicable" and "missing" excluded.
SUMWGT	represents the sum of weights for observations with codes different from "not administered", "invalid", "logical not applicable" and "missing".
N MISS	represents the unweighted number of observations with "not administered", "invalid", "logical not applicable" or "missing" codes.
% MISS	represents the unweighted percentage of observations with "not administered", "invalid", "logical not applicable" or "missing" codes based on the total number of observations.
MEAN	represents the weighted mean.

STDDEV represents the weighted standard deviation.

MODE represents the weighted mode. When there are several values with the same number of observations then the smallest value is taken between these as the mode.

MIN, MAX represent the minimum and maximum values of a variable.

P10, Q1, MEDIAN, Q3, P90 represent the weighted 10th, 25th, 50th, 75th and 90th percentiles of the distributions.

BOXPLOT For some variables with open codes a box plot is produced. The bottom and top edges of the box are located at the sample 25th and 75th percentiles. The center horizontal line is drawn at the sample median and the central plus sign (+) is at the sample mean. It is possible for all of this statistics to fall on the same printer line. The central vertical lines, called "whiskers", extend from the box as far as the data extend, to a distance of at most 1.5 interquartile ranges. (An interquartile range is the distance between the 25th and the 75th sample percentiles). Any value more extreme than this is marked with a zero if it is within three interquartile ranges of the box, or with an asterisk (*) if it is still more extreme.

U OMIT represents the unweighted number of observations with the missing code "missing".

U1, U2, U3, U4, U5, U6, U7, U8 represent the unweighted number of observations for each of the response categories.

N/APP represents the weighted number of observations with the missing-code "logical not applicable". This code is assigned to dependent questions for which the filter question was answered negative.

INVLD represents the weighted number of observations with the missing-code "invalid".

N/ADM represents the weighted number of observations with the missing-code "not administered".

OMIT represents the weighted number of observations with the missing-code "missing".

1, 2, 3, 4, 5, 6, 7, 8 represent the weighted number of observations for each of the response categories.

% OMIT represents the weighted percentage of observations with the missing-code "missing", where the percentage is based on all observations with valid codes including the missing code "missing" (N VLD).

%1, %2, %3, %4, %5, %6, %7, %8 represent the weighted percentages of observations for each of the response categories, where the percentages are based on all observations with valid codes including the missing code "missing" (N VLD).

The first box of each row presents the variable name and the variable label. An "R" at the end of a variable indicates where a variable was recoded at the ICC. A "D" at the end of a variable indicates a derived variable.

If statistics were produced for separate subgroups of observations (e.g. language groups) then these subgroups are presented in the subsequent box divided in as many lines as there are subgroups.

2.2.7.4.1 *Classification of background variables:*
The background variables have been classified according to four types of scales, namely nominal, ordinal, interval, and ratio, for the implied measure. The scale type assigned to each background variable is listed in the International Codebook. The statistics presented in the tables depend on the scale type of the variables.

For *nominal scaled variables (N)* the tables contain: NVLD, N, SUMWGT, %MISS, MODE weighted and unweighted frequency and weighted percentages for each response category.

For *ordinal scaled variables (O,I)* the tables contain: NVLD, N, SUMWGT, %MISS, MEAN, MODE, MIN, 10th 25th 50th 75th 90th weighted percentiles, MAX, weighted and unweighted frequency, (weighted cumulative frequency where this still fits on the same page) and weighted percentages for each response category.

For *interval and rational scaled variables (R,V)* the tables contain: NVLD, N, SUMWGT, %MISS, MEAN, STDDEV, MIN, 10th 25th 50th 75th 90th weighted percentiles, MAX, and box and whisker plots.

When the scaletype was changed during recoding, the statistics presented correspond to the scaletype of the recoded variables.

2.2.7.4.2 *Handling of missing data*
The coding of missing data was undertaken in the countries. A distinction between missing values that were missing because a variable was not administered (code "not administered") and missing values due to non response or invalid responses ("missing") was made.

The ICC recoded variables that could not be answered because of their relationships to other variables to "logical not applicable", e.g. in the case of filter-dependent relations (details are provided in subsequent sections). This was done by specific data verification procedures which depended on the design of the corresponding questions.

The ICC recoded closed variables with codes that do not conform to the range validation criteria specified in the International Codebook to "invalid".

2.2.7.4.3 *Recoding*
Some of the variables were recoded for the analysis at the ICC. The purpose of recoding was to assign reasonable numerical values to the response categories of ordinal variables and in some cases to categorical variables. Recoded variables were indicated by the suffix "R" in the variable name.

Recoding had the function of:

- providing a more direct interpretation of the data,
- enabling the calculation of means and standard deviations,
- enabling the calculation of correlation coefficient between the variable and others variables.

Recoded variables were created after data verification. If a variable that was to be recoded was coded to missing, not administered, not applicable or invalid, then the recoded value was set to missing "." .

2.2.7.5 SUMMARY STATISTICS ON TEST ITEMS

This report presents summary statistics for the student, teacher, and school background questionnaires. These summary statistics are in the folders.

2.2.7.5.1 *Item statistics*

Table 1 contains the items analyzed as rows and the statistics as columns. The items are sorted alphabetically by their variable name. The variable name indicates both the domain (D:Document, E:Expository, N:Narrative, W:Word Recognition) and the passage in which the reading item was located. The header of the table indicates the population and filetype as well as the respective statistics. The labels used to indicate the statistics are explained in the following:

SUMWGT	presents the sum of weights.
P-VAL	[boys/girls] (<domain>, <response option>) represents the weighted proportion of students choosing the corresponding response option. For the subgroups, boys and girls, separate p-values are computed for the correct answer.
P-BIS	[boys/girls] (<domain>, <response option>) represents the point-biserial correlation of the corresponding response option with the domain score. For the subgroups, boys and girls, separate point-biserial correlations are computed for the correct answer.
FLAGS	P: P-value of one or more option is < 2 or > 90 D: A distractor was more attractive than the correct answer B: The point biserial correlation is > .25 for the correct answer or > .1 for one or more distractors S: The mean score of boys and girls differs by more than 10% O: more than 25% of the students omitted the item
OMIT	Percentage of students omitting the item.
NR	Percentage of students not reaching the item.
P-MAL	Percentage of boys with correct answers
P-FM	Percentage of girls with correct answers

2.2.7.5.2 *Test reliabilities*
Table 2 presents the within-domain reliabilities (KR20) of the tests. The sum of weights, the number of items, the variance of scores and the reliability coefficient KR20 are displayed.

2.2.7.5.3 *Score distributions*
Table 3 presents the distribution of raw scores as well as the minimum, maximum, means and standard deviations of scores for the subgroups boys and girls.

2.2.7.5.4 *School results*
Table 4 presents the minimum, maximum, mean and standard deviations of scores for each school.

2.2.7.6 DISKETTES
This section contains the verified raw data files and further documentation on eight $3^{1/2}$" 1.44 MB diskettes in MS-DOS format. The files contained on each of these diskettes are listed in the following. Some of the files have been compressed. To un-compress these files, use the program PKUNZIP which is provided on the diskette labeled "Program files". By invoking the PKUNZIP program without specifying any further parameters, instructions on the use of the PKUNZIP program can be obtained.

CODEBOOK, International Codebook, Population A and B

ASALL01.PRN:	International Codebook, Population A
BSALL01.PRN:	International Codebook, Population B

RAW, Raw Data Files, Population A

AS<Country ID>R.ZIP:	contains the following compressed file:
AS<Country ID>R.RAW:	National Raw Data Files, Population A

RAW, Raw Data Files, Population B

BS<Country ID>R.ZIP:	contains the following compressed file:
BS<Country ID>R.RAW:	National Raw Data Files, Population B

Program files

PKUNZIPF.EXE:	MS-DOS program to un-compress compressed files
MAKESSDA.PRG:	SAS program to create a SAS system file from merged raw datafile, Population A
MAKESSDB.PRG:	SAS program to create a SAS system file from merged raw datafile, Population B
MAKETCHA.PRG:	SAS program to extract the teacher data from the merged SAS system file, Population A
MAKETCHB.PRG:	SAS program to extract the teacher data from the merged SAS system file, Population B
MAKESCHA.PRG:	SAS program to extract the school data from the merged SAS system file, Population A
MAKESCHB.PRG:	SAS program to extract the school data from the merged SAS system file, Population B

RESULT, International Univariates, Pop A

AUNIVAR.ZIP:	contains the following 3 compressed files:
ACUNITAB.LST:	Summary statistics from School Questionnaire, Population A
ATUNITAB.LST:	Summary statistics from Teacher Questionnaire, Population A
ASUNITAB.LST:	Summary statistics from Student Questionnaire, Population A

RESULT, International Univariates, Pop B

BUNIVAR.ZIP:	contains the following 3 compressed files:
BCUNITAB.LST:	Summary statistics from School Questionnaire, Population B
BTUNITAB.LST:	Summary statistics from Teacher Questionnaire, Population B
BSUNITAB.LST:	Summary statistics from Student Questionnaire, Population B

RESULT, International Item Analysis, Population A/B

ACLSITM.ZIP:	contains the following compressed file:
ACLSITM.LST:	Results from the classical item analysis, Population A
BCLSITM.ZIP:	contains the following compressed file:
BCLSITM.LST:	Results from the classical item analysis, Population B
A<Country ID>DW.LOG:	Results from the national Rasch item analysis, Population A, with "not reached" scored as "wrong", Document domain
B<Country ID>DW.LOG:	Results from the national Bigscale Rasch item analysis, Population B, with "not reached" scored as "wrong", Document domain
A<Country ID>DO.LOG:	Results from the national Rasch item analysis, Population A, with "not reached" scored as "omit", Document domain
B<Country ID>DO.LOG:	Results from the national Rasch item analysis, Population B, with "not reached" scored as "omit", Document domain
A<Country ID>EW.LOG:	Results from the national item analysis, Population A, with "not reached" scored as "wrong", Expository domain
B<Country ID>EW.LOG:	Results from the national Rasch item analysis, Population B, with "not reached" scored as "wrong", Expository domain
A<Country ID>EO.LOG:	Results from the national Rasch item analysis, Population A, with "not reached" scored as "omit", Expository domain
B<Country ID>EO.LOG:	Results from the national Rasch item analysis, Population B, with "not reached" scored as "omit", Expository domain

A<Country ID>NW.LOG:	Results from the national Rasch item analysis, Population A, with "not reached" scored as "wrong", Narrative domain
B<Country ID>NW.LOG:	Results from the national Rasch item analysis, Population B, with "not reached" scored as "wrong", Narrative domain
A<Country ID>NO.LOG:	Results from the national Rasch item analysis, Population A, with "not reached" scored as "omit", Narrative domain
B<Country ID>NO.LOG:	Results from the national Rasch item analysis, Population B, with "not reached" scored as "omit", Narrative domain
AALL01DW.LOG:	Results from the international Rasch item analysis (pooled data), Population A, with "not reached" scored as "wrong", Document domain
BALL01DW.LOG:	Results from the international Rasch item analysis (pooled data), Population B, with "not reached" scored as "wrong", Document domain
AALL01DO.LOG:	Results from the international Rasch item analysis (pooled data), Population A, with "not reached" scored as "omit", Document domain
BALL01DO.LOG:	Results from the international Rasch item analysis (pooled data), Population B, with "not reached" scored as "omit", Document domain
AALL01EW.LOG:	Results from the international Rasch item analysis (pooled data), Population A, with "not reached" scored as "wrong", Expository domain
BALL01EW.LOG:	Results from the international Rasch item analysis (pooled data), Population B, with "not reached" scored as "wrong", Expository domain
AALL01EO.LOG:	Results from the international Rasch item analysis (pooled data), Population A, with "not reached" scored as "omit", Expository domain
BALL01EO.LOG:	Results from the international Rasch item analysis (pooled data), Population B, with "not reached" scored as "omit", Expository domain
AALL01NW.LOG:	Results from the international Rasch item analysis (pooled data), Population A, with "not reached" scored as "wrong", Narrative domain
BALL01NW.LOG:	Results from the international Rasch item analysis (pooled data), Population B, with "not reached" scored as "wrong", Narrative domain
AALL01NO.LOG:	Results from the international Rasch item analysis (pooled data), Population A, with "not reached" scored as "omit", Narrative domain

BALL01NO.LOG:	Results from the international Rasch item analysis (pooled data), Population B, with "not reached" scored as "omit", Narrative domain
ASCORETB.ZIP:	contains the following compressed file:
ASCORETB.LST:	Weighted international score distributions for Rasch and classical test scores/ability measures, Population A
BSCORETB.ZIP:	contains the following compressed file:
BSCORETB.LST:	Weighted international score distributions for Rasch and classical test scores/ability measures, Population B
ARELIA.LST:	Test reliabilities from classical item analysis, Population A
BRELIA.LST:	Test reliabilities from classical item analysis, Population B

NRC.DOC, Cleaning Documentation, Weighting Reports, Population A/B

AC<Country ID>.RE?:	Cleaning Report, School data, Population A
AT<Country ID>.RE?:	Cleaning Report, Teacher data, Population A
AS<Country ID>.RE?:	Cleaning Report, Student data, Population A
BC<Country ID>.RE?:	Cleaning Report, School data, Population B
BT<Country ID>.RE?:	Cleaning Report, Teacher data, Population B
BS<Country ID>.RE?:	Cleaning Report, Student data, Population B
AC<Country ID>.LOG:	SAS Inconsistency Report, School data, Population A
AT<Country ID>.LOG:	SAS Inconsistency Report, Teacher data, Population A
AS<Country ID>.LOG:	SAS Inconsistency Report, Student data, Population A
BC<Country ID>.LOG:	SAS Inconsistency Report, School data, Population B
BT<Country ID>.LOG:	SAS Inconsistency Report, Teacher data, Population B
BS<Country ID>.LOG:	SAS Inconsistency Report, Student data, Population B
A<Country ID>.WGT:	SAS Weighting Report, Population A
B<Country ID>.WGT:	SAS Weighting Report, Population B

2.2.8 LIST OF NATIONAL FOLDERS

The following provides a list of the national folders. Note that some of these consist of multiple volumes.

FOLDER	COUNTRIES	RESOURCE	DESCRIPTION
1	Belgium (French)	(Volume 1)	National Assessment Materials
2	Belgium (French)	(Volume 2)	National Assessment Materials Sampling Survey Operations
3	Belgium (French)	(Volume 3)	Survey Operations National Reports and Publications Correspondence

4	Belgium (French)	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
5	Botswana	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials Sampling Survey Operations
6	Botswana	(Volume 2)	Survey Operations National Reports and Publications Correspondence
7	Botswana	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
8	Canada (BC)	(Volume 1)	National Assessment Materials
9	Canada (BC)	(Volume 2)	National Assessment Materials Sampling Survey Operations National Reports and Publications Correspondence
10	Canada (BC)	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
11	Cyprus	(Volume 1)	National Assessment Materials Sampling Survey Operations National Reports and Publications Panel Ratings Correspondence
12	Cyprus	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
13	Denmark	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
14	Denmark	(Volume 2)	Sampling Survey Operations National Reports and Publications Correspondence
15	Denmark	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes

16	Finland	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
17	Finland	(Volume 2)	Sampling Survey Operations
18	Finland	(Volume 3)	Survey Operations National Reports and Publications Panel Ratings Correspondence
19	Finland	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
20	France	(Volume 1)	National Assessment Materials Sampling
21	France	(Volume 2)	Survey Operations
22	France	(Volume 3)	Survey Operations National Reports and Publications Correspondence
23	France	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
24	Germany (ftGDR)	(Volume 1)	National Assessment Materials
25	Germany (ftGDR)	(Volume 2)	Sampling Survey Operations National Reports and Publications Panel Ratings Correspondence
26	Germany (ftGDR)	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
27	Germany (ftFRG)	(Volume 1)	National Assessment Materials Sampling Survey Operations National Reports and Publications
28	Germany (ftFRG)	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
29	Greece	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
30	Greece	(Volume 2)	Sampling Survey Operations National Reports and Publications Panel Ratings Correspondence

31	Greece	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
32	Hong Kong	(Volume 1)	National Assessment Materials
33	Hong Kong	(Volume 2)	Sampling Survey Operations National Reports and Publications Correspondence
34	Hong Kong	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
35	Hungary	(Volume 1)	National Assessment Materials
36	Hungary	(Volume 2)	Sampling Survey Operations National Reports and Publications Correspondence
37	Hungary	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
38	Iceland	(Volume 1)	National Assessment Materials
39	Iceland	(Volume 2)	Sampling Survey Operations National Reports and Publications Correspondence
40	Iceland	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
41	Indonesia	(Volume 1)	National Assessment Materials Sampling Survey Operations National Reports and Publications Correspondence
42	Indonesia	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
43	Ireland	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
44	Ireland	(Volume 2)	National Assessment Materials Sampling Survey Operations
45	Ireland	(Volume 3)	Survey Operations National Reports and Publications Correspondence

46	Ireland	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
47	Italy	(Volume 1)	National Assessment Materials
48	Italy	(Volume 2)	Sampling Survey Operations National Reports and Publications Correspondence
49	Italy	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
50	Latvia	(Volume 1)	National Assessment Materials Sampling Survey Operations Correspondence
51	Netherlands	(Volume 1)	National Assessment Materials
52	Netherlands	(Volume 2)	National Assessment Materials Sampling Survey Operations National Reports and Publications Panel Ratings Correspondence
53	Netherlands	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
54	New Zealand	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
55	New Zealand	(Volume 2)	Sampling Survey Operations
56	New Zealand	(Volume 3)	Survey Operations National Reports and Publications Correspondence
57	New Zealand	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
58	Nigeria	(Volume 1)	National Assessment Materials Sampling Survey Operations Correspondence
59	Nigeria	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
60	Norway	(Volume 1)	National Assessment Materials

61	Norway	(Volume 2)	Sampling Survey Operations
62	Norway	(Volume 3)	Survey Operations National Reports and Publications Correspondence
63	Norway	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
64	Philippines	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials Sampling
65	Philippines	(Volume 2)	Survey Operations National Reports and Publications Correspondence
66	Philippines	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
67	Portugal	(Volume 1)	National Assessment Materials
68	Portugal	(Volume 2)	National Assessment Materials Sampling
69	Portugal	(Volume 3)	Survey Operations National Reports and Publications Correspondence
70	Portugal	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
71	Singapore	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
72	Singapore	(Volume 2)	Sampling Survey Operations National Reports and Publications Correspondence
73	Singapore	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
74	Slovenia	(Volume 1)	National Assessment Materials Sampling
75	Slovenia	(Volume 2)	Survey Operations National Reports and Publications Correspondence

76	Slovenia	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
77	Spain	(Volume 1)	Conceptual Framework and Methodology
78	Spain	(Volume 2)	National Assessment Materials
79	Spain	(Volume 3)	Sampling Survey Operations National Reports and Publications Correspondence
80	Spain	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
81	Sweden	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
82	Sweden	(Volume 2)	Sampling Survey Operations
83	Sweden	(Volume 3)	National Reports and Publications Panel Ratings Correspondence
84	Sweden	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
85	Switzerland	(Volume 1)	National Assessment Materials
86	Switzerland	(Volume 2)	National Assessment Materials
87	Switzerland	(Volume 3)	Sampling Survey Operations National Reports and Publications Correspondence
88	Switzerland	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
89	Thailand	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials Sampling Survey Operations National Reports and Publications Correspondence
90	Thailand	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes

91	Trinidad and Tobago	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials Sampling
92	Trinidad and Tobago	(Volume 2)	Survey Operations
93	Trinidad and Tobago	(Volume 3)	Survey Operations National Reports and Publications Panel Ratings Correspondence
94	Trinidad and Tobago	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
95	United States	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials
96	United States	(Volume 2)	National Assessment Materials Sampling Survey Operations National Reports and Publications
97	United States	(Volume 3)	Correspondence
98	United States	(Volume 4)	Correspondence
99	United States	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
100	Venezuela	(Volume 1)	National Assessment Materials
101	Venezuela	(Volume 2)	Sampling Survey Operations National Reports and Publications Panel Ratings Correspondence
102	Venezuela	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes
103	Zimbabwe	(Volume 1)	Conceptual Framework and Methodology National Assessment Materials Sampling
104	Zimbabwe	(Volume 2)	Survey Operations National Reports and Publications Correspondence
105	Zimbabwe	National Data Documentation and Datafiles	Data verification Rules Data verification Documentation Weighting Report Summary Statistics Diskettes

2.3 INTERNATIONAL FOLDERS (FOLDERS 106-122)

2.3.1 INTERNATIONAL ASSESSMENT INSTRUMENTS

Folders 106-115 document the design and preparation of the international assessment instruments.

Activities for the design and preparation of the reading tests were coordinated by Warwick Elley (New Zealand, chair of the Steering Committee) and Ingvar Lundberg (Sweden, member of the Steering Committee). During the design of the study and the development of the Reading Literacy tests, prodigious efforts were made by researchers and National Research Coordinators to ensure that national circumstances were taken into account in order to ensure maximum content validity in every country. The National Centers played a major role in providing items for the instruments, rating them for suitability, editing them, pilot testing them on judgment samples, and reviewing them before their final administration.

Activities for the design and preparation of the background questionnaires were coordinated by Hans Wagemaker (New Zealand, National Research Coordinator). Past experiences of the IEA studies had some influence on the choice of indicators to be measured and on the style of questions selected for inclusion in the questionnaires. And once again, the National Centers and their National Advisory Committees played a major role in contributing to the conceptual framework and the questionnaires, pilot testing the questionnaires on judgment samples, and reviewing them before their final administration.

2.3.1.1 PROPOSED ITEMS FROM COUNTRIES

During the first meeting, hosted by the National Center for Educational Statistics (NCES) in Washington, D.C. (USA) in November 1988, the domains of reading literacy to be assessed were established. Following this, the Steering Committee requested National Centers to submit test items for the international reading tests. Folders 107-110 contain the proposed passages and items submitted by the National Centers.

2.3.1.2 PILOT TEST VERSION

Appropriate passages and items were selected and adapted by the Steering Committee. Further editing and screening was undertaken with these passages, and finally four parallel forms with common anchor items were prepared which are documented in Folder 111 together with a pilot version of the background questionnaires. These four parallel forms were rotated within classes. These pilot forms contained about three times as many items as were required for the main study tests.

2.3.1.3 SUGGESTIONS FOR REVISION ON PILOT QUESTIONNAIRES

The pilot instruments were trial tested on judgment samples in both populations in all the countries participating in the pilot study. During the pilot study, National Centers sent in comments on the problems they had encountered in translations or on deviations they had made to the background questionnaires to adapt them to their national systems. All these comments and deviations were gathered and

compiled in a master list at the ICC. The national comments and the master list are contained in folders 112-113.

2.3.1.4 PROPOSED FINAL TEST VERSION AND FINAL VERSION

After the pilot study, a committee consisting of members of the Steering Committee and selected National Research Coordinators prepared a proposal as to which items should be dropped, modified, or kept unchanged. Similar work was undertaken with regard to the background questionnaires. Following that, a set of proposed reading tests and background questionnaires for the main study was prepared and submitted to National Centers. These proposed assessment instruments are contained in folder 114.

2.3.1.5 FINAL INSTRUMENTS

All instruments were finalized at the 1990 National Center meeting in Frascati.

The final reading tests for Population A were made up of 106 test items which were administered in two testing sessions (with a length of 35 and 40 minutes respectively). The tests contained 40 items from the Word Recognition test, four reading passages with 22 items from the narrative domain, five passages with 21 items from the expository domain, and six documents with 23 items from the documents domain. Four of the test items required a word or a phrase to be written by the students, two items asked the students to write an extended response in their own words, and the remaining items had multiple-choice response format, with four choices for each item.

The final reading tests for Population B were made up of 89 test items, which were administered in two testing sessions (with a length of 40 and 45 minutes respectively). The tests contained five passages with 29 items from the narrative domain, five passages with 26 items from the expository domain, and 9 documents with 34 items from the document domain. Twenty of the test items required the students to write a word or phrase, two items asked the students to write an extended response in their own words, and the remaining items had multiple-choice response format, with four choices for each item.

The final Student Questionnaires had of two parts: (a) questions on identification, socioeconomic status, home literacy resources, home literacy interaction and others, and (b) questions about the students' voluntary reading activities. For Population B, a third optional part to assess students' perception of their school climate was also added. If the National Centers had national concerns which were not accommodated in the international questionnaire, they were invited to supplement the questionnaires with *international options* and *national options*. International options were coordinated between interested countries and supported and processed by the ICC, whereas national options had to be managed and processed under the responsibility of the respective countries.

The final Teacher Questionnaires were designed to obtain information on the teacher's identity, training, experience, instructional practices, classroom context, and beliefs related to the teaching of breading. As the design of the survey was based on the sampling of intact classes, it was possible to relate teacher variables to student outcomes. Once again, countries were invited to add *international* and *national*

options. The Teacher Questionnaires were administered to the teachers teaching the selected classes in reading.

The final School Questionnaires were designed to obtain information on the school circumstances and policies as viewed by the principal. It is made up of twenty-four questions inquiring about the size and type of school the students in the sampled school attended, the resources available in the community and the school, the size of the school library, details about the staff, the number of days of instruction, extent of absenteeism, special programs for reading, and the way the principal perceived his or her role in the school. Again, countries were invited to add *international* and *national options*

The National Case Study Questionnaires was designed to gather educational, economic, social, political, and cultural data which could help to explain variables in reading achievement across nations or regions. Questions were framed around seven major areas of interest, which included: the structure of the school system, curriculum in reading, teachers, instructional time, socioeconomic conditions, national language and culture, and a final section which requested for any information, which in the opinion of the National Center, would assist in the interpretation of the results.

The final test instruments and background questionnaires are contained in folder 115.

2.3.2 INTERNATIONAL SURVEY OPERATIONS

To assist National Centers in the administration of the IEA Reading Literacy Study, standardized survey procedures were developed at the ICC and countries were provided with a series of survey operation manuals. These manuals are contained in folder 116.

2.3.2.1 SAMPLING MANUAL

The sampling manual was prepared to assist National Centers in the design and execution of the survey samples. It provided guidance for the:

- description of the national the desired, defined, and excluded populations;
- selection and implementation of stratification variables;
- use of sample design tables to decide on the required number of schools and students;
- construction of the sampling frame;
- selection of the sample of school and two parallel replacement samples;
- selection of classes within schools;
- report the execution of the sample design in terms of response rates and marker variables;

- calculation of sampling weights; and for the
- calculation of sampling errors.

2.3.2.2 MANUAL FOR NATIONAL RESEARCH COORDINATORS

The Manual for National Research Coordinators described all operations to be taken care of by the National Centers from the moment the final versions of all instruments arrived at the National center, to the moment the final data sets and accompanying documentation were returned to the ICC.

2.3.2.3 MANUAL FOR SCHOOL COORDINATORS

The manual for School Coordinators described the steps to be undertaken by the School Coordinator from the moment the instruments and the Student Name Forms arrived in the school, to the moment the completed instruments and Student Name Forms were returned to the National Center.

2.3.2.4 MANUAL FOR TEST ADMINISTRATORS

The Manual for Test Administrators described the procedures to be undertaken by the Test Administrators from the moment the first testing session in the classroom began, to the moment the final testing session in the school was completed and all envelopes containing the assessment instruments were returned to the School Coordinator.

2.3.2.5 DATA PROCESSING MANUAL

The Data Processing Manual described the major steps to be undertaken by the National Center starting with the documentation of sampling frame information in computer readable form, through the entry of the data from the assessment materials, up to the data verification of the data in order to ensure that the datafiles which were returned to the ICC would be clean, mergeable, and in a consistent international format. It also described the international codebook, and guidelines and computer programs that were provided to the National Centers for the various stages of data processing.

2.3.2.6 DATA ENTRY MANUAL

The DataEntry User's Manual contained step by step instructions on how to operate the computer program that was provided to National Centers in order to enter data from the printed test instruments into the computer. It also described how to create new datafiles, how to modify the file structure of an existing datafile (for example, in order to add national variables to the codebook), how to change the range validation criteria or labels for some variables, and how to assign certain default or missing values to variables.

2.3.3 INTERNATIONAL CODEBOOK

Folders 117-119 contain the International Codebook for the IEA Reading Literacy Study. The International Codebook consists of two parts (one for Population A and one for Population B) each of which is subdivided into three sections (for student data, teacher data, and school data). (For a detailed description, see section "International Codebook" in chapter "International Electronic Database").

2.3.4 PANEL RATING

In order to provide an external framework against which to evaluate the achievement scores obtained by students at the Population B level, eleven countries took part in a small optional study designed to determine the expected literacy levels in their society. Firstly, a panel of 20 to 40 informed adult judges (i.e. people likely to be sensitive to the literacy demands of public documents and widely used reading materials in their society e.g. people representing business, universities, and health administrators) were selected in each country. Each judge was then asked to read and rate each item in the Population B test in relation to two defined standards of literacy appropriate for their education system. Finally, a comparison of the actual mean scores and the expected scores was made to determine the percentage of students who fell below the "barely literate" criterion and the percentage of students who were considered "high-level literate".

Folder 120 documents the procedures used for this exercise.

Participating countries in this exercise were: Cyprus, Finland, Germany(ftFRG), Greece, Hong Kong, Netherlands, New Zealand, Slovenia, Sweden, Switzerland, Venezuela.

2.3.5 INTERNATIONAL MEETINGS

Folder 121 contains the minutes from important meetings of the National Research Coordinators, the Steering Committee, and the Technical Advisory Committee of the IEA Reading Literacy Study.

2.3.6 CONCEPTUAL FRAMEWORK AND TEST SPECIFICATION

Folder 122 contains:

- the research proposal for the IEA Reading Literacy Study;
- a review of relevant research;
- guidelines for the item writing;
- the test specification for the reading tests.

2.3.7 LIST OF INTERNATIONAL FOLDERS

The following provides a list of the international folders of the study archive.

FOLDER	COUNTRIES	RESOURCE	DESCRIPTION
106	International	Instruments	Proposed Pilot Material
107	International	Instruments	Proposed Items from Countries BEF - DEN
108	International	Instruments	Proposed Items from Countries ENG - GRC
109	International	Instruments	Proposed Items from Countries HUN - NOR

110	International	Instruments	Proposed Items from Countries PHI - TRI
111	International	Instruments	Pilot Test Version Booklet 1 (Part 1 of 2) Booklet 1 (Part 2 of 2) Booklet 2 (Part 1 of 2) Booklet 2 (Part 2 of 2) National Case Study Questionnaire
112	International	Instruments	Suggestions for Revision On Pilot Questionnaires (General)
113	International	Instruments	Suggestions for Revision On Pilot Questionnaires (By Country)
114	International	Instruments	Proposed Final Test Version Population A Population B Final Version National Case Study Questionnaire
115	International	Instruments	Final Instruments, Population A Booklet A1 (Reading Test, Session 1) Booklet A2 (Reading Test, Session 2) Booklet A3 (Student Questionnaire) Teacher Questionnaire School Questionnaire Final Instruments, Population B Booklet B1 (Reading Test, Session 1) Booklet B2 (Reading Test, Session 2) Booklet B3 (Student Questionnaire) Teacher Questionnaire School Questionnaire
116	International	Survey Operations	Field Operation Manual (Main Study) Sampling Manual Manual for National Research Coordinators Manual for School Coordinators Manual for Test Administrators Data Processing Manual Data Entry Manual Data verification Guidelines Field Operation Manual (Pilot Study) Manual for National Research Coordinators Manual for School Coordinators Manual for Test Administrators
117	International	Codebooks (Pilot Study)	Introduction Population A, Student Datafile Population A, Teacher Datafile Population A, School Datafile Population B, Student Datafile Population B, Teacher Datafile Population B, School Datafile

118	International	Codebooks (Main Study)	Population A, Student Datafile Population A, Teacher Datafile Population A, School Datafile Population B, Student Datafile Population B, Teacher Datafile Population B, School Datafile
119	International	Codebooks (Main Study)	Introduction Population A, Merged Datafile Student, Teacher, and School Data Population B, Merged Datafile Student, Teacher, and School Data
120	International	Panel Ratings	
121	International	Meetings	
122	International	Conceptual Framework and Test Specification	Research Proposal Suggestions for Item Writing Test Specification Research Review Item Classification

3. INTERNATIONAL ELECTRONIC DATABASE

3.1 INTRODUCTION

The International Electronic Database of the IEA Reading Literacy Study is located at the IEA Secretariat. It contains the Reading Literacy datafiles as both SAS data sets and cleaned raw datafiles.

The SAS data sets contain data values and associated descriptive information in a form that can be recognized and processed by the SAS Statistical Analysis System, version 6.0.4 or later.

The raw data files are written in ASCII format (referring to the American Standard Code for Information Interchange standards), and contain only data codes. The raw datafiles are in fixed-column format according to the International Codebook. In the raw data file the variables are represented next to each other (in columns) and observations are represented below each other (each observation on one row), and data for each variable are recorded in the same location on the same record for each observation.

3.1.1 STRUCTURAL AND NAMING CONVENTION OF DATAFILES

Each question, test item, or identification code was coded in terms of one or more variables in the datafiles. The variables were thus made up of the set of data values describing a given characteristic of an observation e.g. a question or a test-item. Each variable has a type which is either numeric or character. In addition to their type, the variables have the following attributes: name, label and format. Variable attributes are either explicitly specified in a codebook (for all variables created by the national centers) or defined from the context at their first occurrence (e.g. derived variables).

3.1.1.1 IDENTIFICATION VARIABLES

Each student, teacher and school was identified by a unique identification code (ID). Unique means that each ID occurs only once in each data file. For the data collection and for all steps of the data-processing it was essential that these numbers were consistently used in all instruments and all data files because these IDs were used as references in the merging process when students from the student file were linked to their teachers in the teacher file and teachers in the teacher file were linked to their schools in the school file. A hierarchical numbering system, which repeats the Class ID and School ID in the Student ID, and the School ID and Class ID in the Teacher ID, was used for the IDs.

3.1.1.1.1 *School ID*

The School ID consisted of a unique numeric ID which identified the school from which the student or teacher was selected. Since no country tested more than 999 schools in any one population, a three-digit numerical code was used. These will be denoted by 'AAA' in the following. Format: AAA.

School ID	School
001	school #1 in the Sampling Master List
002	school #2 in the Sampling Master List
...	

3.1.1.1.2 *Class ID*

The Class ID consisted of a unique four-digit numerical code which identified the class from which the student was selected. The Class ID starts with the School ID followed by a one-digit unique numerical identification of the class within the school. These will be denoted by 'B' in the following. Format AAAB.

Class ID	Class
0011	class #1 in school #1
0012	class #2 in school #1
...	
0021	class #1 in school #2
...	

3.1.1.1.3 *Teacher ID*

The Teacher ID consisted of a unique numeric ID which identified the teacher in the school. Normally, there was only one teacher selected from each school, and in this case, the Class ID was used as the Teacher ID. However, in cases where more than one teacher had to be identified (which was normally not the case), then a one-digit unique numerical code was added to the class ID, which will be denoted by 'C' in the following. Format AAAB(C).

If one teacher was the teacher of two or more classes being tested, then the teacher was represented by multiple records in the teacher datafile. For all these records, all the questions referring to the different classes would differ. They would also have a different Teacher ID (because the fourth position in the teacher ID is the class number) for the same teacher.

Teacher ID	Teacher
0011	teacher for class #1 in school #1
0012	teacher for class #2 in school #1
...	
0021	teacher for class #1 in school #2

3.1.1.1.4 *Student ID*

The Student ID consisted of a unique six-digit numerical code which identified the student. The Student ID starts with the School ID followed by the Class ID and then followed by two-digit numerical code for the student within the class within the school, which is denoted by 'DD' in the following. Format: AAABDD.

Student ID	Student
001101	student #1 in class #1 of school #1
001102	student #2 in class #1 of school #1
...	
001201	student #1 in class #2 of school #1
...	
002101	student #1 in class #1 of school #2
...	

3.1.1.1.5 *Stratum ID*

The Stratum ID consisted of a unique two-digit numerical code. The Stratum ID will be denoted by 'EE' in the following. Format: EE.

Stratum ID	Stratum
01	stratum #1 in the Sampling Master List
02	stratum #2 in the Sampling Master List
...	

3.1.2 INTERNATIONAL CODEBOOK

The codebook defines the structure of the raw datafiles and provides the information necessary for the coding of responses from the questionnaire and the achievement tests. The International Codebook consists of two parts (one for Population A and one for Population B) each of which is subdivided into three sections (for student data, teacher data, and school data).

The codebook contains the following pieces of information:

Var. No.: The first column (Var. No.) presents a sequential number for each variable in the Reading Literacy Codebook.

Question No.: The second column (Quest. No.) presents an identification of the background question or test items, and its location in the instruments.

- For identification variables, the question number consists of a prefix, which indicate the population ("A" for population A, or "B" for population B) and the filetype ("S" for Student files, or "T" for teacher file), a slash "/" and the code "ID".
- For background questionnaires, the question number consists of a prefix, which indicate the population and the filetype, a slash "/" and the number of the question in the corresponding questionnaire.
- For the test items, the question number consists of a prefix, which indicate the population, a slash "/", one character for the domain, a short mnemonic code indicating the passage, a slash "/", and a sequential number of the item within the passage.

Variable Name: The third column (Variable Name) presents the name by which the variables are referred to.

Variable Label : The fourth column (Variable Label) presents an extended label for the variable of up to 40 characters.

Code : The fifth column (Code R:Recode) presents the codes for the responses, and the recodes for variables for which recoding is necessary and where recoding is not covered by the general notes on recoding. Whenever actual numerical data are supplied in the response to the questions, this is indicated by the keyword "VALUE". The missing-code presented in the codebook indicates "missing/non-response" values. The "not applicable" code presented in the codebook indicated "not administered" values.

Option: The sixth column Option) presents the response phrase (or an abbreviation of it) that corresponds to the code. For variables that contain actual numeric data, it contains an explanation and the permitted range of the value to be entered.

Location/Format: The seventh column (Location/Format) presents the location and format of the variable in the raw datafile. The format here refers to both the type (where N refers to numeric variables and C refers to character values) and the length (where the numeric code refers to the length of the value and the number of decimal places associated with the values) of each variable.

Comment : For some background variables, a comment indicates special coding instructions.

Coding Instructions : For some background variables, a comment gives additional information necessary for coding. For some achievement variables, it gives special coding instructions for coding free response variables. For participation indicator variables (variable TOKENn), it gives special coding instructions on what codes to assign to indicate whether the student participate in the session are given.

Passage : Provides the passage number with the passage mnemonic encloses in brackets at the end of the passage number.

Item No. : Provides the sequential number of the item within the passage.

Domain : Provides the domain of reading literacy in which the item was classified. The mnemonic codes were used for the domains:

- D=Document, Locating of Information
- E =Expository
- N=Narrative
- W=Word Recognition

Constr. : Provides the skill-constructs into which the item was classified. The following mnemonic codes were used for constructs :

- D=Interpreting directions
- I=Inference
- L =Location of information in graphics
- M=Main idea, message
- P =Paraphrase
- V =Verbatim match between words in text and item

Corr. Answer : Provides the correct answer to the item. For multiple choice items, the response option for the correct answer is indicated. For free response test items, the word "free res" is indicated.

Flags :

"VLD" specifies the range-validation criterion used by the DataEntry program.

"SCR" specifies the sequential number of the variable in which the variable appears on the screen of the DataEntry program.

"CAR" .T. indicates that in the DataEntry program, a value will be carried from one observation to the next observation as the default value during data entry.

"CAT" indicates the measurement category of the variable ("Nom." indicates nominal type, "Ord." indicates ordinal type, "Int." indicate interval type, "Rat." indicate rational variable type, "Val." indicate measurement type (actual numerical figures)

"DEF" indicates a default value for the variable in the DataEntry program.

3.1.3 FILE NAMING CONVENTIONS

For file identification, the ICC adhered very strictly to the specifications of file names given below.

3.1.3.1 FILE NAME DEFINITIONS

All file names have a length of up to 8 characters and an extension of up to 3 characters.

3.1.3.1.1 *Data File Names*

Data file names were constructed as follows:

The first letter indicates the population. 'A' indicates population A, 'B' indicates population B.

The second letter indicates the filetype. 'S' indicates a student file, 'T' indicates a teacher file, 'C' indicates a school file.

The third, fourth and fifth letter indicate the country. A missing country code in a file name of a file sent out by the ICC indicates that the file refers to all countries.

The sixth and seventh letter indicate the data-file ID. This is normally '01'. If you, for any reason, provide more than one population, filetype and country, you should assign '02', '03'... and provide the necessary documentation.

3.1.3.1.2 *Non-data File Names*

Some of the files are not concerned with a particular data file (e.g. programs or documents). This is the case when the first letter of the file name is not A, B or C.

3.1.3.2 FILE EXTENSION DEFINITIONS

The file extensions (in MS/DOS, VAX/VMS, IBM/CMS, CDC/NOS) used at the ICC are listed below in alphabetic order:

.ARC - compressed archive files, refer to section 'Data Compression for the PC' for further information.
.BAT - MS/DOS batch file. Files with this extension can directly be executed from the DOS command level.
.CDF - Codebook in spreadsheet format (can be created by the DataEntry program)
.CDP - Codebook information transformed to SPSS input statements (can be created with the DataEntry program)
.CDS - Codebook information transformed to SAS input statements (can be created with the DataEntry program)
.CDT - Codebook in ASCII text format (the ICC will sent out a printed version. Additionally this text file can be created with the DataEntry program).
.COM - MS/DOS command file. Files with this extension can directly be executed from the DOS command level.
.DBF - DataEntry system files, created by the DataEntry program. The ICC would like to receive files in this format.
.DOC - Documentation file. These files usually contain documentation in ASCII format that refers to the file specified in the filename with the extension .DOC.
.EXE - MS/DOS command file. Files with this extension can directly be executed from the DOS command level.
.SAS - SAS program listing file
.SPS - SPSS program listing file
.SDB - Codebook structure file used by the DataEntry programs.
.RAW - Raw-data file
.REP - Cleaning report file
.RDL - Raw-data file in spreadsheet format

3.1.4 LIST OF DATAFILES (POPULATION A)

COUNTRY	FILE NAME	FILE SIZE (in bytes)	FILE NAME	FILE SIZE (in bytes)
Belgium (French)	ASBEF01W.SD2	8405248	ASBEF01W.RAW	4034920
Canada (BC)	ASCAC01W.SD2	8472832	ASCAC01W.RAW	4069190
Cyprus	ASCYP01W.SD2	4724992	ASCYP01W.RAW	2251390
Denmark	ASDEN01W.SD2	10967296	ASDEN01W.RAW	5279070
Finland	ASFIN01W.SD2	4854016	ASFIN01W.RAW	2312480
France	ASFRA01W.SD2	5849344	ASFRA01W.RAW	2796730
Germany (ftGDR)	ASGER01W.SD2	9627904	ASGER01W.RAW	4627940
Germany (ftFRG)	ASGER02W.SD2	6174976	ASGER02W.RAW	2954670
Greece	ASGRC01W.SD2	11170048	ASGRC01W.RAW	5377410
Hong Kong	ASHKO01W.SD2	10260736	ASHKO01W.RAW	4936370
Hungary	ASHUN01W.SD2	9332992	ASHUN01W.RAW	4484900
Iceland	ASICE01W.SD2	12478720	ASICE01W.RAW	6012150
Indonesia	ASINS01W.SD2	9818368	ASINS01W.RAW	4721810
Ireland	ASIRE01W.SD2	8423680	ASIRE01W.RAW	4043860
Italy	ASITA01W.SD2	6973696	ASITA01W.RAW	3340580
Netherlands	ASNET01W.SD2	5327104	ASNET01W.RAW	2541940
New Zealand	ASNEZ01W.SD2	9382144	ASNEZ01W.RAW	4510230
Norway	ASNOR01W.SD2	7723264	ASNOR01W.RAW	3705630
Portugal	ASPOR01W.SD2	8712448	ASPOR01W.RAW	4183920
Singapore	ASSIN01W.SD2	22812928	ASSIN01W.RAW	11024510
Slovenia	ASYOU01W.SD2	10223872	ASYOU01W.RAW	4917000
Spain	ASSPA01W.SD2	25368832	ASSPA01W.RAW	12262700
Sweden	ASSWE01W.SD2	13437184	ASSWE01W.RAW	6477030
Switzerland	ASSWI01W.SD2	10635520	ASSWI01W.RAW	5118150

Trinidad and Tabago	ASTRI01W.SD2	11403520	ASTRI01W.RAW	5489160
United States	ASUSA01W.SD2	20760832	ASUSA01W.RAW	10026210
Venezuela	ASVEN01W.SD2	14573824	ASVEN01W.RAW	7026840
International	ASALL01W.SD2	269815552		

3.1.5

LIST OF DATAFILES (POPULATION B)

COUNTRY	FILE NAME	FILE SIZE (in bytes)	FILE NAME	FILE SIZE (in bytes)
Belgium (French)	BSBEF01W.SD2	7778048	BSBEF01W.RAW	3599361
Botswana	BSBOT01W.SD2	13494528	BSBOT01W.RAW	6272871
Canada (BC)	BSCAC01W.SD2	13849344	BSCAC01W.RAW	6438813
Cyprus	BSCYP01W.SD2	4089088	BSCYP01W.RAW	1872774
Denmark	BSDEN01W.SD2	11100928	BSDEN01W.RAW	5153421
Finland	BSFIN01W.SD2	3965184	BSFIN01W.RAW	1816143
France	BSFRA01W.SD2	7451392	BSFRA01W.RAW	3447906
Germany (ftGDR)	BSGER01W.SD2	12813056	BSGER01W.RAW	5954157
Germany (ftFRG)	BSGER02W.SD2	5609728	BSGER02W.RAW	2585271
Greece	BSGRC01W.SD2	11179776	BSGRC01W.RAW	5191614
Hong Kong	BSHKO01W.SD2	8977664	BSHKO01W.RAW	4161720
Hungary	BSHUN01W.SD2	9580288	BSHUN01W.RAW	4443558
Iceland	BSICE01W.SD2	10937600	BSICE01W.RAW	5077035
Ireland	BSIRE01W.SD2	10340608	BSIRE01W.RAW	4797831
Italy	BSITA01W.SD2	8797440	BSITA01W.RAW	4068213
Netherlands	BSNET01W.SD2	10633472	BSNET01W.RAW	4936116
New Zealand	BSNEZ01W.SD2	8910080	BSNEZ01W.RAW	4128795
Nogeria	BSNIR01W.SD2	8587520	BSNIR01W.RAW	3126558
Norway	BSNOR01W.SD2	6578432	BSNOR01W.RAW	3038319
Philippines	BSPHI01W.SD2	3148138	BSPHI01W.RAW	12785436
Portugal	BSPOR01W.SD2	9743616	BSPOR01W.RAW	4515993
Singapore	BSSIN01W.SD2	13612800	BSSIN01W.RAW	6328185
Slovenia	BSYOU01W.SD2	9174784	BSYOU01W.RAW	4253910
Spain	BSSPA01W.SD2	23975680	BSSPA01W.RAW	11174745
Sweden	BSSWE01W.SD2	10413824	BSSWE01W.RAW	4832073
Switzerland	BSSWI01W.SD2	18445056	BSSWI01W.RAW	8586840
Thailand	BSTHA01W.SD2	7828736	BSTHA01W.RAW	3620433
Trinidad and Tabago	BSTRI01W.SD2	8605952	BSTRI01W.RAW	3986559
United States	BSUSA01W.SD2	9867520	BSUSA01W.RAW	4577892
Venezuela	BSVEN01W.SD2	12278016	BSVEN01W.RAW	5697342
Zimbabwe	BSZIM01W.SD2	7873792	BSZIM01W.RAW	3645456
International	BSALL01W.SD2	299638016		

4. DATA VERIFICATION RULES AND DATA QUALITY STANDARDS

4.1 INTRODUCTION

This chapter presents the rules and procedures that were applied to the data submitted by National Centers to the ICC.

4.1.1 DETECTING DATA VERIFICATION PROBLEMS

Incoming data were checked by a number of data verification procedures:

- one set of these data verification steps checked the data variables of one observation against the validation criteria specified in the international codebook;
- a second set of data verification steps checked the data variables against certain control variables within an observation;
- a third set of data verification steps checked certain data variables against other data variables for formal consistencies;
- a fourth set of data verification steps checked data variables across observations within a datafile;
- a fifth set of data verification steps checked data variables across observations across datafiles.

The criteria on which the checking was based depended, on the one hand, on the type of the variables (data variables, ID variables, control variables, filter variables, dependent variables) and, on the other hand, on the way and sequence in which questions and items were asked (e.g. for some questions a certain number of responses was required, or responses had to be given in a special way, or there was dependency, or a logical relationship between questions which needs to be checked).

4.1.2 RESOLVING DATA VERIFICATION PROBLEMS

Some data verification problems could be resolved by cross-checking certain data values against other data values within the same observation or by cross-checking against other observations. The ICC tried to do this on the basis of *formal data verification rules* for the problems which are described in the following sections.

Some data verification problems required further inspection and manual resolving whereas other problems could be resolved automatically during detection.

For some data verification problems, data values have been set to special missing values.

Each data verification problem required answers to the following questions, which have been agreed upon individually with the countries for which problems have been found:

- when to set a data value to not administered, logical not applicable or invalid;
- when to correct a data value because of its relationship to other data values;
- when to drop an observation because of invalid, missing or not administered data;
- when to drop a variable;

One aim was to retain maximum data during the process of data verification. Therefore data values, variables and observations were only dropped when retaining them would lead to severe inconsistencies in later analyses. Thus, in some cases minor inconsistencies in the data were retained.

In many cases data verification problems were only flagged by special control variables and it is left to the analysts to decide how to handle these observations or variables during analysis. This was especially the case when there were inconsistencies in the answers of respondents to different questionnaires.

The sequencing of data verification steps as outlined below were essential, especially in the case of ID problems.

4.1.3 CHECKS UNDERTAKEN IN THE COUNTRIES BEFORE CODING AND ENTRY OF DATA

National Centers were asked to undertake a number of checks and take corrective measures where necessary before submitting their data to the ICC. *These checks could only be partially verified at the ICC.* In particular it was requested that:

- the Student Name Forms were verified against the Sampling Master-List for completeness;
- the figures describing the defined target population in the school file were verified against the information in the Sampling Master List;
- the IDs on the Student, Teacher and School Instruments were verified against the corresponding Student Name Form;
- the students answers on age, sex and language tested were verified against the Student Name Form;
- all students (except those that were crossed out in red in the Student Name Form because they had left the school permanently) that were listed in the Student Name Form were entered onto the student file and that the variable EXCLNRD was coded accordingly.

4.1.4 HANDLING OF MISSING DATA

4.1.4.1 CODING OF MISSING DATA

The coding *as undertaken in the countries* required National Centers to distinguish between missing values that are missing because a variable was not administered, and missing values due to non response or invalid responses.

Variables were recoded to the missing code *"not administered (.A)"* when the question or test item had to be dropped because of mistranslation, or because it had been asked, administered or coded in a way that did not allow its use in the international analyses. National Centers were informed when this was done.

When a respondent was not meant to answer a variable because of its logical relationship to other variables, these variables were recoded to the missing code *"logical not applicable (.B)"*. For example, if a respondent gave a negative answer to a filter question, then the corresponding dependent questions were recoded to "logical not applicable" unless all dependent variables indicate that the filter variable was incorrectly coded. In this case the filter variable was recoded.

If there were closed variables whose codes did not conform to the range validation criteria specified in the International Codebook, they were recoded to the missing category *"invalid (.I)"*.

Test item variables were recoded to the missing code *"not reached (.R)"*, when the item was considered as not having been reached by the student during test administration. Items at the end of each testing session, which were not responded to, were considered as *"not reached"*.

The different missing categories are defined below:

4.1.4.1.1 *Omit (SAS code ".")*

Omit values refer to questions/items which a student, teacher or school principal should have answered but which he/she either did not answer or which were *answered* in an invalid way. Some obvious reasons for coding a variable to missing are:

- *No Response* - Where there was no response to a question or an item where there should be one.
- *Two Or More Responses* - Where there were two or more responses when only one answer was allowed.
- *Response Unreadable* - Where the response was unreadable or uninterpretable.

4.1.4.1.2 *Not administered (SAS code ".A")*

"Not administered" values must be distinguished from "missing" data. "Not administered (.A)" was assigned when data were not collected for an observation on a specific variable. In the Pilot Study, it was often the case that "missing" was punched instead of "not administered" or vice versa. Normally, "not administered" codes should occur for several cases at the same time. This was checked. There are some obvious cases:

- *Student Not Present* - If a student was not present in a particular testing session, then all variables referring to that session were supposed to be coded to "not administered". However, if the student received the instrument but did not answer particular questions, then these questions were to be coded to "omit" by the National Center !

- *Booklet Not Received* - If a student did not receive a particular test instrument then all variables referring to that test instrument were to be coded to "not administered".

- *Item Left Out Or Misprinted* - If a particular question or item (or a whole page) was misprinted, left out, or not available to a student, teacher or school then the corresponding variables were to be coded to "not administered" by the National Center.

- *Item mistranslated* - If an item was mistranslated by a National Center, then all observations for this item were also to be coded to "not administered".

4.1.4.1.3 *Logical not applicable (logical relation) (SAS code ".B")*
When a respondent was not meant to answer a variable because of its logical relationship to other variables, these variables were recoded to the missing code *"logical not applicable" (.B).*

4.1.4.1.4 *Invalid (SAS code ".I")*
Data *recorded* in an invalid or inconsistent way have in some cases been recoded to a special missing code *"invalid" (.I).* In this sense, "invalid" means that data were *recorded* in an invalid way, i.e. that the coder coded a variable to a data value that did not conform to the specifications in the International Codebook; this does not necessarily mean that the respondent gave an invalid response. The invalid code ".I" was a special missing code that was assigned at the ICC in special data verification procedures in order to distinguish invalid data from "omit" and "not administered" data.

4.1.4.1.5 *Not reached (SAS code ".R")*
Test items located at the end of each testing session which were left blank had been considered as *"not reached "*

4.1.4.2 IMPUTATION OF MISSING DATA
Imputing for missing data was performed as described in section "Data verification rules applied to inconsistencies".

4.1.4.3 PRESENTATION OF MISSING DATA
Missing data was presented in the univariate reports as follows:

N TOT	represents the unweighted total number of observations for a variable in the datafile.
N OMIT	represents the unweighted number of observations with "omit" codes. This is the number of observations who have been presented the question but did not answer.

N INVLD	represents the unweighted number of observations with "invalid" codes.
N NA	represents the unweighted number of observations with "not administered" codes.
N LOGNA	represents the unweighted number of observations with "logical not applicable" codes. This code is assigned to dependent questions for which the filter question was answered negative.
N ANSWER	represents the unweighted number of observations with "not administered" and "logical not applicable" values excluded, but "omit" and "invalid" included. This is the number of observations who have been presented the questions. The calculation of percentages is based on this value.
N MISS	represents the unweighted number of observations with "not administered", "invalid", "logical not applicable" or "omit" codes, i.e. for which values have any missing category.
N	represents the unweighted number of observations on which the calculation of statistics, except percentages, is based, i.e. the number of observations with "not administered", "invalid", "logical not applicable" and "omit" excluded.
% N OMIT	represents the percentage of N OMIT and N INVLD based on N ANSWER/N INVLD
% MISS	represents the percentage of N MISS based on N TOT.

4.1.5 CONTROL VARIABLES AND PARTICIPATION INFORMATION

4.1.5.1 CODING OF PARTICIPATION INDICATOR VARIABLES

At the beginning of the items or questions in the datafile for each testing session a *"Participation indicator Variable"* has been inserted. These participation indicator variables have the names *TOKEN01, TOKEN02, TOKEN03* and indicate whether the student participated in the corresponding testing session or not. The variable TOKEN01 refers to session 1 (Reading Test), TOKEN02 refers to session 2 (Reading Test) and TOKEN03 refers to session 3 (Questionnaires). These variables were coded from the Student Name Form. National centers were instructed to code these variables as follows:

- the code "3" had to be assigned if the student participated in the corresponding testing session (the normal case);
- the code "2" had to be assigned if the student was absent from the corresponding testing session or excluded. This was indicated by a cross (x) in the columns 6,7,8 of the Student Name Form;
- the code "1" had to be assigned if the booklet of the student was lost. This was indicated in column 9 (Remarks) of the Student Name Form.

4.1.5.2 IDENTIFICATION OF EXCLUDED STUDENTS

In column 9 (Remarks) of the Student Name Form, there was a 1,2,3 or 4 marked next to those students who were excluded from the testing. This was coded in the variable EXCLNRD. The default code for this variable was "omit" which indicates that this student was *not* excluded. If the student was excluded from a testing session, then TOKEN01, TOKEN02 had to be coded to 2. This was cross-checked at the ICC and the necessary corrections were undertaken. National Centers were instructed to code the EXCLNRD variable as follows:

- when the student was absent in all testing sessions or his booklet was lost then this variable had to be coded to "1";
- when the student was mentally or physically handicapped as defined by the National Center to such an extent that the student was deemed not to be able to complete a paper-and-pencil test, the code "2" had to be assigned;
- when the student was so unfamiliar with the language of the test that the student would not even be able to follow the general instructions, the code "3" had to be assigned. This student will be regarded as a non-reader;
- for other reasons (to be defined by the National Center) which would make the testing of the student meaningless the code "4" had to be assigned. This student will be regarded as a non-reader;
- when students were not excluded (the normal case) then this variable had to be coded to "omit".

Students who were excluded had to be marked in columns 6,7,8 by a cross (x) in the Student Name Form.

4.1.5.3 HANDLING OF EXCLUDED STUDENTS

Students defined as excluded in the data file with EXCLNRD=1 were removed from the datafile after merging, and thus were excluded from the calculation of weights and from all analyses.

Students defined as excluded in the data file, due to reasons 2, 3 and 4 (i.e. EXCLNRD=2,3,4) were included in the calculation of student weights. However, these students will be excluded from most of the analyses, i.e. assigned a weight of zero.

Teacher and school weights were based on the student weights including students with EXCLNRD=2,3,4.

In the data verification documentation sent to National Centers, a table was produced for each country listing the weighted number of excluded students for each of the categories.

The following describes the difference between two hypothetical countries, where in one country, 10000 "handicapped" students had already been excluded from the survey population, whereas in the other country, only 2000 had been excluded from

the survey population but 8000 were excluded during the test administration by setting EXCLNRD=2:

	Country 1	*country 2*	
Estimated population size	100000		100000
Excluded by definition		10000	2000
Remainder	90000	98000	
Excluded by EXCLNRD=2		0	8000
Remainder	90000	90000	

The above population figures are derived in the following way: The population size is estimated as the sum of weights on the basis of all students except where EXCLNRD=1. From this the number of students excluded by the survey definition was subtracted. Then the number of students with EXCLNRD=2,3,4 in the population was estimated as the sum of weights of these students and afterwards this figure was subtracted from the above result. The remainder was used as the estimator for the population size in the analyses.

4.1.6 NATIONAL RECODINGS

All National Centers were meant to indicate each and every coding deviation from the international standard in the Deviation Forms. All deviations were checked for whether they had an impact on international comparability. In order to obtain comparability and to retain most of the given information, the ICC performed national recodings where possible.

Recodings were considered in the following cases:

- a country changed the sequence of two items so that item A was placed as item B and they were also coded correspondingly. In this case, the values for item A had to be exchanged by the values for item B;
- a country had asked a question in another way than expected in the international format, e.g. the scale was changed. In this case, the variables had to be recoded to the international scaling if possible;
- a country did not ask a certain question but the missing information could be derived from other variables;
- a country detected suspicious data for a certain question and asked for cross-checking against other variables;
- the answer for a question had to be the be same for all observations because no deviation was possible, e.g. for the type of school. A few countries indicated that they do not have private schools and that all schools were state schools. In this case, the code for this variable was preset, no matter what had been coded.

4.2 DATA VERIFICATION STEPS APPLIED AT THE ICC

4.2.1 VERIFICATION OF FILE INTEGRITY

Files were first checked for file integrity. Files which were submitted in the form of DataEntry system files had a built-in check for file integrity.

4.2.2 VERIFICATION OF CONTROL VARIABLES

4.2.2.1 VERIFICATION OF THE CHECKN VARIABLES

Check variables were inserted in certain places in the international codebook. These check variables always had to be coded to BLANK. The codes "8" and "9" were also interpreted as blank at the ICC because in the Pilot Study it was seen that it was a common error to punch missing instead of blank. A non-blank value for any check variable may be due to one or more of the following reasons:

- a single mispunch on one of the CHECK variables;
- a column shift for the current observation (e.g. a coder might have punched a value for a certain data value twice and thus codes in the following columns were wrong). In this case a check was made to see whether subsequent CHECK variables also had invalid values and corresponding corrections were undertaken;
- a global column shift for the whole datafile which means that the variables were not coded in the format specified in the international codebook.

4.2.2.2 VERIFICATION OF THE PARTICIPATION INDICATOR VARIABLES AGAINST DATA VARIABLES

In the student file, variables were sequenced in blocks, beginning with a participation indicator variable, followed by the data variables and ending with a CHECK variable.

This data verification step checked whether the participation indicator variables matched the data in the corresponding data variables.

The data verification step detected:

- observations for which a participation indicator variable was coded to "3" but two or fewer variables were coded different from "not administered". In this case the corresponding participation indicator variable was recoded to "2";
- observations for which a participation indicator variable was coded to "3" but all items were coded to missing. These cases were only flagged and the participation indicator was <u>not</u> recoded. This means that a student did not answer any of the test items although he or she participated in the test session. These cases should be analyzed carefully since in the Pilot Study it was a common error that "missing" was punched instead of "not administered";
- observations for which the participation indicator variable was coded to "1" but 10 or more of the corresponding questions or test items were coded differently from

"not administered". In this case the participation indicator variable was recoded to "3";

If the TOKENn contained a value different from 1,2,3, then it was coded to "3" if more than 10 items had non-missing codes, otherwise to "2".

4.2.2.3 VERIFICATION OF THE PARTICIPATION INDICATOR VARIABLES AGAINST THE EXCLNRD-VARIABLE

The data verification step "checking participation indicator variables against data variables" was applied before this step because only the corrected participation indicator variables could be cross-checked against the variable EXCLNRD.

If both of the (corrected) variables TOKEN01, TOKEN02 were coded to "3" and the EXCLNRD variable was coded to "1", "2", "3" or "4", this indicated that the participation indicators were inconsistent with the variable EXCLNRD. If the ICC could not resolve the problem, it was assumed that the student was meant to be excluded, but was nevertheless administered the test. In this case, no recoding of the variable EXCLNRD was undertaken, but TOKEN01 and TOKEN02 were recoded to 2. In some cases the National Center indicated that the EXCLNRD variable was miscoded and priority should be given to the TOKEN variables. In this case, the variable EXCLNRD was recoded to missing which is the code for non-excluded students.

When both TOKEN01 and TOKEN02 were coded to "1" or "2" (student absent or excluded) *and* the EXCLNRD variable was coded differently from "1", "2", "3" or "4", then the EXCLNRD variable was recoded to "1".

4.2.3 VERIFICATION OF IDENTIFICATION VARIABLES

For each country a table was produced with the following information to identify inconsistent information, and for cross-checking against the Sampling Master Lists:

School file	*Teacher file*	*Student file*
School ID	School ID	School ID
STRSAMP	Class ID	Class ID
SCHSAMP	Teacher ID	Teacher ID
<pop>CENRSB	<pop>TCLSIZ	
<pop>CENRSG		
<pop>CENRGB		
<pop>CENRGG		

Note that if a teacher taught two classes, he or she was interpreted as two teachers in the analyses and is thus assigned two different Teacher IDs.

4.2.3.1 FORMAL ID CHECK

All IDs were checked for numeric values; non-numeric IDs were set to invalid

If the ID-numbering system proposed by the ICC was used, then Student, Teacher and School IDs within each observation were checked for formal integrity and the following corrections were undertaken:

4.2.3.1.1 *Correction of inconsistent IDs in the student file*

- All IDs matched except the School ID. In this case the School ID was corrected correspondingly.
- The School ID did not conform to the Class ID but the School ID matched the Teacher ID and the Student ID. In this case the Class ID was corrected correspondingly.
- The School ID did not conform to the Teacher ID but the School ID matched the Class ID and the Student ID. In this case the Teacher ID was corrected correspondingly.
- The School ID did not conform to the Student ID but the School ID matched the Teacher ID and the Class ID. In this case the Student ID was corrected correspondingly.
- The Student ID matched the Teacher ID and the Class ID did not conform to the teacher and the Student ID. In this case it was assumed that the Student ID and Teacher ID were correct. The Class ID was corrected correspondingly.
- The Student ID matched the Class ID and the Teacher ID did not conform to the class and the Student ID. In this case it was assumed that the student and the Class ID were correct. The Teacher ID was corrected correspondingly.
- The Teacher ID matched the Class ID and the Student ID did not conform to the class and the Teacher ID. In this case it was assumed that the teacher and the Class ID were correct. The Student ID was corrected correspondingly.

4.2.3.1.2 *Correction of inconsistent IDs in the teacher file*

- The School ID did not conform to the Teacher ID and the Class ID but the Class ID matched the Teacher ID. In this case it was assumed that the Class ID and the Teacher ID were correct. The School ID was corrected correspondingly.
- The Class ID did not conform to the Teacher ID and the School ID but the School ID matched the Teacher ID. In this case it was assumed that the School ID and the Teacher ID were correct. The Class ID was corrected correspondingly.
- The Teacher ID did not conform to the School ID and the Class ID but the Class ID matched the School ID. In this case it was assumed that the Class ID and the School ID were correct. The Teacher ID was corrected correspondingly.

4.2.3.1.3 *Correction of inconsistent Stratum IDs*

The following step was applied if it could be assumed that all records belonging to one school had been entered in a consecutive way. Three schools were checked simultaneously.

If the Stratum ID of the third observation did not equal the Stratum ID of the second observation but the School IDs for all the three observations were the same and the Stratum IDs for the first two observations were equal, then the Stratum ID for the third observation was corrected by assigning the value of the second observation.

Further cross-checks of Stratum IDs were undertaken in the merging process and, in some cases, National Centers were contacted and the necessary corrections made.

4.2.3.2 HANDLING OF DUPLICATE IDS

If there were duplicate Student IDs, Teacher IDs or School IDs and the corresponding records were identical then the second record was dropped. Where several duplicate records were found, then all records after the first were dropped. If the IDs were duplicate but the records were different then an attempt was made to correct the problem manually. If the problem could not be corrected then a consecutive ID was assigned to the second or further records.

4.2.3.3 HANDLING OF INCONSISTENT IDS

If the hierarchical ID system had not been used, the following problems which could occur had to be corrected manually by using the Sampling Master List or by referring back to the countries:

- students with the same Class ID but a different School ID;
- students with the same Teacher ID but a different School ID;
- students with the same Class ID but a different Teacher ID (some of these problems were not necessarily errors in the data. For example, one teacher could teach in more than one school).

Inconsistent stratum IDs across filetypes were corrected manually by cross-reference to the Sampling Master List. In cases where the Stratum IDs of two filetypes were inconsistent with the Sampling Master List the countries were contacted.

4.2.4 MERGING OF STUDENT, TEACHER AND SCHOOL DATA

After the above data verification steps had been undertaken, the student, teacher and school datafiles were merged.

There are several reasons why files might not merge:

- the Teacher ID or School ID is mispunched in the student file;
- the Teacher ID is mispunched in the teacher file;
- the School ID is mispunched in the school file;
- the teacher with this Teacher ID did not return his questionnaire or it was lost;
- the school principal with this School ID did not return his questionnaire or it was lost

4.2.4.1 BOTTOM UP MERGE

First, students were merged to the teachers and then merged to the schools. The following problems were cross-checked:

- *Students with no teacher matches.* If this occurred for a whole class and the following top-down merge check found teachers with no student matches, then a check was made against the documentation and the school file whether these could be matched. If it occurred for single students this indicated problems in the Student ID variables. The IDs of the student were then verified and, if possible, corrected manually.
- *Students with no school matches.* If this occurred for a whole school class and the and the following top-down merge check found schools with no student matches then a check was made against the documentation and the teacher file to see whether these could be matched. If the no-match occurred for single students this indicated problems in the Student ID variables and thus the IDs of the student were verified and, if possible, corrected manually.
- *Teachers with no school matches.* A check was made to see whether there were students for this teacher. If there were no student matches, this indicated a problem in the Teacher ID and the Teacher ID was verified manually.

4.2.4.2 TOP DOWN MERGE

- *School with no teacher matches.* The School IDs were verified against the Sampling Master List and, if necessary and possible, corrected.
- *Schools with no student matches.* The School IDs were verified against the Sampling Master List and, if necessary and possible, corrected. If the school could not be linked to a class, the school record was deleted from the datafile.
- *Teachers with no student matches.* The Teacher IDs were verified against the school file and, if necessary and possible, corrected. If the teacher could not be linked to a class, the teacher record was deleted from the datafile.

4.2.4.3 MERGE INDICATOR VARIABLES

For later analyses, flag variables were created in the SAS datasets indicating merging problems. The following variables were used:

- SYSMST "1" indicates a student with no teacher match
- SYSMTS "1" indicates a teacher with no student match
- SYSMSC "1" indicates a student with no school match
- SYSMTC "1" indicates a teacher with no school match
- SYSMSTC "0" indicates a complete record.
- IN_TCH "1" indicates when teacher contributed in merge T->S

- IN_STD1 "1" indicates when student contributed in merge T->S
- IN_SCH "1" indicates when school contributed in merge C->S,T
- SF_STD "1" indicates first student for teacher
- SL_STD "1" indicates last student for teacher

- FRST_STR "1" indicates first occurrence of stratum in merged file
- FRST_SCH "1" indicates first occurrence of school within stratum in merged file
- FRST_TCH "1" indicates first occurrence of teacher within school in merged file
- FRST_STD "1" indicates first occurrence of student within teacher in merged file

4.2.5 VALUE VALIDATION

This step checks whether the values fit the range validation criteria.

Background questions and test items for which codes rather than open values applied were checked against range validation criterion as defined in the international codebook. Variables for open values were checked against the ranges supplied by the countries in the Valid Range Form.

All values that did not conform were set to the missing code "invalid".

If the number of invalid answers was high, then the National Centers were requested to reconsider the ranges they had supplied and to check their national test instruments for mistranslations.

If a country revised their valid ranges, then this data verification step was applied once again.

When the "Quality of School Life" set questions in the Pop B Student Questionnaire was not administered then all variables referring to this question were recoded to "not administered".

When free response test items (in the Reading Tests) were not administered then all variables referring to these questions were recoded to "not administered".

Special National Options were only checked for the range validation criteria. Missing/not administered was not verified.

4.2.6 CONSISTENCY CHECKS

The criteria on which the consistency checks are based depend on the way and sequence in which questions and items were asked (e.g. for some questions a certain number of responses is required, or responses must be given in a special way).

Some questions can be answered independently of each other, but since they may have logical relationships with other questions within the same test instrument or across instruments, they need to be checked.

For data verification purposes inconsistencies were classified as follows:

- inconsistencies between questions within an observation (e.g. inconsistencies between answers to a dependent and the corresponding filter question);
- inconsistencies for a certain question between observations of a certain subset (e.g. different answers of students of the same class to variables referring to a class as, for example, questions on homework);
- inconsistencies between questions across filetypes, e.g. inconsistencies between answers of the school principal and the teacher;
- inconsistencies between class- and school level aggregates of student questions and teacher and school questions.

4.2.6.1 CONSISTENCY CHECKS FOR ZERO-VALUES VERSUS "OMIT"

The experience with the pilot study revealed that questions with open codes often had "omit" coded where "zero" should have been coded.

This was a particular problem for those questions which did not applied in certain situations. *For example, in a girls-only school the school principal might have left out the question "how many boys are enrolled in the class tested" rather than having answered it with zero.* It was also a problem for questions that had been split, e.g. asked in terms of hours and minutes. In these cases often only hours were given and minutes had then to be interpreted as zero.

All of these variables are listed below. In later sections, the data verification of these variables is described.

- ATSCHED, (AT/3), N YEARS OF SCHOOL EDUCATION
- ATTETRA, (AT/4), N YEARS OF PRE-SERVICE TEACHER TRAINING
- ATSECED, (AT/5), N YEARS OF POST-SECONDARY EDUCATION
- ATOTLAN, (AT/13), N OTHER LANGUAGE STUDENTS
- ATNHELP, (AT/15), N STUDENT NEED REMEDIAL HELP
- ATRHELP, (AT/16), N STUDENT RECEIVE REMEDIAL HELP
- ATTLANH, (AT/18), TIME TEACHING TEST LANGUAGE/HOURS
- ATTEREH, (AT/19), TEACHING AND PRACTICE OF READING

- ATCRNUM, (AT/26), NUMBER OF GROUPS FOR GROUP WORK
- ACENRSB, (AC/2A), TOTAL ENROLLMENT OF FULL-TIME STUD/BOYS
- ACENRSG, (AC/2B), TOTAL ENROLLMENT OF FULL-TIME STUD/GIRLS
- ACENRGB, (AC/3A), TOTAL ENROLLMENT OF STUD IN GRADE/BOYS
- ACENRGG, (AC/3B), TOTAL ENROLLMENT OF STUD IN GRADE/GIRLS
- ACSLIBC, (AC/10), SCHOOL LIBRARY/N BOOKS ADDED LAST YEAR
- ACNFITM, (AC/12A), N OF FULL-TIME TEACHERS/MALE
- ACNFITF, (AC/12B), N OF FULL-TIME TEACHERS/FEMALE
- ACNFSTM, (AC/13A), N OF FULL-TIME SPECIAL TEACHERS/MALE
- ACNFSTF, (AC/13B), N OF FULL-TIME SPECIAL TEACHERS/FEMALE
- ACDAYLO, (AC/16), N DAYS OF INSTRUCTION LOST LAST YEAR
- BSFEDUC, (BS/4), N YEARS FATHERS EDUCATION
- BSMEDUC, (BS/5), N YEARS MOTHERS EDUCATION
- BSSIREH, (BS/25A), HOURS SPENT PER WEEK READING SILENTLY
- BSSIREM, (BS/25B), MINUTES SPENT PER WEEK READING SILENTLY
- BTSCHED, (BT/3), N YEARS OF SCHOOL EDUCATION
- BTTETRA, (BT/4), N YEARS OF PRE-SERVICE TEACHER TRAINING
- BTSECED, (BT/5), N YEARS OF POST-SECONDARY EDUCATION
- BTOTLAN, (BT/9), N OTHER LANGUAGE STUDENTS
- BTNHELP, (BT/10), N STUDENTS NEED REMEDIAL HELP
- BTRHELP, (BT/11), N STUDENT RECEIVE REMEDIAL HELP
- BTTLANH, (BT/13A), TIME TEACHING LANGUAGE OF TEST/HOURS
- BTTLANM, (BT/13B), TIME TEACHING LANGUAGE OF TEST/MINUTES
- BCENRSB, (BC/2A), TOTAL ENROLLMENT OF FULL-TIME STUD/BOYS
- BCENRSG, (BC/2B), TOTAL ENROLLMENT OF FULL-TIME STUD/GIRLS
- BCENRGB, (BC/3A), TOTAL ENROLLMENT OF STUD IN GRADE/BOYS
- BCENRGG, (BC/3B), TOTAL ENROLLMENT OF STUD IN GRADE/GIRLS
- BCSLIBC, (BC/9), SCHOOL LIBRARY/N BOOKS CONTAINED
- BCSLIBA, (BC/10), SCHOOL LIBRARY/N BOOKS ADDED LAST YEAR
- BCNFITM, (BC/12A), N FULL-TIME TEACHERS/MALE
- BCNFITF, (BC/12B), N FULL-TIME TEACHERS/FEMALE
- BCNFSTM, (BC/13A), N FULL-TIME SPECIAL TEACHERS/MALE
- BCNFSTF, (BC/13B), N FULL-TIME SPECIAL TEACHERS/FEM
- BCDAYLO, (BC/16), DAYS OF INSTRUCTION LOST LAST YEAR

- BCSTABS, (BC/17). PERCENTAGE OF STUDENTS ABSENT

4.2.6.2 CONSISTENCY CHECKS FOR ZERO VALUES VERSUS "LOGICAL NOT APPLICABLE"
For some variables the logical not applicable code .B may imply zero in the interpretation.

All of these variables are listed below. In later sections, the data verification of these variables is described.

- ATTETRA, (AT/4) dependent of ATSCHED, (AT/3)
- ATSECED, (AT/5) dependent of ATSCHED, (AT/3)
- ATRHELP, (AT/16) dependent of ATNHELP, (AT/15)
- ATTEREH, (AT/19) dependent of ATTLANH, (AT/18)
- ATGRNUM, (AT/26) dependent of ATGROUP, (AT/24) and ATGRTYP, (AT/25)
- ACYEARP, (AC/1B) dependent of ATYEART, (AC/1A)
- ACENRGB, (AC/3A) dependent of ACENRSB, (AC/2A)
- ACENRGG, (AC/3B) dependent of AVENRSG, (AC/2B)
- ACSLIBC, (AC/9) dependent of ACSCRE1, (AC/8A)
- ACLIBA, (AC/10) dependent of ACSCRE1, (AC/8A)
- ACNFSTM, (AC/13A) dependent of ACNFTTM, (AC/12A)
- ACNFSTF, (AC/13B) dependent of ACNFTTF, (AC/12B)
- BTTETRA, (BT/4) dependent of BTSCHED, (BT/3)
- BTSECED, (BT/5) dependent of BTSCHED, (BT/3)
- BTRHELP, (BT/11) dependent of BTNHELP, (BT/10)
- BCYEARP, (BC/1B) dependent of BCYEART, (BC/1A)
- BCENRGB, (BC/3A) dependent of BCENRSB, (BC/2A)
- BCENRGG, (BC/3B) dependent of BCENRSG, (BC/2B)
- BCSLIBC, (BC/9) dependent of BCSCRE1, (BC/8A)
- BCSLIBA, (BC/10) dependent of BCSCRE1, (BC/8A)
- BCNFSTM, (BC/13A) dependent of BCNFTTM, (BC/12A)
- BCNFSTF, (BC/13B) dependent of BCNFTTF, (BC/12B)

4.2.6.3 CONSISTENCY CHECKS FOR SPLIT VARIABLES
Some questions were asked and/or coded in terms of more than one variable. The data verification of such variables depended on whether they were related to each other or independent and on whether open codes or closed codes were used.

The data verification further depended on the design of the questions. For example, in some cases the question was split into several options and the respondent could either check or omit an option whereas in other cases the respondent could check each response option.

In the cases listed below the question was split into several options and the respondent could either check an option or omit it. In the analysis it was important to know whether a respondent omitted the whole question or did not check a particular option. The coders were supposed to code the checked options to "2", the not-checked options to "1". Only if none of the options were checked, then all options were meant to be coded to "9". It was assumed that, if none of the options had been checked, the question was skipped. It should be noted that this was only obvious for questions for which at least one of the options had to be checked. These cases are marked by an asterisk in the list below.

From the experience of the pilot study, it was noted that the distinction between missing and not-checked was often handled incorrectly. Therefore the following recoding had to be undertaken. If at least one option was checked, then all other options which were coded to missing, were recoded to "not checked"(1). If all options were coded to "1" then all options were recoded to "9" because in this case it was clear that the respondent skipped the entire question. This was only obvious where at least one option had to be checked (marked by an asterisk).

- ASWAY01..11, (AS/16A-K) 3 IMPORTANT WAYS
- *ASMAT1..8, (AS/37A-H) MATERIAL READ ALOUD
- TENST1..5, (AT/29A-E) ENCOURAGE STUDENTS
- *ACSPPR1..5, (AC/18A-D) SPECIAL PROGRAMS
- *ACPROB1..6, (AC/21A-F) 2 PROBLEMS
- *ACPROC1..5, (AC/24A-F) PROCEDURES
- BSWAY01..11, (BS/24A-K) 3 IMPORTANT WAYS
- BTENCO1..5, (BT/16A-E) ENCOURAGE STUDENTS
- *BCSPPR1..5, (BC/18A-D) SPECIAL PROGRAMS
- *BCPROB1..6, (BC/21A-F) 2 PROBLEMS
- *BCPROC1..6, (BC/24A-F) PROCEDURES

4.2.6.4 CONSISTENCY CHECKS FOR SPLIT OPEN ENDED QUESTIONS

Data verification was especially important for split variables with open codes, in particular when those were not answered or only partly answered. (Some of the variables listed below will be treated like split variables although they are coded as independent variables).

Recoding was sometimes necessary in order to avoid an excessive loss of information in the creation of derived variables.

For developing data verification rules the following cases were considered:

- if the question presupposed a composite answer (e.g. years and months or hours and minutes) and only one component of the answer was given, then the missing answer may be interpreted as zero or missing.

- if the question presupposed a simple answer (e.g. number of students who did not have language of the test as their first language) then the missing answer may be better treated as a measure of central tendency (when the question dealt with matters of general policy, e.g. years of teachers" school education) or was left to missing (when it was a question that implied an idiosyncratic answer, e.g. number of students that required or received remedial help in reading)

4.2.6.5 CONSISTENCY CHECKS FOR VARIABLES FOR WHICH ONLY CERTAIN COMBINATIONS OF ANSWERS ARE ALLOWED

There were some questions where just a fixed number of "yes" "no" questions were asked. In these cases the following data verification rule was applied:

If a fixed number X of responses is requested and more than X answers were given, then X responses are selected at random.

There were other questions asking for ratings. In these cases, the requested number of answers was impaired with the number of options provided using the following data verification rules:

If the answers provided were less or equal to the requested number of options, then no recoding would be done; if the answers provided were more than the requested number of options, then the options would be sorted by their response value and a random number. Then the first "n" options (where "n" equals the number of requested options) would be selected.

Variables for which only certain combinations of answers were allowed are listed in the following.

- ATAIM01..12 (AT/22), the teacher was requested to rank 5 of 12 aims in reading instructions with ranks from 1 to 5 where each rank should be given only once.
- ASWAY01..11 (AS/16), the student was requested to check 3 of the 11 possible answers.
- ACPROB1..5 (AC/21), the school principal was requested to check 2 problems.
- ACACTI1..8 (AC/22), the school principal was requested to rank the 8 options with the numbers 1 to 8 where each rank should be given once.
- BSWAY01..11 (BS/24), the student was requested to check 3 of the 11 possible answers. See ASWAY01..11 for further information.
- BTASME1..7 (BT/14), the teacher was asked to rank all the 7 options with numbers from 1 to 7 where each rank should be given once only. See ACACTI1..8 for more details, but notice that the sum equals 28.
- BTENCO1..5 (BT/16), the teacher was asked to check 2 of the 5 options. The data was handled as described above for ACPROB1..5.
- BTAIM01..11 (BT/18), the teacher was requested to rank 5 of 12 aims with ranks from 1 to 5 where each rank should be given only once. The data was handled as described for ATAIM01..12.

- BCPROB1..5 (BC/21), the school principal was asked to check 2 of the 5 options. The data was handled as described for the corresponding question in Pop A (ACPROB1..5).

- BCACTI1..8 (BC/22), the school principal was asked to rank all the 8 options with numbers from 1 to 8 where each rank should be given once only. The data was handled as described for the corresponding question in Pop A (ACACTI1..8).

4.2.6.6 CONSISTENCY CHECKS FOR BETWEEN AND WITHIN QUESTIONS DEPENDENCIES (FILTER QUESTIONS)

There were several filter questions in the questionnaires.

In some of them the relation between filter and dependents questions was made explicit by the statement: "If you answered "No" to question x then please go to question y".

In other questions the dependency was not explicitly stated, but the pattern was the same: the answer to a first question should condition the answer to some following questions.

There were also questions where the dependency was internal, demanding a between options consistency checking or a combined internal and filter questions consistency checking.

In consistent patterns of answers, when a filter question was answered "No", its dependent questions should either have been skipped (in case of explicit dependence) or answered in a negative way. In both cases they would be recoded to "logical not applicable"(.B) and they would be excluded from the analysis, because statistics were based only on filter questions with positive answer.

All filter-dependent relations are outlined below.

Filter variables and dependent variables can have missing, negative or positive values.

- neg. filter variable: filter was set (e.g. there is no school library)

- pos. filter variable: filter was not set (e.g. there is a school library)

- neg. dependent var.: dependent variable conforms to a negative filter variable (e.g. there are no books in the school library)

- pos. dependent var.: dependent variable conforms to a positive filter variable (e.g. there are books in the school library)

All cases where data verification was necessary are marked as "inconsistent" in the following table:

		filter variable		
		positive	negative	omit/invalid/not appl.
dependent variable	positive	consistent	inconsistent	inconsistent
	negative	inconsistent	consistent	inconsistent
	omit	consistent	consistent	consistent
	invalid	consistent	consistent	consistent
	not appl.	consistent	consistent	consistent

4.2.6.7 CONSISTENCY CHECKS FOR LOGICAL RELATIONS BETWEEN QUESTIONS ACROSS FILES

Some questions between datafiles were logically related to each other, for example:

- deriving the number of boys and girls for each class by aggregating the student file. Then checking against ATCLSIZ, BTCLSIZ. Cases with more than 25% differences were flagged.

The following checks were undertaken between the corresponding student, teacher and school files, without any recoding:

- ACENRSB+ACENRSG<ATCLSIZ: This means that there are more students in class than there are in the whole school.
- ACENRGB+ACENRGG<ATGRSIZ: This means that there are more grade level students in class than there are in the whole school.
- ACSCRE1="1".and.ATSCHLI="2", inconsistency in data about the existence of a school library.
- ACSCRE1="2".and.ATSCHLI="1", inconsistency in data about the existence of a school library.
- ACSLBOR="1".and.ATSLBOR="2", inconsistency in data about the possibility of borrowing books from school library.
- ACSLBOR="2".and.ATSLBOR="1", inconsistency in data about the possibility of borrowing books from school library.
- ACNFTTM=0.and.ATSEX="1", inconsistency in data about the existence of male teachers in school.

- ACNFTTF=0.and.ATSEX="2", inconsistency in data about the existence of female teachers in school.
- ACINSTH<<(.or.>>)ATINSTH, inconsistency in data about instructional time per week.
- ATOTLAN=0.and.ASUSLAN$"45", inconsistency in data about the existence of other language students.
- ASBORBO$"345".and.ACSCRE1="1".and.ACAVRE1="1", inconsistency in data about the possibility of borrowing books from a public or school library.
- ASHWKF="1".and.ATASHWK="2", inconsistency in data about having/giving reading homework.
- ASHWKF$"234".and.ATASHWK="1", inconsistency in data about having/giving reading homework.
- ASTEXTF$"234".and.ATTEXTBO="1", inconsistency in data about the availability of textbooks in language lessons.
- BCENRSB+BCENRSG<BTCLSIZ: This means that there are more students in class than there are in the whole school.
- BCSCRE1="1".and.BTSCHLI="2", inconsistency in data about the existence of a school library.
- BCSCRE1="2".and.BTSCHLI="1", inconsistency in data about the existence of a school library.
- BCSCRE1="1".and.BTENCO3="2", inconsistency in data about the possibility of borrowing books from school library.
- BCSLBOR="1".and.BTENCO3="2", inconsistency in data about the possibility of borrowing books from school library.
- BCSLBOR="1".and.BTSLBOR="2", inconsistency in data about the possibility of borrowing books from school library.
- BCSLBOR="2".and.BTSLBOR="1", inconsistency in data about the possibility of borrowing books from school library.
- BCNFTTM=0.and.BTSEX="1", inconsistency in data about the existence of male teachers in school.
- BCNFTTF=0.and.BTSEX="2", inconsistency in data about the existence of female teachers in school.
- BCINSTH<<(.or.>>)BTINSTH, inconsistency in data about instructional time per week.

- BTOTLAN=0.and.BSUSLAN$"45", inconsistency in data about the existence of other language students.
- BSBORBO$"345".and.BCSCRE1="1".and.BCAVRE1="1", inconsistency in data about the possibility of borrowing books from a public or school library.

4.2.7 SPECIFIC DATA VERIFICATION RULES APPLIED TO RESOLVE INCONSISTENCIES

The data verification rules described below were applied to the inconsistencies listed above.

4.2.7.1 DATA VERIFICATION RULES FOR INCONSISTENCIES BETWEEN SPLIT VARIABLES

4.2.7.1.1 *Pop A, student file:*

AS/16 In this question the student was asked to check three of the 11 given ways of becoming a good reader that he/she thought were the three most important. If the student checked more than three a recoding was performed. Three variables from the checked ones were chosen randomly and kept as checked whereas the others were recoded to not checked (ASWAYn := 1). If there were variables that were coded as checked (code:2) then no missing or not applicable was allowed for the remaining variables. If there were variables coded "2" then the not "2"-coded variables were recoded to not checked (ASWAYn := 1).

4.2.7.1.2 *Pop A, teacher file:*

AT/22 In this question the teacher was meant to rank the 5 most important aims that he/she had. If each rank occurred only once and the answers provided were less or equal 5, then no recoding was undertaken; if the answers provided were more than 5 and/or the ranks were not unique, then the following data verification was performed. All options were sorted by their response value and a random number. Then the first "n" options (where "n" equals the minimum of the number of given options/number of requested options) were selected. The option selected first had the highest rank, whereas the option selected last had the lowest rank.

4.2.7.1.3 *Pop A, school file:*

AC/21 In this question the school principals were asked to check the two most serious problems of the 6 given problems. The performed recoding procedure for these variables was the same as for AS/16.

AC/22 In this question the school principals were asked to rank the 8 given answers from "1" for the most important to "8" for least important. The performed recoding procedure for these variables was the same as for AT/22.

4.2.7.1.4 *Pop B, student file:*

BS/24 In this question the student was asked to check three of the 11 given ways of becoming a good reader that he/she thought were the three most important. The performed recoding procedure for these variables was the same as for AS/16.

4.2.7.1.5 *Pop B, teacher file:*

BT/14 In this question the teachers was asked to rank the 7 given answers from "1" for the most frequent to "7" for the least frequent he/she did. The performed recoding procedure for these variables was the same as for AT/22.

BT/16 In this question the teacher had to answer if he/she agreed with a given answer. If one of the 5 given answers was checked then missing codes should not occur for the remaining variables. If one of the variables was coded "2", then all codes "8" and "9" in the remaining variables were recoded to not agreed (BTENCOn := 1).

BT/18 In this question the teacher should have ranked the 5 most important aims he/she had. Each rank had to occur only once. The performed recoding procedure for these variables was the same as for AT/22.

4.2.7.1.6 *Pop B, school file:*

BC/21 In this question the school principals were asked to check the two most serious problems of the 6 given problems. The performed recoding procedure for these variables was the same as for AS/16.

BC/22 In this question the school principals were asked to rank the 8 given answers from "1" for the most important to "8" for least important. The performed recoding procedure for these variables was the same as for AT/22.

4.2.7.2 DATA VERIFICATION RULES APPLIED TO WITHIN- AND BETWEEN QUESTION DEPENDENCIES (FILTER QUESTIONS)

The possible inconsistencies between a filter and its dependent questions were dealt with in different ways:

- there was only one dependent variable: Inconsistent dependent variables were then recoded to "invalid"(.I);
- there was more than one dependent variable: (a) if all dependent variables were positive, then the filter variable was recoded to positive; (b) if all dependent variables were negative, then the filter variable was recoded to negative; (c) if the dependent variables were inconsistent with themselves, then priority was given to the filter variable;
- eventually, after all this data verification of filter-dependent relations, all dependent variables were recoded to "logical not applicable"(.B) where the corresponding filter was set (negative).

4.2.7.2.1 *Pop A, student file:*

AS/19 If the student answered that he/she did not get reading homework in the filter question (ASHWKF = 1) but he/she answered the time that he/she spent on doing them in question AS/20 (ASHWKT > 1) and he/she answered that he/she was asked questions about the reading homework in

AS/21 (ASQUEST = 2,3,4 or 5) and he/she got help by doing them (ASHWKH = 2,3 or 4) then the filter question was set to "invalid" (ASHWKT := .I)

AS/19 If the student answered that he/she got reading homework in the filter question (ASHWKF = 2,3 or 4) or he/she did not answer this question (ASHWKF < .Z) but he/she answered that he/she did not get reading homework in AS/21 (ASQUEST = 1) and in AS/22 (ASHWKH = 1) then the filter question was set to "one" (ASHWKT := 1)

AS/19 If - after all of these recodings - the filter question said that the student did not get reading homework (ASHWKF = 1) then all dependent variables were set to "logical not applicable" (ASHWKT := .B; ASQUEST := .B; ASHWKH := .B).

AS/37 If the student answered that he/she read all of the mentioned materials at home (he/she checked all dependent variables (ASMAT2 = 2, ASMAT3 = 2,..., ASMAT8 = 2)) but he/she also answered that he/she read nothing at home (ASMAT1 = 2) or this variable was not validly coded (ASMAT1 < .Z) then the filter question was set to not checked (ASMAT1 = 1).

AS/34 If the student answered that he/she did not read aloud at home (ASALOUD = 2) or he/she did not answer this question at all (ASALOUD < .Z) but he/she answered all dependent questions consistently in a positive way (ASALOUF = 2,3 or 4), (ASALOUW = 2,3 or 4) and (ASMAT1 = 1) then the filter variable was set to "Yes" (ASALOUD := 2).

AS/34 If the student answered that he/she read at home (ASALOUD = 1) or he/she did not answer this question at all (ASALOUD < .Z) but he/she answered all dependent questions consistently in a negative way (ASALOUF = 1), (ASALOUW = 1) and (ASMAT1 = 2) then the filter variable was set to "No" (ASALOUD := 1).

AS/34 If - after all of these recodings - the filter question said that the student did not read at home (ASALOUD = 1) then all dependent variables were set to "logical not applicable" (ASALOUF := .B, ASALOUW := .B, ASMAT1 := .B, ASMAT2 := .B,..., ASMAT8 := .B).

AS/37 If - after all of these recodings - the filter question said that the student read nothing at home (ASMAT1 = 2) then all dependent variables were set to "logical not applicable" (ASMAT2 := .B, ASMAT3 := .B,..., ASMAT8 := .B).

4.2.7.2.2 *Pop A, teacher file:*

AT/24 If the teacher answered that he/she did not divide students into groups or that he/she did not answer the question at all (ATGROUP <= 1) but he/she answered both dependent variables in a positive way (ATGRTYP > 1) and (ATGRNUM > 1) then the filter variable was recoded to "Yes" (ATGROUP := 2).

AT/24 If the teacher answered that he/she divided students into groups or he/she did not answer the question at all (ATGROUP # 1) but he/she answered both dependent variables in a negative way (ATGRTYP = 1) and

(ATGRNUM < 2) then the filter variable was recoded to "No" (ATGROUP := 1).

AT/24 If - after all of these recodings - the filter question said that the teacher does not divide students into groups (ATGROUP = 1) then both dependent variables were set to "logical not applicable" (ATGRTYP := .B) and (ATGRNUM := .B).

AT/33 If the teacher answered that he/she did not assign homework in reading to the class tested or he/she did not answer the question at all (ATASHWK <= 1) but he/she answered both dependent variables in a positive way (ATREHWK > 1) and (ATMIHWK > 1) then the filter variable was recoded to "Yes" (ATASHWK := 2).

AT/33 If the teacher answered that he/she assigned homework in reading to the class tested (ATASHWK = 2) but he/she answered in the dependent variable that this occurs never (ATREHWK = 1) then the dependent variable was set to "invalid" (ATREHWK := .I).

AT/33 If - after all of these recodings - the filter question said that the teacher did not assign homework in reading to the class tested (ATASHWK = 1) then both dependent variables were set to "logical not applicable" (ATREHWK := .B) and (ATMIHWK := .B).

AT/36 If the teacher answered that the class did not have classroom library or he/she did not answer the question at all (ATCLLIB <= 1) but he/she answered both dependent variables in a positive way (ATBODIT >= 1) and (ATCLBOR > 1) (ATDIMAG did not influence this recoding because it is common to have libraries without magazines) then the filter variable was recoded to "Yes" (ATCLLIB := 2).

AT/36 If - after this recoding - the filter question said that there was no classroom library (ATCLLIB = 1) then all dependent variables were set to "logical not applicable" (ATBODIT := .B), (ATDIMAG := .B) and (ATCLBOR := .B).

AT/40 If the teacher answered that there was no school library or he/she did not answer the question at all (ATSCHLI <= 1) but he/she answered both dependent variables in a positive way (ATVISIT >= 1) and (ATSLBOR = 2) then the filter variable was recoded to "Yes" (ATSCHLI := 2).

AT/40 If - after this recoding - the filter question said that there was no school library (ATSCHLI = 1) then both dependent variables were set to "logical not applicable" (ATVISIT := .B) and (ATSLBOR := .B).

AT/45 If the teacher answered that there were no staff meetings at his/her school (ATSTAMF = 1) but he/she answered all dependent questions in a positive way (ATSTAM1 > .Z, ATSTAM2 > .Z,..., ATSTAM6 > .Z) and not all were coded "4" (sum(ATSTAM1,2..., 6) < 24) then the filter variable was set to "invalid" (ATSTAMF := .I).

AT/45 If - after this check - the answer of the teacher was that there are no staff meetings at his/her school (ATSTAMF = 1) and not all of the dependent

variables were coded "not in any staff meetings" (ATSTAM1 # 4, ATSTAM2 # 4,..., ATSTAM6 # 4) then all dependent variable were set to "logical not applicable" (ATSTAM1 := .B,ATSTAM2 := .B,..., ATSTAM6 := .B).

4.2.7.2.3 *Pop A, school file:*

AC/8a If the school principal answered something different than "Yes" to the question if there was a school library as a school resource (ACSCRE1 <=1) but he/she answered the number of books that it contained in a positive way (ACSLIBC > 0) and even the number of books that were added in the last year (ACSLIBA > 0) then it was assumed that there was a school library (ACSCRE1 := 2).

AC/8a If - after this check - the filter question indicated that there was no school library (ACSCRE1 = 1) then the dependent variables was set to "logical not applicable" (ACSLIBC := .B) and (ACSLIBA := .B).

AC/18 If the school principal did not answer that there were special programs in the school (ACSPPR1 # 1) but two or more of the programs were checked (sum(ACSPPR2,3,...,5) > 5) then the filter variable was recoded correspondingly (ACSPPR1 := 1).

AC/18 If the school principal did not answer that there were no special programs in the school (ACSPPR1 # 2) but none of the programs were checked (sum(ACSPPR2,3,...,5) = 4) then the filter variable was recoded correspondingly (ACSPPR1 := 2).

AC/18 If the answer of the school principal - after these recodings - was that there were no special programs in the school (ACSPPR1 = 2), then all dependent variables were recoded to "logical not applicable" (ACSPPR2 := .B, ACSPPR3 := .B,..., ACSPPR5 := .B).

AC/21 If the school principal answered that there were no serious problems in the school or if he or she did not answer this question (ACPROB6 # 1) but two or more of the problems were checked (sum(ACPROB1,2,..., 5) > 6) then the filter variable was recoded correspondingly (ACPROB6 := 1).

AC/21 If the school principal did not answer that there were no serious problems in the school (ACPROB6 # 2) but none of the problems were checked (sum(ACPROB1,2,..., 5) = 5) then the filter variable was recoded correspondingly (ACPROB6 := 2).

AC/21 If the answer of the school principal - after these recodings - was that there were serious problems in the school (ACPROB6 = 2) then the dependent variables were recoded to "logical not applicable" (ACPROB1 := .B, ACPROB2 := .B,..., ACPROB5 := .B).

4.2.7.2.4 *Pop B, student file:*

BS/7 If the student answered that he/she had a job or regular family responsibilities (BSJOB = 2) but he/she answered that he/she had no job or family responsibilities when he/she was asked how much time he/she

spent on doing it (BSTIJOB = 1) then the dependent variable was recoded to "invalid" (BSTIJOB := .I).

BS/7 If the student did not answer if he/she has a job or regular family responsibilities (BSJOB < .Z) but he/she answered the amount of time he spent on doing it (BSTIJOB > 1) then the filter variable was recoded to "Yes" (BSTJOB := 2).

BS/7 If the student did not answer if he/she had a job or regular family responsibilities (BSJOB < .Z) but he/she answered that he/she had no job or family responsibilities when he/she was asked how much time he/she spent on doing it (BSTIJOB = 1) then the filter variable was recoded to "No" (BSJOB := 1).

BS/7 If the student answered that he/she had no job or regular family responsibilities (BSJOB = 1) but he/she answered the amount of time spending on a job or family responsibilities (BSTIJOB > 1) then the dependent variable was recoded to "logical not applicable" (BSTIJOB := .B).

BS/16 If the student answered that he never got homework (BSHWKF = 1) then the amount of time that he/she spent on doing it was recoded to "zero" (BSHWKT = 1).

BS/18 If the student answered that he/she did not get homework in the language of the test or if he/she did not answer this question (BSHWKTF <= 1) but he/she answered the amount of time spent on it (BSHWKTT > 2) and he/she answered that he/she was asked questions about it in the class (BSHWKQ = 2,3,4 or 5) then the filter variable was recoded to "invalid" (BSHWKTF := .I).

BS/18 If the student did not answer that he/she did not get homework in the language of the test (BSHWKTF # 1) but he/she answered that he/she did not get homework in the language of the test when he/she was asked about the amount of time spending on it (BSHWKTT = 1) and he/she answered the same when he/she was asked if he/she had been asked questions about them in the class (BSHWKQ = 1) then the filter variable was recoded to "never" (BSHWKTF := 1).

BS/18 If - after these two checks - the students answer was that he/she never got homework in the language of the test (BSHWKTF = 1) then both dependent variables were set to "logical not applicable" (BSHWKTT := .B) and (BSHWKQ := .B).

4.2.7.2.5 *Pop B, teacher file:*

BT/22 If the teacher answered that there was no school library or he/she did not answer the question at all (BTSCHLI <= 1) but he/she answered both dependent variables in a positive way (BTVISIT >= 1) and (BTSLBOR = 2) then the filter variable was recoded to "Yes" (BTSCHLI := 2).

BT/22 If - after this recoding - the filter question said that there was no school library (BTSCHLI = 1) then both dependent variables were set to "logical not applicable" (BTVISIT := .B) and (BTSLBOR := .B).

BT/22 If the teacher answered that there was no school library (BTSCHLI = 1) but he/she answered that he/she encouraged students to borrow books from the school library (BTENCO3 = 2) then the dependent variable was recoded to "invalid" (BTENCO3 := .I).

BT/27 If the teacher answered that there were no staff meetings at his/her school (BTSTAMF = 1) but he/she answered all dependent questions in a positive way (BTSTAM1 > .Z, BTSTAM2 > .Z,..., BTSTAM6 > .Z) then the filter variable was set to "invalid" (BTSTAMF := .I).

BT/27 If - after this check - the answer of the teacher was that there were no staff meetings at his/her school (BTSTAMF = 1) then all dependent variable were set to "logical not applicable" (BTSTAM1 := .B, BTSTAM2 := .B,..., BTSTAM6 := .B).

4.2.7.2.6 *Pop B, school file:*

BC/8a If the school principal answered something different than "Yes" to the question if there was a school library as a school resource (BCSCRE1 <=1) but he/she answered the number of books that it contained in a positive way (BCSLIBC > 0) and even the number of books that were added in the last year (BCSLIBA > 0) then it was assumed that there was a school library (BCSCRE1 := 2).

BC/8a If - after this check - the filter question indicated that there was no school library (BCSCRE1 = 1) then the dependent variables was set to "logical not applicable" (BCSLIBC := .B) and (BCSLIBA := .B).

BC/18 If the school principal did not answer that there were special programs in the school (BCSPPR1 # 1) but two or more of the programs were checked (sum(BCSPPR2,3,...,5) > 5) then the filter variable was recoded correspondingly (BCSPPR1 := 1).

BC/18 If the school principal did not answer that there were no special programs in the school (BCSPPR1 # 2) but none of the programs were checked (sum(BCSPPR2,3,...,5) = 4) then the filter variable was recoded correspondingly (BCSPPR1 := 2).

BC/18 If the answer of the school principal - after these recodings - was that there were no special programs in the school (BCSPPR1 = 2), then all dependent variables were recoded to "logical not applicable" (BCSPPR2 := .B, BCSPPR3 := .B,..., BCSPPR5 := .B).

BC/21 If the school principal answered that there were no serious problems in the school or if he or she did not answer this question (BCPROB6 # 1) but two or more of the problems were checked (sum(BCPROB1,2,..., 5) > 6) then the filter variable was recoded correspondingly (BCPROB6 := 1).

BC/21 If the school principal did not answer that there were no serious problems in the school (BCPROB6 # 2) but none of the problems were checked (sum(BCPROB1,2,..., 5) = 5) then the filter variable was recoded correspondingly (BCPROB6 := 2).

BC/21 If the answer of the school principal - after these recodings - was that there were serious problems in the school (BCPROB6 = 2) then the dependent variables were recoded to "logical not applicable" (BCPROB1 := .B, BCPROB2 := .B,..., BCPROB5 := .B).

4.2.7.3 DATA VERIFICATION RULES FOR BETWEEN VARIABLE INCONSISTENCIES

4.2.7.3.1 *Pop A, student file:*

AS/1 If there was a valid response for the age in years of the student (ASAGEY > .Z) but no valid response for the age in months (ASAGEM < .Z) then the months were imputed by the mean (ASAGEM := 5.5).

AS/1 If there was a valid response for the months (ASAGEM > .Z) but no valid response for the years (ASAGEY = .) then the number of months were set to invalid (ASAGEM := .I).

4.2.7.3.2 *Pop A, teacher file:*

AT/3 (ATSCHED:Years of school education) No recoding was undertaken in the case of missing values.

AT/4 (ATTETRA:Years of pre-service teacher-training) No recoding was undertaken in the case of missing values.

AT/5 (ATSECED:Years of post-secondary education) No recoding was undertaken in the case of missing values.

AT/12' If there was a valid response for the total number of students (ATCLSIZ > .Z) but no valid response for the total number of grade level being tested students (ATGRSIZ < .Z) then the number of students in the grade was imputed by the total number of students (ATGRSIZ := ATCLSIZ).

AT/12 If there was no valid response for the total number of students (ATCLSIZ < .Z) but there was a valid response for the total number of grade level being tested students (ATGRSIZ > .Z) then the total number of students was imputed by the number of students in the grade level (ATCLSIZ := ATGRSIZ).

AT/12 If there was a value for the total number of students less than the total number of grade level being tested students (ATGRSIZ > ATCLSIZ) then the number of students in the grade was imputed by the total number of students (ATGRSIZ := ATCLSIZ).

AT/13 (ATOTLAN:Number of other language students) No recoding was undertaken in the case of missing values.

AT/15 (ATNHELP:Number of students who need remedial help in reading) No recoding was undertaken in the case of missing values.

AT/16 (ATRHELP:Number of students who receive remedial help) No recoding was undertaken in the case of missing values.

AT/15/16 If the number of students receiving remedial help was larger than the number of students that needed remedial help (ATRHELP > ATNHELP) and the number of students needing remedial help was valid (ATNHELP >

.Z) then the number of students receiving remedial help was set to the number of students needing remedial help (ATRHELP := ATNHELP).

AT/17 If the number of hours of total instructional time in a typical week had a valid value (ATINSTH > .Z) but the number of minutes was missing (ATINSTM = .) then the number of minutes was set to zero (ATINSTM := 0).

AT/17 If there was a valid response for the minutes (ATINSTM > .Z) but no valid response for the hours (ATINSTH < .Z) then the number of minutes was set to invalid (ATINSTM := .I).

AT/18 If the number of hours of teaching and learning the test language had a valid value (ATTLANH > .Z) but the number of minutes was missing (ATTLANM = .) then the number of minutes was set to zero (ATTLANM := 0).

AT/18 If there was a valid response for the minutes (ATTLANM > .Z) but the response for the hours was omitted (ATTLANH = .) then the number of hours was set to zero (ATTLANH := 0).

AT/19 If the number of hours of teaching and practice of reading the test language had a valid value (ATTEREH > .Z) but the number of minutes was omitted (ATTEREM = .) then the number of minutes was set to zero (ATTEREM := 0).

AT/19 If there was a valid response for the minutes (ATTEREM > .Z) but no valid response for the hours (ATTEREH < .Z) then the number of minutes was set to invalid (ATTEREM := .I).

4.2.7.3.3 *Pop A, school file:*

AC/1 If the time of being a principal in the total career was smaller than the time of being a principal in the present school (ACYEARP > ACYEART) and the time of being a principal in the total career was valid (ACYEART > .Z) then the number of years as a principal in the present school was recoded to the number of years as a principal in the total career (ACYEARP := ACYEART).

AC/2 If there was a valid value for the total number of enrolled full-time students - boys (ACENRSB > .Z) but the number of girls was omitted (ACENRSG = .) then the number of girls was recoded to zero (ACENRSG := 0).

AC/2 If there was a valid value for the number of girls (ACENRSG > .Z) but the number of boys was omitted (ACENRSB = .) then the number of boys was recoded to zero (ACENRSB := 0).

AC/2 If both values were zero (number of girls (ACENRSG = 0) and number of boys (ACENRSB = 0)) then both values were set to invalid (ACENRSB := .I , ACENRSG := .I).

AC/3 If there was a valid value for the total enrolled full-time grade being tested students - boys (ACENRGB > .Z) but the number of girls was omitted

(ACENRGG = .) then the number of girls was recoded to zero (ACENRGG := 0).

AC/3 If there was a valid value for the number of girls (ACENRGG > .Z) but the number of boys was omitted (ACENRGB = .) then the number of boys was recoded to zero (ACENRGB := 0).

AC/3 If both values were zero (number of girls (ACENRGG = 0) and number of boys (ACENRGB = 0)) then both values were set to invalid (ACENRGB := .I , ACENRGG := .I).

AC/12 If there was a valid value for the number of full-time teaching teachers - male (ACNFTTM > .Z) but the number of female teachers was omitted (ACNFTTF = .) then the number of female teachers was recoded to zero (ACNFTTF := 0).

AC/12 If there was a valid value for the number of female teachers (ACNFTTF > .Z) but the number of male teachers was omitted (ACNFTTM = .) then the number of male teachers was recoded to zero (ACNFTTM := 0).

AC/13 If there was a valid value for the number of full-time special teachers - male (ACNFSTM > .Z) but the number of female special teachers was omitted (ACNFSTF = .) then the number of female special teachers was recoded to zero (ACNFSTF := 0).

AC/13' If there was a valid value for the number of female special teachers (ACNFSTF > .Z) but the number of male special teachers was omitted (ACNFSTM = .) then the number of male special teachers was recoded to zero (ACNFSTM := 0).

AC/14 If the number of hours of total instructional time in a typical week for all subjects had a valid value (ACINSTH > .Z) but the number of minutes was omitted (ACINSTM = .) then the number of minutes was set to zero (ACINSTM := 0).

AC/14 If there was a valid response for the minutes (ACINSTM > .Z) but no valid response for the hours (ACINSTH < .Z) then the number of minutes was set to invalid (ACINSTM := .I).

AC/16 (ACDAYLO:Days of instruction lost in the last school year) No recoding was undertaken in the case of missing values.

4.2.7.3.4 *Pop B, student file:*

BS/1 If there was a valid response for the age in years of the student (BSAGEY > .Z) but no valid response for the age in months (BSAGEM < .Z) then the number of months was imputed by the mean (BSAGEM := 5.5).

BS/1 If there was a valid response for the months (BSAGEM > .Z) but the number of years was omitted (BSAGEY = .) then the number of months was set to invalid (BSAGEM := .I).

BS/25 If there was a valid response for the hours of silent reading in class (BSSIREH > .Z) but the response for the minutes was omitted (BSSIREM = .) then the minutes were recoded to zero (BSSIREM := 0).

BS/25 If there was a valid response for the minutes (BSSIREM > .Z) but the number of hours was omitted (BSSIREH = .) then the number of hours was set to zero (BSSIREH := 0).

4.2.7.3.5 *Pop B, teacher file:*

BT/3 (BTSCHED:Years of school education) No recoding was undertaken in the case of missing values.

BT/4 (BTTETRA:Years of pre-service teacher-training) No recoding was undertaken in the case of missing values.

BT/5 (BTSECED:Years of post-secondary education) No recoding was undertaken in the case of missing values.

BT/13 (BTOTLAN:Number of other language students) No recoding was undertaken in the case of missing values.

BT/15 (BTNHELP:Number of students who need remedial help in reading) No recoding was undertaken in the case of missing values.

BT/16 (BTRHELP:Number of students who receive remedial help) No recoding was undertaken in the case of missing values.

BT/15/16 If the number of students receiving remedial help was larger than the number of students that needed remedial help (BTRHELP > BTNHELP) and the number of students needing remedial help was valid (BTNHELP > .Z) then the number of students receiving remedial help was set to the number of students needing remedial help (BTRHELP := BTNHELP).

BT/17 If the number of hours of total instructional time in a typical week had a valid value (BTINSTH > .Z) but the number of minutes was omitted (BTINSTM = .) then the number of minutes was set to zero (BTINSTM := 0).

BT/17 If there was a valid response for the minutes (BTINSTM > .Z) but no valid response for the hours (BTINSTH < .Z) then the number of minutes was set to invalid (BTINSTM := .I).

BT/18 If the number of hours of teaching and learning the test language had a valid value (BTTLANH > .Z) but the number of minutes was omitted (BTTLANM = .) then the number of minutes was set to zero (BTTLANM := 0).

BT/18 If there was a valid response for the minutes (BTTLANM > .Z) but the number of hours was omitted (BTTLANH = .) then the number of hours was set to zero (BTTLANH := 0).

4.2.7.3.6 *Pop B, school file:*

BC/1 If the time of being a principal in the total career was smaller than the time of being a principal in the total career was valid (BCYEART > .Z) then

the number of years as a principal in the present school was recoded to the number of years as a principal in the total career (BCYEARP := BCYEART).

BC/2 If there was a valid value for the total number of enrolled full-time students - boys (BCENRSB > .Z) but the number of girls was omitted (BCENRSG = .) then the number of girls was recoded to zero (BCENRSG := 0).

BC/2 If there was a valid value for the number of girls (BCENRSG > .Z) but the number of boys was omitted (BCENRSB = .) then the number of boys was recoded to zero (BCENRSB := 0).

BC/2 If both values were zero (number of girls (BCENRSG = 0) and number of boys (BCENRSB = 0)) then both values were set to invalid (BCENRSB := .I , BCENRSG := .I).

BC/3 If there was a valid value for the total enrolled full-time grade being tested students - boys (BCENRGB > .Z) but the number of girls was omitted (BCENRGG = .) then the number of girls was recoded to zero (BCENRGG := 0).

BC/3 If there was a valid value for the number of girls (BCENRGG > .Z) but the number of boys was omitted (BCENRGB = .) then the number of boys was recoded to zero (BCENRGB := 0).

BC/3 If both values were zero (number of girls (BCENRGG = 0) and number of boys (BCENRGB = 0)) then both values were set to invalid (BCENRGB := .I , BCENRGG := .I).

BC/12 If there was a valid value for the number of full-time teaching teachers - male (BCNFTTM > .Z) but the number of female teachers was omitted (BCNFTTF = .) then the number of female teachers was recoded to zero (BCNFTTF := 0).

BC/12 If there was a valid value for the number of female teachers (BCNFTTF > .Z) but the number of male teachers was omitted (BCNFTTM = .) then the number of male teachers was recoded to zero (BCNFTTM := 0).

BC/13 If there was a valid value for the number of full-time special teachers - male (BCNFSTM > .Z) but the number of female special teachers was omitted (BCNFSTF = .) then the number of female special teachers was recoded to zero (BCNFSTF := 0).

BC/13 If there was a valid value for the number of female special teachers (BCNFSTF > .Z) but the number of male special teachers was omitted (BCNFSTM = .) then the number of male special teachers was recoded to zero (BCNFSTM := 0).

BC/14 If the number of hours of total instructional time in a typical week for all subjects had a valid value (BCINSTH > .Z) but the number of minutes was omitted (BCINSTM = .) then the number of minutes was set to zero (BCINSTM := 0).

BC/14 If there was a valid response for the minutes (BCINSTM > .Z) but no valid response for the hours (BCINSTH < .Z) then the number of minutes was set to invalid (BCINSTM := .I).

BC/16 (BCDAYLO:Days of instruction lost in the last school year) No recoding was undertaken.

Chapter 7

Development of International Constructs

Ingvar Lundberg and Dirk Hastedt

Introduction

One unique aspect of the IEA Reading Literacy Study was the development and testing of a formal model for the development of reading literacy (See Lundberg and Munck in Elley, 1994, p. 210). The model grew out of previous research, experience, and discussions with the NRCs. The model relates community, home, school factors to school conditions, teacher characteristics, and teacher instructional behaviors to reading outcomes. The constructs presented above were composites of various questionnaire items and test questions. The questionnaire items and test questions were regarded as indicators of the various constructs. In this chapter, we set forth the way in which the various constructs were developed from the questionnaire items and test questions.

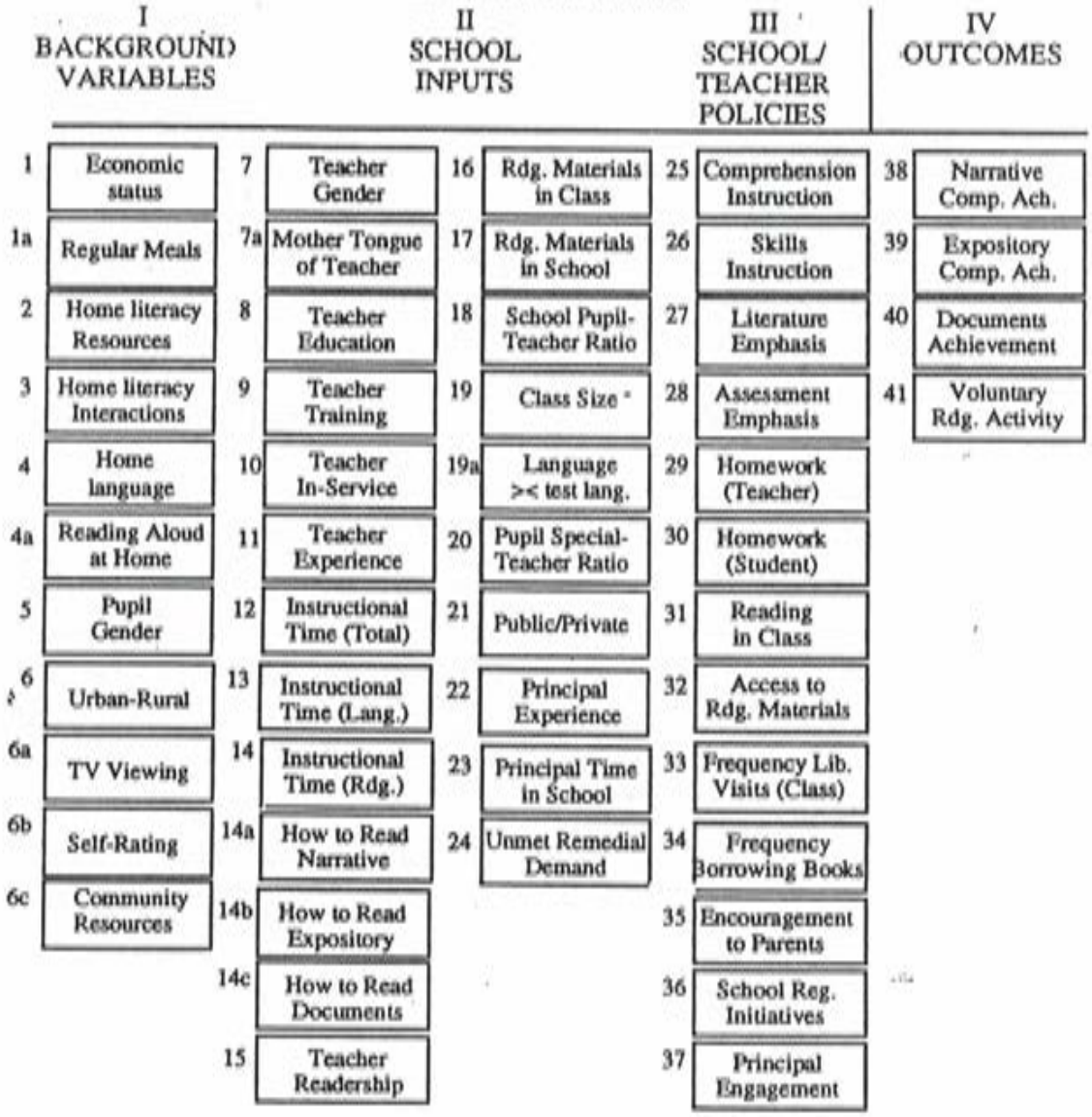
Development of International Constructs

The term construct is used here to refer to a measure of something that is not directly observable but is literally constructed by us to summarize or account for the regularity or relationships that we find in the questionnaire responses. By establishing a limited number of constructs we also reduce the enormous data set (some 600 variables) to a more manageable size.

The establishment of international constructs for Population A was based on a pooled set of data from 27 countries. Data from schools and teachers were complete, whereas student data was a random selection of 400 individuals from each country. (The computing capacity at the ICC could not meet the demands of a total and completely merged data set.)

The construct work was thoroughly discussed at the NRC meeting in Denmark. Progress reports have also been sent out after the meeting. The general conceptual framework which guided the construct work is presented in Figure 1.

Figure 1. Overall theoretical framework of the reading literacy indicators: Population A



Groups of variables which seemed to be conceptually related were subjected to principal components analyses (the international data set). The obtained factors were checked separately for each country. Construct reliabilities were also estimated. In most cases the constructs also seemed to be applicable to individual countries, which means that they explain about the same amount of variance, are about equally reliable, and have similar factor loadings.

In each country, some information was undoubtedly lost by using the internationally defined constructs. Several variables which might have been useful in one country were dropped on the international level. That is the price to be paid for an international analysis. Within a given country, a different set of constructs might actually "explain" more variance or handle the variables in a better way. So, for the national analyses to be undertaken by each NRC for his or her data set, each national data set should be processed on the basis of its own merits and conditions.

The international perspective meant that the full range of variation between countries was utilized, which implied increased chances of detecting potent determinants of differences in achievement between countries. For example, within one country, the variance in a specific instructional practice might be extremely small (all teachers in the country do the same) whereas in an international perspective, however, the variation might be considerable.

The internationally defined constructs also facilitate descriptive work. With a limited set of constructs, where each construct has the same meaning across countries, comparative analyses will be simpler than if all 600 variables had to be considered.

After having established a set of reliable constructs, the work with the international multivariate modeling could take place. The general assumption underlying the model is that aspects of the context within which schooling takes place (characteristics of schools, classrooms and teachers) influence the teaching that occurs which, in turn, influences the learning process and student outcomes. The model serves several purposes. The discussions that were held in developing the model helped to clarify the nature and complexity of the study. The model also provided a framework for examining, presenting, and interpreting the data collected. Finally, the model provided a starting point for multivariate analyses which focused on the relationships among the constructs and associated variables.

To sum up: International constructs were needed. However, a price was paid for them. A number of variables which might be useful in several countries were discarded. The common pool is still big enough to yield a reasonable level of reliability.

NRCs were recommended to rename those constructs which they modify (by adding or deleting variables) for their national analyses. This avoids the

confusion that could arise when national results are compared with results produced for the international analysis. The following tables contain the international constructs for Population A. We also established six constructs for reading activities for Population B (which are attached).

For each construct, we have specified the set of variables which define the construct, and for each variable the factor loading is given. The loading indicates the relative weight of a given variable.

The PCA procedure

The type of factor analysis used in this study is the principal components analysis (PCA). It was originally defined by Pearson in 1901. The PCA computes the Eigenvalues and Eigenvectors of the correlation matrix (or, if specified, the covariance matrix) of the construct defining variables. The number of factors (or principal components) will normally be taken as the number of Eigenvalues greater than one. This ensures that the number of factors is adequate in relation to the variance that can be explained. However, each factor is the linear combination of the original variables with coefficients equal to the calculated Eigenvector. The factors are sorted by descending order of the belonging Eigenvalue (equivalent to the variance of the component).
The first j factors are a least-square solution of the equation: $Y = X*B + E$ where Y is an **nxp** matrix of the centered observed variables; X is the **nxj** matrix of scores on the first j principal components; B is the **jxp** matrix of Eigenvectors; E is an **nxp** matrix of residuals. The task is to minimize the determinant of $(E'E)$.

The often mentioned percentage of explained variance can easily be calculated by dividing the Eigenvalue belonging to a factor by the number of defining variables.

Geometrical clarification of the PCA procedure

If there are n variables to be analyzed, then you should imagine an n-dimensional space were each axis represents one variable and each observation is represented by one dot in this space. The PCA draws a line in this space which minimizes the sum of distances between the line and the dots. This line represents the first factor of the construct. The minimization of the distances is equivalent with the maximization of the variance accounted for by this factor. For each dimension, the coefficients that leads to this line are the loadings of the variables on the factor.
The second factor is represented by a line that is orthogonal to the first line and that maximizes the remaining variance. The two factors can be rotated in this space to increase the sum of the explained variances. This can be done with a rotation matrix. The results are two purified factors.
The factor scores are calculated by rescaling the loadings, so that they sum up to unity and multiplying these rescaled loadings with the values of the corresponding variables.

Alternatives

One alternative to the PCA would have been the Maximum-Likelihood (ML) factor analysis which gives in most cases better estimates. However, an ML factor analysis needs up to a hundred times more computational time, so this was not judged to be a reasonable alternative.

An alternative for using the FACTOR procedure of SAS™ to calculate the principal components would have been the PRINCOMP procedure, but only the FACTOR procedure normalizes the scoring coefficients and allows rotation of the factors.

The steps that have been undertaken for establishing constructs in this study:
The producing of constructs has been performed in several steps.

(1) The first step was to define constructs that make sense logically. This was done to prevent the producing of constructs out of variables that correlate highly but that do not fit together conceptually.

(2) The second step was an explorative factor analysis. For this purpose internationally sampled datasets were produced where every participating country was weighted equally. For the variables in a predefined construct a correlation matrix with Pearson correlation coefficients was produced with SAS™ PROC CORR procedure. Then a factor analysis was performed with the SAS™ PROC FACTOR METHOD = PRINCIPAL procedure where the number of factors was given by the number of

Eigenvalues of the correlation matrix about one (MINEIGEN = 1 option). For clarifying the loading pattern a rotation of the factors was performed with the ROTATE = VARIMAX option.

(3) While inspecting the results of this analysis the first modifications of the predefined constructs were made. Some variables showed no variance and had therefore to be dropped. Some variables did not correlate with any other variables in the construct and had therefore to be dropped. Also the number of factors that would be specified in the next run was established by looking at the Eigenvalues: After ordering the Eigenvalues from the highest to the lowest a drastic drop could be found from one of the values to its neighbor. The number of the factors was then defined by the number of values that were higher than this drop point.

(4) Then the factor analysis was rerun. The step (2) was performed again without losing sense of the conceptual basis of the construct. Then step (3) was next, and so on.

For some constructs this loop had to be performed several times. Sometimes it lead to a point where no sensible construct was left - then one of the earlier decisions was obviously wrong and had therefore to be revised. Then the new settings were used as the starting point for a new loop.

Some of the predefined constructs were purified by this procedure, whereas some of the constructs were split up into two or more factors, which were then separated and rerun in another factor analysis later on.

(5) Then several checks on the quality of the factors were undertaken:

(a) The percentage of explained variance was calculated by dividing the explained variance by the number of variables in a factor. The reliability of the factors was estimated by:

$$p = \frac{n}{n-1} * \frac{\lambda - 1}{\lambda}$$

where n is the number of variables and

λ is the first Eigenvalue of the correlation matrix

This reliability estimate is equivalent to the Crombach coefficient [a].
Factors with a small percentage of explained variance (below about 40%) or a small reliability (below about 0.6) have not been accepted. (The numbers in brackets should not be taken without regarding the number of variables in a construct. The more variables there are in a construct the smaller is normally the percentage of explained variance. This also explains that there is an upper boundary for the number of variables that can be taken in a construct at about 10. This natural boundary is only exceeded for the construct BSFSFF.)

(b) The next check was to run a factor analysis with the internationally defined constructs for each country. Here it is important that the above mentioned limits for the percentage of explained variance and the reliability are valid for the majority of countries. There should not be too many small or negative loadings of variables on a factor.

(c) Another challenge for the constructs is that the meaning of a construct in a country does not differ too much from the international meaning. For checking this challenge a measure of internationality was defined in the following way:

$$D = \frac{\sqrt{\sum_{i=1}^{n} (l_i - \lambda_i)^2}}{n}$$

where n is the number of defining variables of the construct,

l_i is the international loading of the i-th variable, and

λ_i is the loading of the i-th variable in the examined country.

The result of this computing task showed that the national meaning of the constructs does not differ very much from the international meaning. Values from 0.02 up to 0.07 could be found for most of the constructs in all countries. Only a few values reached 0.1, but this also seemed to be acceptable.

The last set of tables that will be presented here are tables that include some statistics of the construct scores. For each construct there is a table including one row for each country with the number of valid observations, the mean and the standard deviation of the scores.

Pop A

Teaching Instructional Strategies

Comprehension Instruction

TQ	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACT15	Dramatize stories	.62	.32	.48	.43	.37	.33	.44	.48	.58	.64	.46	.45	.42	.55	.56	.50
ACT17	Orally summarizing rdg.	.61	.59	.66	.52	.48	.46	.55	.71	.56	.46	.55	.45	.55	.47	.70	.65
ACT18	Relating experience	.70	.63	.64	.58	.68	.63	.63	.74	.52	.60	.69	.55	.71	.65	.67	.70
ACT20	Making predictions	.63	.60	.65	.53	.45	.56	.51	.40	.52	.62	.56	.66	.59	.69	.60	.57
ACT21	Diagramming content	.64	.57	.60	.60	.13	.37	.58	.35	.50	.64	.32	.52	.57	.68	.30	.65
ACT22	Looking for theme	.77	.75	.79	.67	.76	.69	.65	.51	.60	.63	.66	.59	.74	.61	.73	.67
ACT23	Generalization & inference	.78	.73	.70	.67	.77	.79	.75	.67	.72	.64	.72	.77	.68	.63	.72	.60
ACT24	Study style & structure	.73	.57	.72	.66	.56	.74	.62	.62	.73	.69	.67	.68	.43	.59	.57	.68
ACT25	Compare pictures & story	.67	.52	.77	.56	.62	.64	.61	.58	.43	.45	.60	.69	.53	.73	.54	.72
ACT26	Student discussion	.64	.50	.65	.49	.43	.54	.47	.44	.43	.59	.63	.56	.52	.54	.48	.67
	Variance accounted for	.46	.35	.45	.33	.31	.35	.34	.32	.32	.36	.36	.36	.34	.38	.36	.41
	Reliability estimate	.87	.79	.87	.78	.75	.79	.79	.76	.77	.80	.80	.80	.78	.82	.80	.84

TQ	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACT15	Dramatize stories	.24	.52	.43	.27	.54	.13	.50	.39	.24	.48	.38	.54
ACT17	Orally summarizing rdg.	.63	.60	.66	.44	.62	.48	.61	.58	.53	.70	.74	.75
ACT18	Relating experience	.66	.70	.62	.51	.70	.58	.66	.60	.63	.61	.70	.72
ACT20	Making predictions	.37	.58	.54	.60	.59	.61	.64	.66	.50	.70	.75	.72
ACT21	Diagramming content	.52	.56	.54	.47	.68	.62	.64	.57	.62	.55	.54	.63
ACT22	Looking for theme	.79	.74	.74	.65	.71	.53	.69	.62	.60	.69	.74	.75
ACT23	Generalization & inference	.79	.72	.79	.68	.67	.77	.73	.69	.76	.70	.73	.72
ACT24	Study style & structure	.67	.75	.58	.70	.68	.55	.66	.67	.50	.57	.66	.72
ACT25	Compare pictures & story	.58	.66	.52	.69	.73	.41	.67	.59	.64	.64	.63	.48
ACT26	Student discussion	.36	.70	.54	.58	.69	.50	.51	.56	.59	.59	.55	.63
	Variance accounted for	.35	.43	.36	.33	.44	.29	.40	.36	.33	.39	.42	.45
	Reliability estimate	.79	.85	.81	.77	.86	.73	.83	.80	.78	.83	.85	.87

Assessment of low order skills

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ARE1	Word Recognition	.73	.69	.74	.63	.57	.66	.73	.77	.78	.77	.77	.77	.61	.78	.52	.81
ARE2	Vocabulary	.75	.62	.86	.52	.74	.78	.53	.78	.81	.65	.77	.65	.67	.75	.42	.72
ARE5	Background Knowledge	.73	.73	.65	.43	.60	.55	.76	.67	.68	.45	.64	.61	.73	.60	.68	.41
ARE6	Sentence Understanding	.75	.59	.71	.41	.70	.78	.71	.34	.56	.32	.74	.72	.68	.60	.65	.71
ARE8	Reading Study Skills	.71	.73	.69	.49	.66	.61	.53	.34	.56	.24	.63	.66	.60	.42	.50	.63
ARE9	Amount of Reading	.60	.69	.55	.52	.43	.45	.53	.73	.71	.69	.51	.67	.43	.59	.70	.49
ARE10	Decoding	.65	.68	.73	.65	.69	.55	.65	.49	.65	.38	.58	.65	.57	.55	.65	.59
	Variance Accounted for	.50	.46	.50	.28	.40	.40	.41	.38	.47	.27	.45	.46	.38	.39	.36	.40
	Reliability estimate	.83	.80	.84	.57	.75	.75	.76	.73	.81	.55	.79	.80	.73	.74	.70	.75

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ARE1	Word Recognition	.36	.70	.72	.56	.72	.63	.67	.84	.77	.39	.72	.71
ARE2	Vocabulary	.74	.75	.69	.45	.69	.68	.72	.85	.73	.22	.64	.73
ARE5	Background Knowledge	.77	.73	.64	.57	.67	.54	.66	.82	.60	.65	.39	.67
ARE6	Sentence Understanding	.72	.76	.65	.49	.71	.45	.66	.86	.73	.29	.63	.74
ARE8	Reading Study Skills	.74	.73	.69	.75	.71	.62	.60	.82	.75	.75	.77	.65
ARE9	Amount of Reading	.08	.64	.60	.74	.55	.35	.53	.74	.59	.78	.45	.66
ARE10	Decoding	.69	.74	.64	.59	.60	.66	.53	.60	.74	.72	.81	.59
	Variance Accounted for	.40	.52	.44	.36	.45	.33	.39	.63	.49	.34	.42	.46
	Reliability estimate	.75	.85	.78	.71	.79	.66	.74	.90	.83	.68	.77	.81

High demands and structure

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
VIE5	Reading Should be Assessed	.67	.59	.53	.51	.64	.75	.54	.53	.67	.55	.69	.67	.40	.61	.77	.51
VIE14	Correct Mistakes	.67	.52	.66	.49	.62	.66	.67	.70	.70	.33	.41	.68	.74	.66	.44	.55
VIE15	Feedback	.66	.42	.63	.67	.70	.73	.44	.38	.55	.50	.75	.77	.61	.73	.21	.63
VIE18	Enhance Vocabulary	.72	.77	.68	.61	.65	.38	.64	.77	.66	.63	.38	.52	.63	.56	.80	.61
VIE19	Structure & Vocabulary	.68	.74	.80	.55	.58	.68	.59	.45	.63	.71	.46	.42	.60	.70	.37	.69
	Variance Accounted for	.46	.39	.44	.32	.41	.43	.34	.34	.41	.32	.31	.39	.37	.43	.32	.36
	Reliability estimate	.71	.61	.69	.48	.64	.67	.51	.51	.65	.46	.45	.61	.57	.66	.47	.56

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
VIE5	Reading Should be Assessed	.80	.65	.62	.53	.41	.57	.70	.61	.61	.68	.68	.19
VIE14	Correct Mistakes	.46	.58	.50	.59	.45	.57	.48	.71	.69	.46	.65	.60
VIE15	Feedback	.52	.56	.47	.63	.60	.49	.59	.72	.58	.64	.65	.71
VIE18	Enhance Vocabulary	.05	.74	.71	.59	.68	.75	.56	.61	.33	.63	.67	.75
VIE19	Structure & Vocabulary	.56	.59	.64	.64	.74	.56	.60	.63	.48	.62	.76	.68
	Variance Accounted for	.29	.39	.36	.36	.35	.35	.35	.43	.30	.37	.47	.38
	Reliability estimate	.38	.61	.55	.55	.54	.54	.54	.67	.43	.57	.71	.60

Encouragement to read

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
STR11	Encourage Children to Read	.87	.90	.84	.82	.89	.88	.86	.86	.89	.78	.89	.87	.86	.81	.88	.74
STR12	Encourage Library Use	.87	.90	.84	.82	.89	.88	.86	.86	.89	.78	.89	.87	.86	.81	.88	.74
	Variance Accounted for	.76	.81	.71	.67	.80	.78	.73	.74	.79	.61	.81	.76	.74	.66	.77	.54
	Reliability estimate	.68	.77	.59	.51	.75	.72	.64	.66	.73	.35	.76	.68	.66	.49	.71	.15

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
STR11	Encourage Children to Read	.90	.86	.89	.80	.87	.90	.82	.88	.85	.78	.85	.80
STR12	Encourage Library Use	.90	.86	.89	.80	.87	.90	.82	.88	.85	.78	.85	.80
	Variance Accounted for	.81	.73	.80	.65	.75	.81	.67	.77	.73	.61	.72	.64
	Reliability estimate	.77	.64	.75	.45	.67	.77	.51	.71	.63	.35	.61	.44

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ASME3	Student Interest	.59	.76	.71	.73	.63	.85	.78	.06	.26	.74	.72	.24	.53	.58	.74	.71
METH5	Knowledge of Student Rdg Int.	.74	.83	.74	.76	.62	.64	.79	.56	.32	.82	.70	.77	.69	.67	.72	.60
METH7	Informal Observation	.56	.60	.56	.38	.56	.49	.64	.84	.84	.31	.74	.80	.68	.78	.53	.48
METH8	Interviews	.73	.73	.81	.72	.69	.62	.62	.78	.74	.69	.82	.80	.71	.71	.66	.54
	Variance Accounted for	.44	.54	.51	.45	.39	.44	.51	.41	.35	.45	.55	.49	.43	.47	.45	.35
	Reliability estimate	.57	.72	.68	.59	.48	.58	.68	.51	.39	.59	.73	.65	.56	.63	.58	.37

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ASME3	Student Interest	.64	.74	.64	.71	.74	.52	.79	.47	.69	.78	.78	.83
METH5	Knowledge of Student Rdg Int.	.82	.75	.70	.62	.76	.72	.69	.88	.69	.78	.77	.66
METH7	Informal Observation	.64	.66	.29	.67	.55	.58	.55	.46	.57	.47	.34	.68
METH8	Interviews	.64	.69	.73	.74	.70	.69	.62	.82	.56	.71	.80	.80
	Variance Accounted for	.48	.51	.38	.47	.48	.40	.45	.39	.40	.48	.49	.56
	Reliability estimate	.63	.68	.45	.63	.64	.49	.58	.48	.50	.65	.65	.73

General emphasis on assessment

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACT4	Comprehension Tests	.65	.53	.66	.60	.74	.25	.15	.75	.74	.54	.27	.77	.65	.48	.53	.66
METH3	Exercises in Workbooks	.61	.68	.83	.20	.61	.75	.83	.46	.33	.04	.68	.26	.57	.72	.42	.61
METH9	Tests In Workbooks	.69	.66	.83	.58	.59	.83	.83	.77	.58	.73	.76	-.07	.56	.76	.60	.58
ASME1	Multiple Choice Q's	.66	.64	.72	.63	.47	.25	.33	.30	.43	.75	.61	.59	.46	.42	.61	.60
ASME6	Open-ended Tests	.60	.46	.11	.74	.74	.43	.16	.62	.71	.68	.26	.70	.63	.70	.70	.56
	Variance Accounted for	.41	.36	.47	.34	.41	.31	.31	.37	.34	.37	.31	.30	.33	.40	.34	.37
	Reliability estimate	.64	.55	.72	.51	.63	.45	.44	.57	.51	.58	.44	.42	.50	.62	.51	.57

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACT4	Comprehension Tests	.53	.75	.66	.41	.35	.40	.52	.62	.60	.30	.47	.60
METH3	Exercises in Workbooks	.64	.84	.28	.64	.34	.54	.58	.69	.48	.70	.83	.78
METH9	Tests In Workbooks	.71	.73	.52	.78	.57	.74	.69	.56	.63	.76	.79	.63
ASME1	Multiple Choice Q's	.37	.67	.70	.27	.69	.63	.67	.71	.50	.69	.71	.56
ASME6	Open-ended Tests	.51	.70	.70	.49	.64	.56	.51	.49	.70	.55	-.03	.60
	Variance Accounted for	.32	.55	.35	.30	.29	.34	.36	.39	.35	.39	.41	.41
	Reliability estimate	.47	.79	.54	.42	.38	.52	.55	.60	.53	.61	.64	.64

Phonics teaching

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACT1	Letter-Sound Relationships	.85	.85	.92	.91	.79	.86	.85	.80	.83	.89	.85	.82	.85	.74	.82	.82
ACT2	Word Attack Skills	.80	.79	.86	.89	.61	.57	.79.	.79	.84	.83	.77	.83	.80	.76	.74	.84
ARE7	Phonics Skills	.77	.62	.69	.79	.65	.67	.77	.05	.34	.72	.68	.72	.55	.53	.78	.50
	Variance Accounted for	.65	.58	.68	.75	.47	.51	.65	.42	.50	.67	.59	.63	.55	.46	.61	.55
	Reliability estimate	.73	.63	.77	.83	.45	.51	.73	.32	.51	.75	.66	.69	.60	.42	.69	.58

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACT1	Letter-Sound Relationships	.90	.88	.85	.87	.76	.75	.84	.83	.78	.79	.88	.80
ACT2	Word Attack Skills	.91	.78	.66	.90	.72	.70	.84	.71	.72	.71	.79	.65
ARE7	Phonics Skills	.67	.83	.78	.56	.64	.60	.44	.73	.66	.66	.84	.76
	Variance Accounted for	.70	.69	.59	.62	.50	.47	.53	.58	.52	.52	.70	.55
	Reliability estimate	.78	.78	.65	.70	.50	.44	.56	.64	.54	.53	.79	.59

Teacher readership (Expository)

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
FRRE3	Reading Books on History	.79	.78	.81	.75	.78	.74	.81	.79	.81	.79	.76	.86	.82	.80	.83	.78
FRRE4	Reading Books on Art	.79	.81	.86	.78	.72	.72	.70	.76	.78	.84	.67	.84	.83	.74	.79	.83
FRRE5	Reading Books on Science	.80	.75	.84	.80	.81	.78	.75	.66	.59	.74	.79	.80	.86	.65	.81	.82
	Variance Accounted for	.63	.60	.70	.60	.60	.56	.57	.55	.54	.62	.55	.69	.70	.54	.66	.65
	Reliability estimate	.71	.67	.78	.67	.61	.61	.62	.59	.57	.70	.59	.78	.78	.58	.74	.74

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
FRRE3	Reading Books on History	.76	.81	.81	.75	.76	.17	.81	.83	.71	.80	.85	.77
FRRE4	Reading Books on Art	.86	.81	.70	.75	.84	.83	.84	.76	.75	.86	.85	.70
FRRE5	Reading Books on Science	.85	.72	.79	.77	.76	.83	.73	.83	.77	.80	.84	.76
	Variance Accounted for	.68	.61	.59	.57	.62	.47	.63	.65	.55	.67	.72	.56
	Reliability estimate	.76	.69	.65	.62	.69	.44	.71	.73	.60	.76	.80	.61

Teacher readership (Literature)

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
FRRE6	Reading Novels	.64	.50	.63	.75	.61	.64	.64	.68	.67	.66	.60	.49	.66	.70	.61	.59
FRRE7	Reading Poems	.78	.76	.81	.76	.81	.84	.80	.77	.77	.66	.84	.63	.79	.84	.78	.81
FRRE8	Reading Plays	.72	.75	.72	.76	.79	.77	.75	.63	.68	.76	.80	.69	.75	.79	.77	.65
FRRE9	Reading Children's Books	.63	.60	.64	.68	.48	.61	.57	.68	.51	.73	.69	.67	.38	.64	.66	.50
	Variance Accounted for	.48	.44	.49	.55	.47	.52	.48	.48	.45	.49	.55	.39	.44	.55	.50	.42
	Reliability estimate	.64	.58	.66	.72	.72	.69	.64	.63	.59	.66	.72	.48	.58	.73	.66	.53

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
FRRE6	Reading Novels	.64	.61	.59	.71	.63	.70	.66	.59	.61	.68	.53	.67
FRRE7	Reading Poems	.75	.83	.77	.72	.75	.70	.78	.74	.69	.81	.84	.78
FRRE8	Reading Plays	.81	.78	.72	.66	.82	.70	.75	.68	.66	.77	.75	.75
FRRE9	Reading Children's Books	.63	.63	.61	.48	.62	.69	.64	.57	.67	.64	.55	.75
	Variance Accounted for	.51	.52	.46	.42	.50	.49	.50	.42	.43	.53	.46	.54
	Reliability estimate	.68	.69	.61	.54	.67	.65	.66	.55	.56	.70	.61	.72

Teacher readership (Professional reading)

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
FRRE1	Articles on Teaching	.90	.84	.90	.90	.89	.91	.88	.83	.87	.91	.90	.87	.88	.90	.90	.86
FRRE2	Articles on Reading	.90	.84	.90	.90	.89	.91	.88	.83	.87	.91	.90	.87	.88	.90	.90	.86
	Variance Accounted for	.80	.71	.82	.81	.79	.83	.78	.69	.76	.83	.80	.75	.78	.81	.80	.74
	Reliability estimate	.75	.60	.77	.76	.68	.80	.72	.56	.69	.80	.75	.67	.72	.76	.76	.65

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
FRRE1	Articles on Teaching	.89	.88	.89	.83	.88	.85	.85	.90	.87	.89	.91	.87
FRRE2	Articles on Reading	.89	.88	.89	.83	.88	.85	.85	.90	.87	.89	.91	.87
	Variance Accounted for	.80	.78	.80	.68	.77	.72	.72	.81	.76	.80	.82	.76
	Reliability estimate	.74	.72	.75	.53	.71	.60	.60	.76	.68	.74	.79	.68

Reading materials in school

Sch Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACSLIBC	No. of Books in School Lib.	.89	.93	.87		.86	.85	.91	.92	.87	.92	.86	.87	.91	.81	.96	.86
ACSLIBA	No. of Books Added	.89	.93	.87		.86	.85	.91	.92	.87	.92	.86	.87	.91	.81	.96	.86
	Variance Accounted for	.79	.86	.76		.73	.73	.82	.84	.75	.85	.74	.75	.82	.65	.91	.73
	Reliability estimate	.74	.84	.68		.64	.62	.78	.81	.67	.83	.65	.68	.78	.47	.91	.63

Sch Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACSLIBC	No. of Books in School Lib.	.89	.85	.92	.83	.80	.86	.82	.96	.93	.87	.80	.82
ACSLIBA	No. of Books Added	.89	.85	.92	.83	.80	.86	.82	.96	.93	.87	.80	.82
	Variance Accounted for	.79	.72	.84	.69	.64	.74	.68	.92	.87	.75	.63	.68
	Reliability estimate	.74	.61	.81	.54	.44	.64	.52	.92	.85	.67	.42	.52

Note: Blank cells in a construct for a particular country indicate that there was no variance for these variables in this country. Therefore no loadings could be computed.

Home literacy interaction

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
SPRHTL	People Read to Stds in Test Lang.	.62	.64	.69	.51	.40	.41	.67	.56	.61	.41	.82	.34	.47	.76	.54	.67
SREATL	Frequency Reading at Home	.78	.76	.82	.69	.82	.84	.73	.79	.78	.72	.84	.79	.76	.77	.77	.75
SASKRE	Freq. Asked at Home on Reading	.63	.58	.53	.56	.55	.61	.60	.61	.64	.58	.43	.71	.55	.66	.59	.68
SALOUF	Frequency Reading Aloud at Home	.62	.64	.64	.52	.76	.79	.59	.64	.67	.57	.24	.75	.67	.56	.62	.62
	Variance Accounted for	.44	.44	.46	.32	.43	.47	.42	.43	.46	.34	.40	.45	.39	.48	.40	.46
	Reliability estimate	.58	.57	.61	.31	.56	.62	.55	.55	.61	.34	.51	.59	.48	.64	.50	.61

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
SPRHTL	People Read to Stds in Test Lang.	.55	.68	.51	.65	.81	.59	.52	.57	.52	.71	.45	.77
SREATL	Frequency Reading at Home	.82	.81	.79	.79	.83	.80	.75	.81	.79	.74	.95	.80
SASKRE	Freq. Asked at Home on Reading	.42	.58	.61	.52	.54	.74	.66	.61	.63	.68	.51	.65
SALOUF	Frequency Reading Aloud at Home	.80	.58	.70	.65	.38	.67	.66	.71	.68	.59	.95	.45
	Variance Accounted for	.45	.45	.44	.43	.45	.50	.43	.46	.44	.47	.57	.46
	Reliability estimate	.59	.59	.57	.56	.59	.66	.55	.61	.57	.62	.74	.61

Read aloud

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
SMAT2	Material Read Aloud Newspaper	.63	.61	.58	.64	.66	.67	.52	.66	.55	.55	.68	.55	.69	.40	.60	.59
SMAT3	Mat. Read Aloud Magazine	.61	.64	.70	.68	.64	.66	.65	.60	.48	.71	.62	.60	.62	.31	.67	.63
SMAT6	Mat. Read Aloud Comics	.57	.37	.62	.44	.52	.55	.52	.54	.47	.54	.39	.50	.63	.49	.48	.56
SMAT7	Mat. Read Aloud Letters	.57	.51	.62	.35	.53	.46	.57	.44	.50	.49	.58	.57	.56	.56	.55	.65
SMAT8	Mat. Rd.Ald. Words on TV Screen	.62	.51	.58	.55	.60	.58	.54	.54	.69	.51	.56	.61	.70	.60	.50	.67
	Variance Accounted for	.36	.29	.39	.30	.35	.34	.32	.32	.30	.32	.33	.32	.41	.23	.32	.38
	Reliability estimate	.56	.38	.60	.42	.54	.52	.46	.46	.41	.47	.50	.47	.65	.18	.46	.60

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
SMAT2	Material Read Aloud Newspaper	.62	.68	.62	.64	.63	.60	.60	.59	.64	.66	.06	.50
SMAT3	Mat. Read Aloud Magazine	.53	.65	.66	.69	.71	.66	.65	.61	.67	.62	.66	.63
SMAT6	Mat. Read Aloud Comics	.55	.45	.57	.46	.69	.56	.52	.63	.51	.61	-.79	.57
SMAT7	Mat. Read Aloud Letters	.50	.56	.54	.50	.38	.51	.64	.51	.37	.58	.05	.62
SMAT8	Mat. Rd. Ald Words on TV Screen	.60	.61	.63	.58	.44	.54	.65	.59	.51	.58	.05	.59
	Variance Accounted for	.31	.36	.38	.34	.34	.33	.38	.35	.30	.37	.21	.34
	Reliability estimate	.46	.55	.60	.50	.51	.50*	.60	.53	.43	.58	.08	.52

This construct could not be analyzed with PCA and listwise deletion because too many observations had to be dropped. Then the correlation matrix became singular and could not be factor analyzed. Therefore pairwise deletion was used to produce the correlation matrix for the construct. The above is the result of the factor analysis that followed the computing of the correlation matrix with pairwise deletion.

Community resources

Sch Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
AVRE1	Public Library	.72	.80	.78	.81	.63	.88	.80	.61	.60	.68	.77	.77	.56	.70	.74	.71
AVRE2	Bookstore	.84	.83	.75	.88	.87	.96	.83	.84	.88	.81	.72	.93	.82	.89	.84	.84
AVRE3	Secondary Level Schools	.81	.85	.76	.86	.92	.96	.88	.87	.82	.59	.58	.91	.79	.76	.68	.85
AVRE4	Higher Educ. Institutions	.70	.53	.73	.70	.62	.44	.53	.58	.74	.65	.49	.44	.81	.83	.70	.77
	Explained Variance	.59	.58	.57	.66	.59	.70	.60	.54	.58	.47	.42	.62	.57	.63	.55	.64
	Reliability estimate	.77	.76	.75	.83	.77	.86	.77	.72	.76	.63	.54	.80	.75	.81	.73	.81

Sch Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
AVRE1	Public Library		.92	.42	.75		.75	.70	.72		.70	.73	.86
AVRE2	Bookstore		.87	.84	.79		.86	.85	.76		.82	.67	.88
AVRE3	Secondary Level Schools		.88	.86	.82		.83	.83	.86		.77	.65	.83
AVRE4	Higher Educ. Institutions		.57	.64	.72		.56	.60	.61		.63	.76	.76
	Explained Variance		.68	.51	.60		.58	.56	.55		.54	.49	.70
	Reliability estimate		.84	.68	.77		.76		.73		.71	.66	.85

Note: Blank cells in a construct for a particular country indicate that there was no variance for these variables in this country. Therefore no loadings could be computed.

Voluntary reading

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
BOOKF	Books for Fun	.55	.63	.55	.59	.53	.45	.59	.55	.51	.58	.67	.54	.60	.57	.53	.66
COMIF	Comics	.63	.68	.66	.65	.63	.45	.63	.65	.56	.67	.68	.69	.62	.67	.63	.69
MAGAF	Magazines	.70	.64	.72	.74	.67	.70	.65	.70	.70	.69	.67	.75	.63	.78	.65	.61
NEWSF	Newspapers	.66	.52	.58	.62	.65	.68	.60	.63	.67	.57	.51	.67	.65	.66	.61	.56
DIRCF	Directions	.53	.56	.49	.58	.52	.58	.53	.599	.64	.57	.54	.55	.47	.63	.58	.53
	Variance Accounted for	.38	.37	.37	.41	.37	.34	.36	.39	.38	.38	.38	.42	.36	.44	.36	.38
	Reliability estimate	.59	.57	.57	.64	.57	.52	.56	.61	.60	.60	.60	.65	.55	.69	.56	.59

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
BOOKF	Books for Fun	.34	.58	.56	.69	.54	.53	.66	.51	.57	.62	.49	.65
COMIF	Comics	.56	.48	.52	.70	.70	.66	.71	.50	.67	.72	.54	.75
MAGAF	Magazines	.62	.64	.66	.70	.72	.72	.61	.66	.74	.70	.70	.76
NEWSF	Newspapers	.65	.65	.71	.56	.59	.65	.55	.69	.62	.69	.66	.68
DIRCF	Directions	.59	.71	.47	.59	.42	.58	.53	.58	.57	.58	.49	.64
	Variance Accounted for	.32	.35	.35	.42	.36	.40	.38	.35	.41	.44	.34	.49
	Reliability estimate	.46	.53	.53	.66	.56	.62	.59	.54	.63	.68	.52	.74

Principal engagement

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
TEVALU	Eval. by School Principal	.62	.57	.31	.57	.70	.70	.53	.59	.30	.52	.42	.47	.54	-.05	.68	.71
TPRIN1	Discuss Standards	.75	.77	.70	.81	.71	.59	.76	.70	.51	.76	.64	.60	.70	.69	.69	.82
TPRIN2	Ask for Results	.66	.66	.67	.63	.66	.44	.59	.59	.51	.68	.79	.74	.50	.54	.64	.84
TPRIN3	Suggest Methods	.76	.74	.68	.79	.45	.21	.72	.64	.57	.84	.76	.12	.78	.82	.73	.80
TPRIN4	Encourage Teacher Content	.56	.62	.05	.26	.53	.65	.60	.54	.65	.54	.59	.52	.32	.76	.46	.56
TPRIN5	Teacher Development	.61	.72	.42	.50	.53	.70	.63	.47	.61	.66	.60	.31	.52	.75	.64	.38
TPRIN6	Suggest Content	.72	.66	.74	.67	.38	.28	.52	.62	.58	.81	.76	.50	.63	.58	.63	.62
	Variance Accounted for	.45	.46	.32	.40	.33	.30	.39	.36	.30	.49	.44	.25	.34	.41	.41	.48
	Reliability estimate	.80	.81	.64	.75	.67	.60	.74	.70	.60	.82	.79	.50	.68	.76	.76	.65

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
TEVALU	Eval. by School Principal	.74	.58	.47	.55		.36	.75	.55	.19	.57	.35	.72
TPRIN1	Discuss Standards	.66	.56	.73	.72		.52	.77	.59	.76	.82	.68	.79
TPRIN2	Ask for Results	.48	.59	.77	.70		.59	.72	.33	.61	.77	.62	.73
TPRIN3	Suggest Methods	.74	.62	.82	.74		.69	.71	.45	.76	.84	.70	.78
TPRIN4	Encourage Teacher Content	.54	.62	.42	.71		.41	.49	.64	.63	.49	.42	.74
TPRIN5	Teacher Development	.73	.53	.42	.75		.32	.53	.55	.70	.66	.37	.70
TPRIN6	Suggest Content	.55	.58	.77	.72		.77	.76	.44	.67	.78	.72	.71
	Variance Accounted for	.41	.34	.42	.49		.30	.47	.27	.41	.51	.33	.55
	Reliability estimate	.76	.68	.77	.83		.61	.81	.55	.76	.84	.66	.86

Note: Blank cells in a construct for a particular country indicate that there was no variance for these variables in this country. Therefore no loadings could be computed.

Staff meetings

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
TSTAM1	Curriculum Content	.79	.73	.83	.76	.92	.77	.76	.84	.84	.70	.86	.68	.72	.63	.72	.88
TSTAM2	Presentation of Subject	.79	.81	.78	.80	.92	.70	.77	.83	.84	.64	.91	.78	.72	.82	.84	.87
TSTAM3	Development of Teachers	.70	.69	.44	.70	-.02	.69	.75	.23	.56	.76	.60	.26	.72	.74	.76	.13
	Variance Accounted for	.57	.55	.50	.57	.57	.52	.58	.48	.57	.49	.64	.38	.52	.54	.60	.52
	Reliability estimate	.63	.60	.50	.63	.62	.54	.63	.46	.63	.48	.72	.17	.53	.57	.67	.53

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
TSTAM1	Curriculum Content	.66	.78	.84	.74	.83	.79	.76	.77	.82	.84	.78	.79
TSTAM2	Presentation of Subject	.81	.72	.46	.76	.85	.78	.79	.55	.85	.77	.77	.82
TSTAM3	Development of Teachers	.80	.58	.83	.65	.70	.42	.59	.73	.46	.73	.76	.66
	Variance Accounted for	.57	.48	.53	.52	.63	.47	.52	.47	.54	.61	.59	.58
	Reliability estimate	.63	.47	.56	.53	.71	.43	.54	.44	.57	.69	.65	.64

Reading in class

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ASTEXTF	Textbooks	.68	.66	.67	.57	.57	.64	.68	.66	.67	.59	.61	.61	.69	.57	.67	.56
ASSTORF	Storybooks	.53	.57	.54	.49	.56	.53	.67	.66	.63	.46	.64	.60	.55	.62	.58	.62
ASWORKF	Workbooks	.67	.64	.62	.55	.65	.72	.67	.58	.59	.55	.59	.59	.62	.64	.64	.64
ASEXERF	Exercises	.71	.68	.64	.65	.56	.70	.66	.74	.70	.66	.58	.51	.67	.65	.69	.69
ASINFOF	Looking up Information	.53	.57	.47	.60	.52	.45	.60	.39	.41	.57	.53	.55	.45	.57	.47	.53
	Variance Accounted for	.39	.39	.35	.33	.33	.38	.43	.38	.37	.32	.35	.33	.36	.38	.38	.37
	Reliability estimate	.62	.61	.54	.49	.48	.60	.67	.60	.58	.48	.54	.49	.56	.58	.59	.58

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ASTEXTF	Textbooks	.61	.68	.68	.58	.69	.69	.67	.70	.62	.58	.70	.76
ASSTORF	Storybooks	.53	.63	.57	.34	.63	.63	.63	.51	.62	.66	.56	.69
ASWORKF	Workbooks	.61	.63	.69	.68	.73	.63	.64	.55	.67	.65	.66	.72
ASEXERF	Exercises	.59	.67	.72	.65	.74	.71	.69	.71	.63	.68	.69	.72
ASINFOF	Looking up Information	.36	.55	.43	.57	.41	.56	.44	.59	.51	.61	.46	.65
	Variance Accounted for	.30	.40	.39	.33	.42	.42	.38	.38	.38	.40	.39	.50
	Reliability estimate	.42	.63	.61	.50	.66	.65	.60	.59	.59	.63	.60	.75

Literature Emphasis

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	INS	IRL	ITA
ATACT05	Silent reading in library	.56	.60	.52	.57	.50	.49	.56	.48	.70	.73	.51	.58	.56	.60	.58	.51
ATACT07	Listen to students in groups	.55	.51	.60	.48	.41	.70	.53	.50	.57	.51	.54	.60	.44	.59	.46	.48
ATACT09	Discussion of books	.70	.67	.63	.68	.61	.43	.65	.54	.62	.68	.74	.68	.64	.61	.78	.69
ATACT12	Learning library skills	.65	.67	.57	.60	.56	.63	.57	.66	.62	.65	.59	.61	.51	.76	.67	.70
ATACT13	Reading plays or dramas	.57	.37	.67	.59	.45	.60	.54	.49	.47	.71	.65	.56	.32	.53	.45	.47
ATACT19	Read students writing	.58	.50	.56	.52	.56	.63	.54	.71	.37	.45	.60	.51	.41	.52	.48	.45
ATACT27	Read other subject areas	.34	.40	.39	.66	.40	.37	.31	.24	.08	.67	.30	.19	.37	.52	.42	.22
ATENST4	Encourage stds./hold discussions	.43	.38	.45	.42	.52	-.03	.63	.33	.48	.26	-.07	.21	.40	.31	.61	.45
	Variance accounted for	.31	.28	.31	.33	.26	.28	.30	.26	.27	.36	.29	.27	.22	.32	.32	.27
	Reliability estimate	.68	.63	.68	.70	.59	.63	.67	.60	.62	.75	.65	.62	.48	.70	.70	.61

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATACT05	Silent reading in library	.11	.52	.58	.64	.47	.62	.48	.59	.63	.56	.34	.64
ATACT07	Listen to students in groups	-.23	.55	.42	.52	.59	.54	.62	.52	.38	.44	.64	.65
ATACT09	Discussion of books	.69	.78	.65	.75	.57	.40	.57	.70	.75	.67	.65	.54
ATACT12	Learning library skills	.49	.64	.58	.68	.61	.66	.55	.59	.62	.70	.53	.72
ATACT13	Reading plays or dramas	.22	.58	.04	.54	.59	.63	.58	.42	.15	.55	.60	.52
ATACT19	Read students writing	.58	.38	.35	.58	.56	.56	.54	.58	.55	.62	.70	.64
ATACT27	Read other subject areas	-.09	.18	.42	.50	.38	.39	.28	.23	.21	.31	.11	.52
ATENST4	Encourage stds./hold discussions	.73	.56	.40	.56	.51	.08	.35	.51	.58	.15	.59	.37
	Variance accounted for	.21	.30	.22	.36	.29	.27	.26	.28	.27	.28	.30	.34
	Reliability estimate	.47	.67	.48	.75	.65	.61	.59	.64	.62	.64	.67	.72

Story reading aloud

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	INS	IRL	ITA
ATACT08	Listen to students in class	.77	.82	.84	.71	.85	.80	.81	.79	.82	.53	.77	.79	.89	.70	.80	.63
ATENST3	Encourage stds./read attr. stories	.54	.52	.25	.09	.53	.72	.68	.58	.49	.49	.35	.68	.77	.23	.66	.49
ATSTR09	Strategies/read aloud to children	.75	.80	.85	.45	.89	.89	.78	.78	.79	.45	.67	.72	.89	.71	.78	.73
ATVIE06	Views/read to children every day	.69	.70	.78	.64	.76	.58	.60	.64	.66	.77	.62	.56	.08	.67	.77	.56
ATVIE21	Views/read self-written texts	.40	.29	.19	.54	.02	-.27	.33	.39	-.18	.19	.08	.36	.37	-.18	.37	.13
	Variance accounted for	.42	.43	.43	.28	.47	.47	.44	.42	.40	.27	.31	.41	.46	.30	.48	.30
	Reliability estimate	.65	.67	.66	.37	.72	.72	.68	.66	.63	.33	.44	.64	.71	.43	.73	.42

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATACT08	Listen to students in class	.86	.83	.82	.76	.73	.83	.79	.81	.75	.76	.86	.85
ATENST3	Encourage stds./read attr. stories	.61	.57	.30	.36	.58	.50	.39	.46	.67	.59	.70	.39
ATSTR09	Strategies/read aloud to children	.88	.80	.78	.76	.75	.60	.77	.86	.81	.47	.84	.76
ATVIE06	Views/read to children every day	.64	.53	.81	.29	.59	.45	.70	.64	.52	.56	.69	.39
ATVIE21	Views/read self-written texts	.52	.47	.43	.49	.31	.27	.26	.28	.15	.28	.30	.05
	Variance accounted for	.51	.43	.44	.32	.37	.32	.38	.42	.39	.31	.50	.32
	Reliability estimate	.76	.67	.68	.47	.58	.46	.60	.65	.61	.44	.75	.47

Active teaching of comprehension

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRE	ITA
ATACT10	Rdg.act./systematic vocabulary	.59	.27	.19	.48	.24	.24	.39	.60	.41	.28	.46	.37	.57	.55	.54	.54
ATARE04	Rdg.aspects/literary appreciation	.57	.43	.59	.42	.32	.26	.69	.38	.58	.55	.48	.59	.46	.52	.62	.41
ATMETH2	Methods/tch.-made vocab. tests	.64	.69	.33	.61	.59	.58	.55	.39	.32	.63	.47	.55	.46	.42	.60	.50
ATMETH4	Methods/formal tests of compreh.	.53	.33	.16	.42	.37	.44	.32	.50	.34	.41	.41	.40	.41	.39	.38	.44
ATSTR01	Strategies/introduce background	.59	.55	.57	.16	.58	.54	.41	.62	.35	.33	.72	.23	.63	.53	.41	.56
ATSTR02	Strategies/ask children to describe	.61	.45	.67	.59	.64	.65	.54	.72	.46	.61	.68	.70	.72	.64	.63	.56
ATSTR07	Strategies/show how to understnd.	.62	.64	.70	.48	.67	-.24	.59	.56	.66	.56	.63	.65	.58	.62	.57	.65
ATSTR08	Strategies/compare stories	.68	.53	.77	.62	.62	.61	.63	.66	.64	.57	.67	.51	.67	.62	.52	.62
ATSTR13	Strategies/use self-prepared mats.	.50	.54	.59	.58	.36	.65	.46	.60	.28	.50	.53	.44	.44	.49	.56	.48
	Variance accounted for	.35	.26	.30	.25	.26	.25	.27	.32	.22	.26	.33	.26	.31	.29	.30	.29
	Reliability estimate	.77	.64	.71	.63	.65	.62	.67	.74	.56	.64	.74	.65	.72	.69	.70	.69

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATACT10	Rdg.act./systematic vocabulary	.39	.38	.56	.38	.51	.57	.43	.56	.55	.39	.37	.58
ATARE04	Rdg.aspects/literary appreciation	.54	.54	.43	.58	.57	.40	.56	.50	.69	.62	.55	.73
ATMETH2	Methods/tch.-made vocab. tests	.38	.42	.40	.55	.63	.54	.48	.53	.41	.44	.40	.65
ATMETH4	Methods/formal tests of compreh.	.42	.37	.32	.30	.39	.53	.37	.24	.30	.05	.29	.69
ATSTR01	Strategies/introduce background	.54	.51	.37	.64	.54	.58	.50	.44	.47	.61	.47	.57
ATSTR02	Strategies/ask children to describe	.56	.68	.64	.21	.63	.56	.64	.74	.61	.61	.54	.62
ATSTR07	Strategies/show how to understnd.	.61	.63	.66	.43	.56	.61	.56	.67	.62	.60	.64	.59
ATSTR08	Strategies/compare stories	.72	.74	.60	.64	.62	.54	.68	.66	.54	.59	.63	.43
ATSTR13	Strategies/use self-prepared mats.	.30	.45	.23	.61	.60	.62	.56	.23	.23	.58	.56	.51
	Variance accounted for	.26	.29	.24	.25	.32	.31	.29	.29	.26	.28	.26	.36
	Reliability estimate	.64	.70	.60	.63	.73	.72	.69	.69	.65	.68	.64	.78

Comprehension through graded materials

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ATVIE01	Views/expect accuracy	.59	.27	.58	.46	.51	.32	.44	.58	.48	.50	.40	.30	.65	.54	.65	.53
ATVIE04	Views/should enjoy reading	.42	.56	.23	.31	.34	-.34	.48	.33	.44	.07	.59	.48	.41	.47	.41	.33
ATVIE09	Views/must understand why rdg.	.66	.38	.65	.63	.69	.60	.36	.64	.44	.58	.42	.64	.66	.78	.66	.86
ATVIE13	Views/insuff.comprehension skills	.52	.48	.16	.37	.33	.49	.54	.13	.24	.53	.52	.06	.41	.03	.41	.37
ATVIE16	Views/no new bk until finish last	.49	.41	.54	.26	.40	.44	.58	.48	.37	.49	.33	.24	.33	-.11	.33	.28
ATVIE22	Views/understand what they read	.69	.67	.68	.77	.70	.77	.56	.66	.75	.57	.70	.80	.65	.74	.65	.81
ATVIE23	Views/choose own books	.52	.47	-.04	.53	.21	.43	.41	.35	.52	.40	.26	.17	.39	.01	.39	.14
	Variance accounted for	.32	.23	.23	.26	.24	.26	.24	.24	.24	.23	.23	.21	.27	.24	.27	.29
	Reliability estimate	.64	.44	.43	.51	.47	.52	.47	.46	.46	.43	.45	.36	.54	.47	.54	.59

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATVIE01	Views/expect accuracy	.26	.26	.57	.55	.58	.62	.36	.22	.34	.36	.50	.31
ATVIE04	Views/should enjoy reading	-.45	.21	.45	.63	.23	.49	.59	.46	.49	.42	.27	.53
ATVIE09	Views/must understand why rdg.	.58	.66	.59	.76	.65	.76	.52	.73	.74	.76	.62	.59
ATVIE13	Views/insuff.comprehension skills	.39	.45	.56	.46	.51	.42	.16	.33	.45	.60	.48	.63
ATVIE16	Views/no new bk until finish last	.67	.39	.44	.35	.63	.36	.47	.41	.51	.35	.53	.40
ATVIE22	Views/understand what they read	.67	.81	.65	.76	.79	.69	.69	.84	.72	.73	.76	.61
ATVIE23	Views/choose own books	-.32	.43	.39	.24	.29	-.10	.38	.32	.46	.37	.19	.53
	Variance accounted for	.25	.25	.28	.32	.31	.28	.23	.27	.30	.29	.26	.28
	Reliability estimate	.50	.50	.57	.65	.63	.58	.45	.54	.61	.59	.53	.56

UNITAB_B.LST

Univariates for the construct: (BSFSFF) Reading Activities I: Fiction and Faction 17:41 Wednesday, September 9, 1992 1

OBS	ID_CNTRY	N	MEAN	STD
1	80	2124	-0.32096	0.80751
2	110	3051	0.33632	1.08682
3	155	4077	-0.37999	0.85515
4	200	1113	0.26628	0.95363
5	215	3277	-0.56193	0.72019
6	275	1196	-0.41713	0.67915
7	280	2267	-0.20294	0.83042
8	290	1815	-0.43667	0.68869
9	295	4035	-0.47674	0.70291
10	305	3105	0.27034	0.98772
11	345	2889	-0.35187	0.81951
12	350	2925	0.39754	0.86270
13	355	3193	-0.61837	0.73917
14	380	2797	-0.20540	0.85967
15	390	2301	0.16193	0.97933
16	565	3259	-0.63816	0.68917
17	570	2586	-0.46547	0.77669
18	585	0	.	.
19	590	1880	-0.65369	0.71971
20	625	9677	1.13366	1.03856
21	635	2815	0.00927	0.89818
22	690	4762	0.15193	0.90155
23	715	7193	0.01712	0.92460
24	755	2990	-0.47083	0.79074
25	760	5490	-0.41596	0.73094
26	780	2525	0.40279	0.88548
27	795	1974	0.70690	1.11996
28	840	3100	-0.16576	0.99264
29	850	2128	0.20941	1.09723
30	856	2838	-0.08252	0.84405
31	870	1814	0.63694	0.92957

Univariates for the construct: (BSFSDR) Reading Activities II: Document Reading 17:41 Wednesday, September 9, 1992 2

OBS	ID_CNTRY	N	MEAN	STD
1	80	2282	-0.58259	0.80544
2	110	3834	0.59286	0.94972
3	155	4263	-0.04457	1.02287
4	200	1212	-0.11672	0.87426
5	215	3435	-0.33300	0.80202
6	275	1243	-0.09455	0.93248
7	280	2296	-0.12626	0.84372
8	290	1798	-0.41479	0.79016
9	295	3998	-0.37871	0.78017
10	305	3289	-0.16470	0.89907
11	345	2952	-0.34744	0.85031
12	350	3107	0.09200	0.79722
13	355	3434	-0.25819	1.01847
14	380	3110	0.05042	0.96922
15	390	2756	-0.06083	0.96073
16	565	3340	-0.37190	0.89909
17	570	2761	0.20394	1.02295
18	585	0	.	.
19	590	2005	-0.25396	0.93963
20	625	9693	0.55822	1.01115
21	635	3142	-0.51126	0.87022
22	690	4741	0.06214	0.93458
23	715	7754	-0.41867	0.85250
24	755	3214	-0.29858	0.91551
25	760	5891	-0.26757	0.86464
26	780	2680	0.13000	0.92548
27	795	2533	0.78707	1.07630
28	840	3154	0.01785	1.02337
29	850	2421	-0.19977	0.99724
30	856	2980	-0.07826	0.86562
31	870	2159	0.77241	0.86699

Univariates for the construct: (BSFSSH) Reading Activities III: School & Homework 17:41 Wednesday, September 9, 1992 3

OBS	ID_CNTRY	N	MEAN	STD
1	80	2231	-0.79099	0.95213
2	110	3822	0.34701	0.89595
3	155	4183	-0.05922	0.96415
4	200	1220	0.10454	0.90503
5	215	3418	-0.30445	0.97161
6	275	1226	-0.15043	0.96551
7	280	2303	-0.03542	1.00831
8	290	1807	-0.37136	0.91040
9	295	4030	-0.45414	0.89615
10	305	3539	0.29231	0.88505
11	345	2956	-0.26188	0.99864
12	350	3128	-0.30761	0.85485
13	355	3305	-0.11731	0.95729
14	380	2904	0.25640	0.98419
15	390	2884	0.57635	0.95181
16	565	3371	-0.60115	0.98751
17	570	2449	-0.37834	0.95858
18	585	0	.	.
19	590	2015	-0.12659	0.97129
20	625	9694	0.59949	0.88091
21	635	2993	-0.54068	0.90036
22	690	4693	0.14875	0.72033
23	715	7620	-0.06017	0.98830
24	755	3172	-0.10936	0.97189
25	760	5867	-0.31134	0.98160
26	780	2717	0.23337	0.83489
27	795	2326	0.56182	0.84237
28	840	3038	0.31376	0.96680
29	850	2632	-0.35871	0.90562
30	856	2954	-0.14416	0.87723
31	870	2085	0.57632	0.77557

Univariates for the construct: (BSFSRR) Reading Activities IV: Recreation
Reading 17:41 Wednesday, September 9, 1992 4

OBS	ID_CNTRY	N	MEAN	STD
1	80	2387	-0.23011	1.02111
2	110	4235	-1.08722	0.93412
3	155	4354	-0.29109	1.00697
4	200	1284	0.34791	0.91332
5	215	3571	-0.07608	0.93476
6	275	1274	0.79871	0.77009
7	280	2397	-0.13054	0.94291
8	290	1804	-0.28008	0.89396
9	295	4024	-0.17265	0.96635
10	305	3491	0.20113	0.98448
11	345	3039	-0.32704	0.95495
12	350	3170	0.35078	0.82615
13	355	3560	0.15819	0.97662
14	380	3271	0.07804	0.90385
15	390	2913	0.17891	0.98807
16	565	3401	-0.45910	1.01608
17	570	2810	-0.18730	0.92406
18	585	0	.	.
19	590	2051	0.27146	0.90683
20	625	9701	0.03526	0.94268
21	635	3154	-0.15113	0.97915
22	690	4776	0.32962	0.84897
23	715	8016	-0.08920	1.04934
24	755	3291	0.45450	0.90721
25	760	6064	-0.02101	0.98118
26	780	2726	0.05022	0.87597
27	795	2676	0.29174	0.91001
28	840	3194	-0.09208	1.00785
29	850	2695	0.21325	0.99880
30	856	3116	0.02783	0.91770
31	870	2293	-0.71308	0.95809

OBS	ID_CNTRY	N	MEAN	STD
1	80	2286	-0.30344	0.93943
2	110	3872	-0.18253	0.84381
3	155	4237	-0.30924	0.95787
4	200	1208	0.28344	1.07784
5	215	3497	-0.18395	0.96478
6	275	1219	-0.02379	0.98785
7	280	2332	-0.19281	0.97375
8	290	1801	0.09731	1.02977
9	295	4015	-0.16698	0.99766
10	305	3362	0.09628	0.99658
11	345	2994	-0.20104	0.91583
12	350	3115	0.56941	1.05293
13	355	3393	-0.06953	0.96581
14	380	3065	-0.08480	1.00138
15	390	2801	-0.07788	0.98150
16	565	3366	-0.21437	0.91098
17	570	2737	-0.15859	0.95825
18	585	0	.	.
19	590	1977	0.00363	1.00668
20	625	9703	0.42776	0.86463
21	635	3011	0.10709	0.95845
22	690	4776	0.08261	1.02529
23	715	7669	-0.24940	0.97768
24	755	3166	-0.00074	1.05511
25	760	5834	-0.09359	1.00338
26	780	2675	0.13047	0.88232
27	795	2424	0.26226	1.06906
28	840	3177	0.06027	1.03971
29	850	2448	0.41015	1.09993
30	856	3028	0.26542	1.05321
31	870	2117	0.19967	0.89523

Univariates for the construct: (BSFSNP) Reading Activities VI: News & Politics 17:41 Wednesday, September 9, 1992 6

OBS	ID_CNTRY	N	MEAN	STD
1	80	2409	-0.67690	0.85166
2	110	4351	-0.04666	0.86991
3	155	4364	-0.47863	0.85450
4	200	1278	0.40366	0.93388
5	215	3615	-0.10267	1.01628
6	275	1267	0.12505	0.95164
7	280	2414	-0.50589	0.89914
8	290	1804	-0.15044	1.01940
9	295	4024	-0.23876	0.98612
10	305	3522	0.26555	1.00807
11	345	3068	0.50695	0.95744
12	350	3200	0.19942	0.95575
13	355	3590	-0.32189	0.98116
14	380	3292	-0.16576	0.95917
15	390	2923	-0.59454	0.85645
16	565	3414	-0.39582	0.94306
17	570	2858	-0.27765	0.91496
18	585	0	.	.
19	590	2072	0.17205	0.95729
20	625	9703	0.32097	0.83231
21	635	3187	-0.07979	0.94131
22	690	4782	0.46181	0.91355
23	715	8002	-0.60990	0.85086
24	755	3333	0.10732	0.99445
25	760	6144	-0.12741	1.02821
26	780	2728	0.60780	0.79793
27	795	2757	0.27151	0.93570
28	840	3207	0.02071	0.96691
29	850	2759	0.07149	0.99712
30	856	3115	-0.30941	0.95252
31	870	2368	0.24582	0.92019

unitabc.lst

Univariates for the construct: (ACFSRM) Reading Materials in School
20:07 Wednesday, September 9, 1992 1

OBS	NAME	MEAN	N	STD
1	BEF01	-0.17499	93	0.18050
2	CAC01	0.46688	130	0.49365
3	CYP01	.	0	.
4	DEN01	0.66129	153	0.55133
5	FIN01	-0.11350	51	0.11081
6	FRA01	-0.17133	103	0.11146
7	GER01	-0.19999	105	0.06483
8	GER02	-0.22973	41	0.07117
9	GRC01	-0.23113	122	0.08984
10	HKO01	0.08740	67	0.31801
11	HUN01	0.27041	134	0.40743
12	ICE01	0.10773	149	0.21852
13	INS01	-0.08944	161	0.28121
14	IRE01	-0.10930	95	0.23058
15	ITA01	-0.17726	71	0.10022
16	NET01	-0.23369	82	0.04732
17	NEZ01	0.19621	157	0.28578
18	NOR01	-0.04246	171	0.26882
19	POR01	-0.24857	127	0.05962
20	SIN01	0.65006	205	0.59319
21	SPA01	-0.02139	270	0.24476
22	SWE01	0.04987	105	0.35868
23	SWI01	-0.11122	162	0.18260
24	TRI01	-0.19820	104	0.10912
25	USA01	0.40804	144	0.49940
26	VEN01	-0.12822	79	0.29965
27	YOU01	0.31914	137	0.35326

Univariates for the construct: (ACFSCR) Community Resources
20:07 Wednesday, September 9, 1992 2

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.03189	141	0.39644
2	CAC01	0.14597	140	0.26043
3	CYP01	0.12545	177	0.28097
4	DEN01	0.03569	158	0.30241
5	FIN01	0.02792	70	0.36675
6	FRA01	-0.13865	120	0.44368
7	GER01	-0.02731	136	0.34498
8	GER02	-0.07296	95	0.30661
9	GRC01	0.04131	165	0.29813
10	HKO01	-0.02051	113	0.32577
11	HUN01	0.10560	141	0.24938
12	ICE01	0.00070	161	0.42955
13	INS01	-0.62993	136	0.61066
14	IRE01	0.00165	116	0.33186
15	ITA01	0.08535	128	0.37936
16	NET01	.	0	.
17	NEZ01	0.14437	168	0.27952
18	NOR01	0.03754	171	0.30483
19	POR01	-0.47572	101	0.60964
20	SIN01	.	0	.
21	SPA01	-0.05114	264	0.44816
22	SWE01	0.08360	116	0.29226
23	SWI01	.	0	.
24	TRI01	-0.20040	150	0.43167
25	USA01	0.17490	161	0.18895
26	VEN01	-0.14775	104	0.56913
27	YOU01	-0.09816	139	0.35106

UNITABS.lst

Univariates for the construct: (ASFSHL) Home Literacy Interaction
23:05 Wednesday, September 9, 1992 1

OBS	_NAME_	MEAN	N	STD
1	BEF01	0.02603	1618	0.91202
2	CAC01	0.27951	1530	1.00520
3	CYP01	0.15814	1000	0.84517
4	DEN01	0.31905	2443	0.92658
5	FIN01	-0.18929	1005	0.89643
6	FRA01	-0.15108	876	0.86887
7	GER01	-0.18802	1529	0.82562
8	GER02	0.09456	1039	0.90343
9	GRC01	0.06864	1947	0.81266
10	HKO01	-0.11050	1128	0.81307
11	HUN01	0.00636	2605	0.87896
12	ICE01	0.01183	3015	0.90798
13	INS01	-0.33079	1068	0.84115
14	IRE01	0.31273	1698	0.88688
15	ITA01	0.28876	1358	0.96170
16	NET01	-0.29836	948	0.88139
17	NEZ01	0.11935	1703	0.94434
18	NOR01	0.16485	1387	0.88330
19	POR01	0.46208	1782	0.93802
20	SIN01	0.38832	2650	0.99001
21	SPA01	-0.05297	4503	0.94756
22	SWE01	-0.01580	3388	0.89703
23	SWI01	-0.17444	2371	0.85684
24	TRI01	0.68548	2293	1.03444
25	USA01	-0.33013	3809	0.62587
26	VEN01	0.41787	2200	0.98727
27	YOU01	0.27097	2541	0.92632

OBS	_NAME_	MEAN	N	STD
1	BEF01	0.17153	1646	0.90856
2	CAC01	0.18348	1582	1.09081
3	CYP01	-0.13723	1041	0.89241
4	DEN01	0.34163	2495	1.03249
5	FIN01	0.08254	1030	0.93850
6	FRA01	-0.23392	913	0.84940
7	GER01	-0.20797	1751	0.80377
8	GER02	0.18228	1143	0.97272
9	GRC01	.	0	.
10	HKO01	-0.27375	1152	0.84172
11	HUN01	-0.11498	2669	0.91261
12	ICE01	-0.06973	3106	1.01678
13	INS01	-0.32373	1162	0.69753
14	IRE01	0.09001	1738	0.96747
15	ITA01	-0.02733	1394	0.96803
16	NET01	0.10748	975	0.89173
17	NEZ01	0.13676	1737	1.03363
18	NOR01	0.03489	1464	0.99394
19	POR01	0.14620	1806	0.99577
20	SIN01	.	0	.
21	SPA01	.	0	.
22	SWE01	0.10712	3448	0.94773
23	SWI01	0.05750	2517	0.88450
24	TRI01	0.35063	2391	1.10785
25	USA01	-0.73494	4064	0.35818
26	VEN01	0.05593	2380	0.97302
27	YOU01	-0.07955	2770	0.93367

Univariates for the construct: (ASFSVR) Voluntary Reading
23:05 Wednesday, September 9, 1992 3

OBS	_NAME_	MEAN	N	STD
1	BEF01	0.04330	2535	0.98349
2	CAC01	-0.29180	2478	0.93205
3	CYP01	0.20170	1411	0.96591
4	DEN01	-0.03690	3286	0.96092
5	FIN01	0.58844	1463	0.84972
6	FRA01	0.01595	1796	0.93402
7	GER01	-0.38004	2052	0.89589
8	GER02	-0.16181	1403	0.95849
9	GRC01	0.14706	3301	0.94807
10	HKO01	-0.03424	3160	0.87068
11	HUN01	0.21420	2844	0.89779
12	ICE01	-0.12117	3752	0.97043
13	INS01	-0.17935	2974	1.04806
14	IRE01	-0.15754	2602	0.91856
15	ITA01	-0.20629	2064	0.94009
16	NET01	-0.27514	1634	0.90180
17	NEZ01	-0.19229	2896	0.88514
18	NOR01	0.06221	2308	0.91575
19	POR01	0.11020	2636	1.04910
20	SIN01	0.12646	7215	0.93224
21	SPA01	-0.20873	7958	1.01041
22	SWE01	0.17769	4146	0.93101
23	SWI01	-0.11493	3217	0.97721
24	TRI01	-0.02939	3476	1.04414
25	USA01	-0.20931	5844	0.88583
26	VEN01	0.29628	3450	1.19416
27	YOU01	0.26452	3139	0.95455

Univariates for the construct: (ASFSRC) Reading in Class
23:05 Wednesday, September 9, 1992 4

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.57994	2514	0.99334
2	CAC01	-0.36101	2440	0.95950
3	CYP01	0.58624	1411	0.68052
4	DEN01	0.10094	3262	0.78586
5	FIN01	0.32189	1460	0.72127
6	FRA01	-0.17937	1783	1.06801
7	GER01	-0.62260	2190	0.91815
8	GER02	-0.31346	1529	0.88256
9	GRC01	0.66524	3244	0.69534
10	HKO01	-0.12789	3210	0.82901
11	HUN01	0.15191	2800	0.63663
12	ICE01	-0.41516	3732	0.96791
13	INS01	-0.12645	2953	0.89505
14	IRE01	0.38362	2566	0.85628
15	ITA01	-0.02107	2050	0.99797
16	NET01	-0.64670	1648	0.83429
17	NEZ01	-0.28624	2856	1.04798
18	NOR01	-0.10578	2263	0.96479
19	POR01	0.44440	2624	0.81770
20	SIN01	0.72225	7222	0.80190
21	SPA01	0.10984	7981	1.07337
22	SWE01	-0.16069	4080	0.90498
23	SWI01	-0.57103	3209	0.90049
24	TRI01	0.29307	3472	0.89506
25	USA01	0.30021	5936	0.94102
26	VEN01	-0.01151	3443	1.14842
27	YOU01	0.14121	3140	0.87392

UNITABT

Univariates for the construct: (ATFSCI) Comprehension Instruction
11:31 Thursday, September 10, 1992 1

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.40455	133	0.30739
2	CAC01	0.22734	142	0.35412
3	CYP01	0.51150	270	0.29331
4	DEN01	-0.26573	197	0.27952
5	FIN01	-0.14027	64	0.29909
6	FRA01	-0.42532	112	0.28130
7	GER01	-0.16772	129	0.25640
8	GER02	0.01531	87	0.24238
9	GRC01	0.44370	154	0.36316
10	HKO01	-0.20666	143	0.31517
11	HUN01	0.47352	135	0.25610
12	ICE01	-0.49292	247	0.30875
13	INS01	0.05738	161	0.32166
14	IRE01	-0.17886	108	0.38656
15	ITA01	0.36003	122	0.38164
16	NET01	-0.29391	88	0.30043
17	NEZ01	0.10445	161	0.35795
18	NOR01	-0.13348	169	0.33549
19	POR01	0.33030	120	0.33423
20	SIN01	-0.01096	197	0.36723
21	SPA01	-0.11324	224	0.39133
22	SWE01	-0.18019	217	0.33339
23	SWI01	-0.20356	211	0.28144
24	TRI01	0.08657	189	0.35699
25	USA01	0.14579	281	0.35306
26	VEN01	0.16020	111	0.43028
27	YOU01	0.17455	122	0.22064

Univariates for the construct: (ATFSAL) Assessment of Low Order Skills 11:31 Thursday, September 10, 1992 2

OBS	NAME	MEAN	N	STD
1	BEF01	-0.06803	133	0.37921
2	CAC01	0.05504	138	0.36249
3	CYP01	0.19627	242	0.25073
4	DEN01	-0.04704	159	0.38349
5	FIN01	-0.20302	68	0.32626
6	FRA01	-0.18194	114	0.42591
7	GER01	-0.15297	143	0.43020
8	GER02	-0.05777	98	0.33938
9	GRC01	0.26635	164	0.20907
10	HKO01	-0.04175	150	0.34649
11	HUN01	0.07908	131	0.31243
12	ICE01	-0.10327	212	0.40137
13	INS01	0.28954	162	0.20308
14	IRE01	-0.04805	108	0.34758
15	ITA01	0.20713	134	0.27471
16	NET01	-0.49100	91	0.39621
17	NEZ01	-0.16592	154	0.41821
18	NOR01	-0.15212	149	0.40288
19	POR01	0.35196	131	0.19208
20	SIN01	-0.02484	187	0.37479
21	SPA01	0.14804	204	0.28583
22	SWE01	-0.64191	199	0.62506
23	SWI01	-0.39829	196	0.47094
24	TRI01	0.34370	201	0.17727
25	USA01	0.26760	281	0.24018
26	VEN01	0.15954	106	0.33800
27	YOU01	0.27700	126	0.16566

Univariates for the construct: (ATFSHD) High Demand and
Structure 11:31 Thursday, September 10, 1992 3

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.10414	137	0.35346
2	CAC01	-0.50944	147	0.32673
3	CYP01	0.15783	301	0.28143
4	DEN01	-0.31833	198	0.31711
5	FIN01	-0.20901	68	0.42497
6	FRA01	-0.22091	125	0.32813
7	GER01	-0.17726	145	0.31791
8	GER02	0.02286	98	0.28520
9	GRC01	0.37683	166	0.27345
10	HKO01	0.35455	146	0.20087
11	HUN01	0.50007	141	0.29463
12	ICE01	-0.05179	260	0.33619
13	INS01	0.54543	171	0.24016
14	IRE01	-0.04881	116	0.27043
15	ITA01	0.29978	144	0.30315
16	NET01	-0.07851	92	0.21637
17	NEZ01	-0.33769	171	0.30659
18	NOR01	-0.18891	174	0.30497
19	POR01	0.35143	148	0.30525
20	SIN01	-0.04542	203	0.29599
21	SPA01	0.28671	270	0.29769
22	SWE01	-0.48014	232	0.33897
23	SWI01	-0.26363	225	0.26079
24	TRI01	0.17888	209	0.31928
25	USA01	-0.26436	292	0.33668
26	VEN01	0.41335	129	0.30290
27	YOU01	-0.17227	133	0.31777

Univariates for the construct: (ATFSER) Encouragement to Read
11:31 Thursday, September 10, 1992 4

OBS	NAME	MEAN	N	STD
1	BEF01	-0.02205	146	0.45359
2	CAC01	0.19536	145	0.29534
3	CYP01	0.25354	209	0.28448
4	DEN01	-0.15535	202	0.38484
5	FIN01	-0.08037	68	0.42806
6	FRA01	-0.03039	129	0.46405
7	GER01	-0.43186	145	0.45221
8	GER02	-0.31719	97	0.44692
9	GRC01	0.12735	174	0.35711
10	HKO01	-0.19635	156	0.45660
11	HUN01	0.09610	142	0.34398
12	ICE01	-0.01847	272	0.40975
13	INS01	0.16999	170	0.32211
14	IRE01	0.09562	116	0.39309
15	ITA01	0.07472	146	0.35763
16	NET01	-0.16699	96	0.45513
17	NEZ01	0.11170	175	0.34192
18	NOR01	-0.17555	177	0.40463
19	POR01	0.12309	144	0.37044
20	SIN01	0.20602	204	0.32631
21	SPA01	0.17110	299	0.34930
22	SWE01	-0.02313	227	0.36931
23	SWI01	-0.38398	224	0.41459
24	TRI01	0.10553	208	0.36715
25	USA01	0.15534	298	0.32541
26	VEN01	0.08879	126	0.41022
27	YOU01	0.09668	136	0.36738

OBS	NAME	MEAN	N	STD
1	BEF01	-0.16106	129	0.47867
2	CAC01	0.26708	137	0.31226
3	CYP01	-0.08078	269	0.40618
4	DEN01	0.24505	190	0.33166
5	FIN01	0.23649	70	0.28696
6	FRA01	-0.24726	116	0.43568
7	GER01	-0.10349	143	0.36141
8	GER02	-0.01044	99	0.36084
9	GRC01	-0.09181	163	0.44411
10	HKO01	-0.20166	150	0.44334
11	HUN01	0.32620	138	0.20908
12	ICE01	-0.01305	253	0.44464
13	INS01	0.19611	162	0.32014
14	IRE01	-0.27427	115	0.41702
15	ITA01	0.25620	127	0.28309
16	NET01	-0.23647	93	0.45478
17	NEZ01	0.12889	167	0.34115
18	NOR01	0.02587	164	0.33886
19	POR01	0.05071	136	0.41016
20	SIN01	-0.18371	200	0.43265
21	SPA01	-0.08121	257	0.42339
22	SWE01	-0.15912	209	0.34413
23	SWI01	-0.23847	205	0.38221
24	TRI01	0.05379	176	0.40780
25	USA01	0.02122	288	0.39723
26	VEN01	0.00673	110	0.50715
27	YOU01	0.26411	119	0.26688

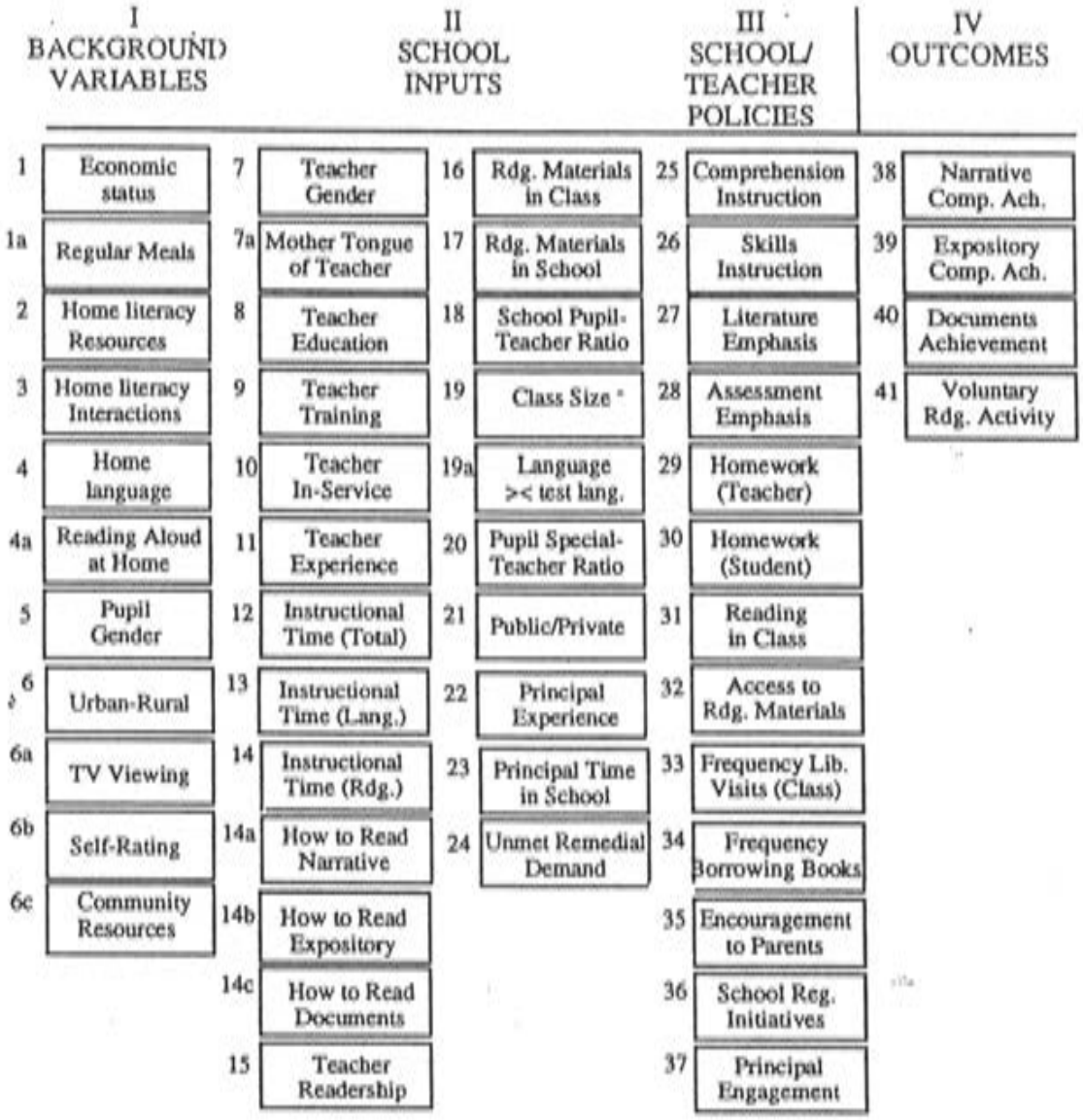
Univariates for the construct: (ATFSAE) General Emphasis on Assessment 11:31 Thursday, September 10, 1992 6

OBS	NAME	MEAN	N	STD
1	BEF01	0.04570	144	0.33936
2	CAC01	-0.62191	134	0.47515
3	CYP01	0.18819	303	0.29005
4	DEN01	-0.25327	188	0.41498
5	FIN01	-0.11279	64	0.27248
6	FRA01	0.06492	117	0.30750
7	GER01	-0.41237	139	0.36811
8	GER02	-0.25516	95	0.35531
9	GRC01	0.23171	171	0.31284
10	HKO01	-0.02240	142	0.29006
11	HUN01	0.31157	142	0.19293
12	ICE01	-0.18411	240	0.33362
13	INS01	0.20220	163	0.26762
14	IRE01	0.19341	113	0.33028
15	ITA01	0.27351	140	0.27057
16	NET01	-0.27625	93	0.34852
17	NEZ01	-0.40193	161	0.52478
18	NOR01	0.09339	155	0.33576
19	POR01	0.42081	141	0.20679
20	SIN01	0.17992	201	0.25125
21	SPA01	0.14606	271	0.33388
22	SWE01	-0.25170	213	0.44610
23	SWI01	-0.28422	208	0.36375
24	TRI01	0.24253	198	0.28951
25	USA01	0.15246	292	0.33303
26	VEN01	0.12717	116	0.38845
27	YOU01	0.19929	131	0.23517

Univariates for the construct: (ATFSPT) Phonics Teaching
11:31 Thursday, September 10, 1992 7

OBS	NAME	MEAN	N	STD
1	BEF01	-0.22622	132	0.41119
2	CAC01	0.15991	138	0.37436
3	CYP01	-0.30947	277	0.49505
4	DEN01	0.03548	174	0.31819
5	FIN01	-0.10878	68	0.29189
6	FRA01	-0.35168	113	0.44866
7	GER01	0.00283	135	0.31413
8	GER02	0.15996	89	0.27241
9	GRC01	-0.13506	156	0.47004
10	HKO01	0.21703	151	0.32568
11	HUN01	0.06794	135	0.38838
12	ICE01	-0.36255	241	0.45697
13	INS01	0.27187	165	0.19687
14	IRE01	0.23492	118	0.29918
15	ITA01	0.13298	125	0.37046
16	NET01	-0.21148	86	0.44680
17	NEZ01	0.05574	163	0.39235
18	NOR01	-0.19042	167	0.38881
19	POR01	0.13869	124	0.46011
20	SIN01	0.20059	199	0.21998
21	SPA01	0.14432	270	0.33475
22	SWE01	-0.49063	203	0.39929
23	SWI01	-0.22428	211	0.36504
24	TRI01	0.35219	210	0.21820
25	USA01	0.14426	286	0.39808
26	VEN01	0.07211	113	0.38226
27	YOU01	0.14212	124	0.25459

Figure 1. Overall theoretical framework of the reading literacy indicators: Population A



Groups of variables which seemed to be conceptually related were subjected to principal components analyses (the international data set). The obtained factors were checked separately for each country. Construct reliabilities were also estimated. In most cases the constructs also seemed to be applicable to individual countries, which means that they explain about the same amount of variance, are about equally reliable, and have similar factor loadings.

In each country, some information was undoubtedly lost by using the internationally defined constructs. Several variables which might have been useful in one country were dropped on the international level. That is the price to be paid for an international analysis. Within a given country, a different set of constructs might actually "explain" more variance or handle the variables in a better way. So, for the national analyses to be undertaken by each NRC for his or her data set, each national data set should be processed on the basis of its own merits and conditions.

The international perspective meant that the full range of variation between countries was utilized, which implied increased chances of detecting potent determinants of differences in achievement between countries. For example, within one country, the variance in a specific instructional practice might be extremely small (all teachers in the country do the same) whereas in an international perspective, however, the variation might be considerable.

The internationally defined constructs also facilitate descriptive work. With a limited set of constructs, where each construct has the same meaning across countries, comparative analyses will be simpler than if all 600 variables had to be considered.

After having established a set of reliable constructs, the work with the international multivariate modeling could take place. The general assumption underlying the model is that aspects of the context within which schooling takes place (characteristics of schools, classrooms and teachers) influence the teaching that occurs which, in turn, influences the learning process and student outcomes. The model serves several purposes. The discussions that were held in developing the model helped to clarify the nature and complexity of the study. The model also provided a framework for examining, presenting, and interpreting the data collected. Finally, the model provided a starting point for multivariate analyses which focused on the relationships among the constructs and associated variables.

To sum up: International constructs were needed. However, a price was paid for them. A number of variables which might be useful in several countries were discarded. The common pool is still big enough to yield a reasonable level of reliability.

NRCs were recommended to rename those constructs which they modify (by adding or deleting variables) for their national analyses. This avoids the

confusion that could arise when national results are compared with results produced for the international analysis. The following tables contain the international constructs for Population A. We also established six constructs for reading activities for Population B (which are attached).

For each construct, we have specified the set of variables which define the construct, and for each variable the factor loading is given. The loading indicates the relative weight of a given variable.

The PCA procedure

The type of factor analysis used in this study is the principal components analysis (PCA). It was originally defined by Pearson in 1901. The PCA computes the Eigenvalues and Eigenvectors of the correlation matrix (or, if specified, the covariance matrix) of the construct defining variables. The number of factors (or principal components) will normally be taken as the number of Eigenvalues greater than one. This ensures that the number of factors is adequate in relation to the variance that can be explained. However, each factor is the linear combination of the original variables with coefficients equal to the calculated Eigenvector. The factors are sorted by descending order of the belonging Eigenvalue (equivalent to the variance of the component).
The first j factors are a least-square solution of the equation: $Y = X*B + E$ where Y is an **nxp** matrix of the centered observed variables; X is the **nxj** matrix of scores on the first j principal components; B is the **jxp** matrix of Eigenvectors; E is an **nxp** matrix of residuals. The task is to minimize the determinant of $(E'E)$.

The often mentioned percentage of explained variance can easily be calculated by dividing the Eigenvalue belonging to a factor by the number of defining variables.

Geometrical clarification of the PCA procedure

If there are n variables to be analyzed, then you should imagine an n-dimensional space were each axis represents one variable and each observation is represented by one dot in this space. The PCA draws a line in this space which minimizes the sum of distances between the line and the dots. This line represents the first factor of the construct. The minimization of the distances is equivalent with the maximization of the variance accounted for by this factor. For each dimension, the coefficients that leads to this line are the loadings of the variables on the factor.
The second factor is represented by a line that is orthogonal to the first line and that maximizes the remaining variance. The two factors can be rotated in this space to increase the sum of the explained variances. This can be done with a rotation matrix. The results are two purified factors.
The factor scores are calculated by rescaling the loadings, so that they sum up to unity and multiplying these rescaled loadings with the values of the corresponding variables.

Alternatives

One alternative to the PCA would have been the Maximum-Likelihood (ML) factor analysis which gives in most cases better estimates. However, an ML factor analysis needs up to a hundred times more computational time, so this was not judged to be a reasonable alternative.

An alternative for using the FACTOR procedure of SAS™ to calculate the principal components would have been the PRINCOMP procedure, but only the FACTOR procedure normalizes the scoring coefficients and allows rotation of the factors.

The steps that have been undertaken for establishing constructs in this study:
The producing of constructs has been performed in several steps.

(1) The first step was to define constructs that make sense logically. This was done to prevent the producing of constructs out of variables that correlate highly but that do not fit together conceptually.

(2) The second step was an explorative factor analysis. For this purpose internationally sampled datasets were produced where every participating country was weighted equally. For the variables in a predefined construct a correlation matrix with Pearson correlation coefficients was produced with SAS™ PROC CORR procedure. Then a factor analysis was performed with the SAS™ PROC FACTOR METHOD = PRINCIPAL procedure where the number of factors was given by the number of

Eigenvalues of the correlation matrix about one (MINEIGEN = 1 option). For clarifying the loading pattern a rotation of the factors was performed with the ROTATE = VARIMAX option.

(3) While inspecting the results of this analysis the first modifications of the predefined constructs were made. Some variables showed no variance and had therefore to be dropped. Some variables did not correlate with any other variables in the construct and had therefore to be dropped. Also the number of factors that would be specified in the next run was established by looking at the Eigenvalues: After ordering the Eigenvalues from the highest to the lowest a drastic drop could be found from one of the values to its neighbor. The number of the factors was then defined by the number of values that were higher than this drop point.

(4) Then the factor analysis was rerun. The step (2) was performed again without losing sense of the conceptual basis of the construct. Then step (3) was next, and so on.

For some constructs this loop had to be performed several times. Sometimes it lead to a point where no sensible construct was left - then one of the earlier decisions was obviously wrong and had therefore to be revised. Then the new settings were used as the starting point for a new loop.

Some of the predefined constructs were purified by this procedure, whereas some of the constructs were split up into two or more factors, which were then separated and rerun in another factor analysis later on.

(5) Then several checks on the quality of the factors were undertaken:

(a) The percentage of explained variance was calculated by dividing the explained variance by the number of variables in a factor. The reliability of the factors was estimated by:

$$p = \frac{n}{n-1} * \frac{\lambda - 1}{\lambda}$$

where n is the number of variables and

λ is the first Eigenvalue of the correlation matrix

This reliability estimate is equivalent to the Crombach coefficient [a].
Factors with a small percentage of explained variance (below about 40%) or a small reliability (below about 0.6) have not been accepted. (The numbers in brackets should not be taken without regarding the number of variables in a construct. The more variables there are in a construct the smaller is normally the percentage of explained variance. This also explains that there is an upper boundary for the number of variables that can be taken in a construct at about 10. This natural boundary is only exceeded for the construct BSFSFF.)

(b) The next check was to run a factor analysis with the internationally defined constructs for each country. Here it is important that the above mentioned limits for the percentage of explained variance and the reliability are valid for the majority of countries. There should not be too many small or negative loadings of variables on a factor.

(c) Another challenge for the constructs is that the meaning of a construct in a country does not differ too much from the international meaning. For checking this challenge a measure of internationality was defined in the following way:

$$D = \frac{\sqrt{\sum_{i=1}^{n} (l_i - \lambda_i)^2}}{n}$$

where n is the number of defining variables of the construct,

l_i is the international loading of the i-th variable, and

λ_i is the loading of the i-th variable in the examined country.

The result of this computing task showed that the national meaning of the constructs does not differ very much from the international meaning. Values from 0.02 up to 0.07 could be found for most of the constructs in all countries. Only a few values reached 0.1, but this also seemed to be acceptable.

The last set of tables that will be presented here are tables that include some statistics of the construct scores. For each construct there is a table including one row for each country with the number of valid observations, the mean and the standard deviation of the scores.

Pop A

Teaching Instructional Strategies

Comprehension Instruction

TQ	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACT15	Dramatize stories	.62	.32	.48	.43	.37	.33	.44	.48	.58	.64	.46	.45	.42	.55	.56	.50
ACT17	Orally summarizing rdg.	.61	.59	.66	.52	.48	.46	.55	.71	.56	.46	.55	.45	.55	.47	.70	.65
ACT18	Relating experience	.70	.63	.64	.58	.68	.63	.63	.74	.52	.60	.69	.55	.71	.65	.67	.70
ACT20	Making predictions	.63	.60	.65	.53	.45	.56	.51	.40	.52	.62	.56	.66	.59	.69	.60	.57
ACT21	Diagramming content	.64	.57	.60	.60	.13	.37	.58	.35	.50	.64	.32	.52	.57	.68	.30	.65
ACT22	Looking for theme	.77	.75	.79	.67	.76	.69	.65	.51	.60	.63	.66	.59	.74	.61	.73	.67
ACT23	Generalization & inference	.78	.73	.70	.67	.77	.79	.75	.67	.72	.64	.72	.77	.68	.63	.72	.60
ACT24	Study style & structure	.73	.57	.72	.66	.56	.74	.62	.62	.73	.69	.67	.68	.43	.59	.57	.68
ACT25	Compare pictures & story	.67	.52	.77	.56	.62	.64	.61	.58	.43	.45	.60	.69	.53	.73	.54	.72
ACT26	Student discussion	.64	.50	.65	.49	.43	.54	.47	.44	.43	.59	.63	.56	.52	.54	.48	.67
	Variance accounted for	.46	.35	.45	.33	.31	.35	.34	.32	.32	.36	.36	.36	.34	.38	.36	.41
	Reliability estimate	.87	.79	.87	.78	.75	.79	.79	.76	.77	.80	.80	.80	.78	.82	.80	.84

TQ	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACT15	Dramatize stories	.24	.52	.43	.27	.54	.13	.50	.39	.24	.48	.38	.54
ACT17	Orally summarizing rdg.	.63	.60	.66	.44	.62	.48	.61	.58	.53	.70	.74	.75
ACT18	Relating experience	.66	.70	.62	.51	.70	.58	.66	.60	.63	.61	.70	.72
ACT20	Making predictions	.37	.58	.54	.60	.59	.61	.64	.66	.50	.70	.75	.72
ACT21	Diagramming content	.52	.56	.54	.47	.68	.62	.64	.57	.62	.55	.54	.63
ACT22	Looking for theme	.79	.74	.74	.65	.71	.53	.69	.62	.60	.69	.74	.75
ACT23	Generalization & inference	.79	.72	.79	.68	.67	.77	.73	.69	.76	.70	.73	.72
ACT24	Study style & structure	.67	.75	.58	.70	.68	.55	.66	.67	.50	.57	.66	.72
ACT25	Compare pictures & story	.58	.66	.52	.69	.73	.41	.67	.59	.64	.64	.63	.48
ACT26	Student discussion	.36	.70	.54	.58	.69	.50	.51	.56	.59	.59	.55	.63
	Variance accounted for	.35	.43	.36	.33	.44	.29	.40	.36	.33	.39	.42	.45
	Reliability estimate	.79	.85	.81	.77	.86	.73	.83	.80	.78	.83	.85	.87

Assessment of low order skills

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ARE1	Word Recognition	.73	.69	.74	.63	.57	.66	.73	.77	.78	.77	.77	.77	.61	.78	.52	.81
ARE2	Vocabulary	.75	.62	.86	.52	.74	.78	.53	.78	.81	.65	.77	.65	.67	.75	.42	.72
ARE5	Background Knowledge	.73	.73	.65	.43	.60	.55	.76	.67	.68	.45	.64	.61	.73	.60	.68	.41
ARE6	Sentence Understanding	.75	.59	.71	.41	.70	.78	.71	.34	.56	.32	.74	.72	.68	.60	.65	.71
ARE8	Reading Study Skills	.71	.73	.69	.49	.66	.61	.53	.34	.56	.24	.63	.66	.60	.42	.50	.63
ARE9	Amount of Reading	.60	.69	.55	.52	.43	.45	.53	.73	.71	.69	.51	.67	.43	.59	.70	.49
ARE10	Decoding	.65	.68	.73	.65	.69	.55	.65	.49	.65	.38	.58	.65	.57	.55	.65	.59
	Variance Accounted for	.50	.46	.50	.28	.40	.40	.41	.38	.47	.27	.45	.46	.38	.39	.36	.40
	Reliability estimate	.83	.80	.84	.57	.75	.75	.76	.73	.81	.55	.79	.80	.73	.74	.70	.75

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ARE1	Word Recognition	.36	.70	.72	.56	.72	.63	.67	.84	.77	.39	.72	.71
ARE2	Vocabulary	.74	.75	.69	.45	.69	.68	.72	.85	.73	.22	.64	.73
ARE5	Background Knowledge	.77	.73	.64	.57	.67	.54	.66	.82	.60	.65	.39	.67
ARE6	Sentence Understanding	.72	.76	.65	.49	.71	.45	.66	.86	.73	.29	.63	.74
ARE8	Reading Study Skills	.74	.73	.69	.75	.71	.62	.60	.82	.75	.75	.77	.65
ARE9	Amount of Reading	.08	.64	.60	.74	.55	.35	.53	.74	.59	.78	.45	.66
ARE10	Decoding	.69	.74	.64	.59	.60	.66	.53	.60	.74	.72	.81	.59
	Variance Accounted for	.40	.52	.44	.36	.45	.33	.39	.63	.49	.34	.42	.46
	Reliability estimate	.75	.85	.78	.71	.79	.66	.74	.90	.83	.68	.77	.81

High demands and structure

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
VIE5	Reading Should be Assessed	.67	.59	.53	.51	.64	.75	.54	.53	.67	.55	.69	.67	.40	.61	.77	.51
VIE14	Correct Mistakes	.67	.52	.66	.49	.62	.66	.67	.70	.70	.33	.41	.68	.74	.66	.44	.55
VIE15	Feedback	.66	.42	.63	.67	.70	.73	.44	.38	.55	.50	.75	.77	.61	.73	.21	.63
VIE18	Enhance Vocabulary	.72	.77	.68	.61	.65	.38	.64	.77	.66	.63	.38	.52	.63	.56	.80	.61
VIE19	Structure & Vocabulary	.68	.74	.80	.55	.58	.68	.59	.45	.63	.71	.46	.42	.60	.70	.37	.69
	Variance Accounted for	.46	.39	.44	.32	.41	.43	.34	.34	.41	.32	.31	.39	.37	.43	.32	.36
	Reliability estimate	.71	.61	.69	.48	.64	.67	.51	.51	.65	.46	.45	.61	.57	.66	.47	.56

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
VIE5	Reading Should be Assessed	.80	.65	.62	.53	.41	.57	.70	.61	.61	.68	.68	.19
VIE14	Correct Mistakes	.46	.58	.50	.59	.45	.57	.48	.71	.69	.46	.65	.60
VIE15	Feedback	.52	.56	.47	.63	.60	.49	.59	.72	.58	.64	.65	.71
VIE18	Enhance Vocabulary	.05	.74	.71	.59	.68	.75	.56	.61	.33	.63	.67	.75
VIE19	Structure & Vocabulary	.56	.59	.64	.64	.74	.56	.60	.63	.48	.62	.76	.68
	Variance Accounted for	.29	.39	.36	.36	.35	.35	.35	.43	.30	.37	.47	.38
	Reliability estimate	.38	.61	.55	.55	.54	.54	.54	.67	.43	.57	.71	.60

Encouragement to read

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
STR11	Encourage Children to Read	.87	.90	.84	.82	.89	.88	.86	.86	.89	.78	.89	.87	.86	.81	.88	.74
STR12	Encourage Library Use	.87	.90	.84	.82	.89	.88	.86	.86	.89	.78	.89	.87	.86	.81	.88	.74
	Variance Accounted for	.76	.81	.71	.67	.80	.78	.73	.74	.79	.61	.81	.76	.74	.66	.77	.54
	Reliability estimate	.68	.77	.59	.51	.75	.72	.64	.66	.73	.35	.76	.68	.66	.49	.71	.15

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
STR11	Encourage Children to Read	.90	.86	.89	.80	.87	.90	.82	.88	.85	.78	.85	.80
STR12	Encourage Library Use	.90	.86	.89	.80	.87	.90	.82	.88	.85	.78	.85	.80
	Variance Accounted for	.81	.73	.80	.65	.75	.81	.67	.77	.73	.61	.72	.64
	Reliability estimate	.77	.64	.75	.45	.67	.77	.51	.71	.63	.35	.61	.44

Taking student interest into account

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ASME3	Student Interest	.59	.76	.71	.73	.63	.85	.78	.06	.26	.74	.72	.24	.53	.58	.74	.71
METH5	Knowledge of Student Rdg Int.	.74	.83	.74	.76	.62	.64	.79	.56	.32	.82	.70	.77	.69	.67	.72	.60
METH7	Informal Observation	.56	.60	.56	.38	.56	.49	.64	.84	.84	.31	.74	.80	.68	.78	.53	.48
METH8	Interviews	.73	.73	.81	.72	.69	.62	.62	.78	.74	.69	.82	.80	.71	.71	.66	.54
	Variance Accounted for	.44	.54	.51	.45	.39	.44	.51	.41	.35	.45	.55	.49	.43	.47	.45	.35
	Reliability estimate	.57	.72	.68	.59	.48	.58	.68	.51	.39	.59	.73	.65	.56	.63	.58	.37

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ASME3	Student Interest	.64	.74	.64	.71	.74	.52	.79	.47	.69	.78	.78	.83
METH5	Knowledge of Student Rdg Int.	.82	.75	.70	.62	.76	.72	.69	.88	.69	.78	.77	.66
METH7	Informal Observation	.64	.66	.29	.67	.55	.58	.55	.46	.57	.47	.34	.68
METH8	Interviews	.64	.69	.73	.74	.70	.69	.62	.82	.56	.71	.80	.80
	Variance Accounted for	.48	.51	.38	.47	.48	.40	.45	.39	.40	.48	.49	.56
	Reliability estimate	.63	.68	.45	.63	.64	.49	.58	.48	.50	.65	.65	.73

General emphasis on assessment

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACT4	Comprehension Tests	.65	.53	.66	.60	.74	.25	.15	.75	.74	.54	.27	.77	.65	.48	.53	.66
METH3	Exercises in Workbooks	.61	.68	.83	.20	.61	.75	.83	.46	.33	.04	.68	.26	.57	.72	.42	.61
METH9	Tests In Workbooks	.69	.66	.83	.58	.59	.83	.83	.77	.58	.73	.76	-.07	.56	.76	.60	.58
ASME1	Multiple Choice Q's	.66	.64	.72	.63	.47	.25	.33	.30	.43	.75	.61	.59	.46	.42	.61	.60
ASME6	Open-ended Tests	.60	.46	.11	.74	.74	.43	.16	.62	.71	.68	.26	.70	.63	.70	.70	.56
	Variance Accounted for	.41	.36	.47	.34	.41	.31	.31	.37	.34	.37	.31	.30	.33	.40	.34	.37
	Reliability estimate	.64	.55	.72	.51	.63	.45	.44	.57	.51	.58	.44	.42	.50	.62	.51	.57

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACT4	Comprehension Tests	.53	.75	.66	.41	.35	.40	.52	.62	.60	.30	.47	.60
METH3	Exercises in Workbooks	.64	.84	.28	.64	.34	.54	.58	.69	.48	.70	.83	.78
METH9	Tests In Workbooks	.71	.73	.52	.78	.57	.74	.69	.56	.63	.76	.79	.63
ASME1	Multiple Choice Q's	.37	.67	.70	.27	.69	.63	.67	.71	.50	.69	.71	.56
ASME6	Open-ended Tests	.51	.70	.70	.49	.64	.56	.51	.49	.70	.55	-.03	.60
	Variance Accounted for	.32	.55	.35	.30	.29	.34	.36	.39	.35	.39	.41	.41
	Reliability estimate	.47	.79	.54	.42	.38	.52	.55	.60	.53	.61	.64	.64

Phonics teaching

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACT1	Letter-Sound Relationships	.85	.85	.92	.91	.79	.86	.85	.80	.83	.89	.85	.82	.85	.74	.82	.82
ACT2	Word Attack Skills	.80	.79	.86	.89	.61	.57	.79.	.79	.84	.83	.77	.83	.80	.76	.74	.84
ARE7	Phonics Skills	.77	.62	.69	.79	.65	.67	.77	.05	.34	.72	.68	.72	.55	.53	.78	.50
	Variance Accounted for	.65	.58	.68	.75	.47	.51	.65	.42	.50	.67	.59	.63	.55	.46	.61	.55
	Reliability estimate	.73	.63	.77	.83	.45	.51	.73	.32	.51	.75	.66	.69	.60	.42	.69	.58

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACT1	Letter-Sound Relationships	.90	.88	.85	.87	.76	.75	.84	.83	.78	.79	.88	.80
ACT2	Word Attack Skills	.91	.78	.66	.90	.72	.70	.84	.71	.72	.71	.79	.65
ARE7	Phonics Skills	.67	.83	.78	.56	.64	.60	.44	.73	.66	.66	.84	.76
	Variance Accounted for	.70	.69	.59	.62	.50	.47	.53	.58	.52	.52	.70	.55
	Reliability estimate	.78	.78	.65	.70	.50	.44	.56	.64	.54	.53	.79	.59

Teacher readership (Expository)

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
FRRE3	Reading Books on History	.79	.78	.81	.75	.78	.74	.81	.79	.81	.79	.76	.86	.82	.80	.83	.78
FRRE4	Reading Books on Art	.79	.81	.86	.78	.72	.72	.70	.76	.78	.84	.67	.84	.83	.74	.79	.83
FRRE5	Reading Books on Science	.80	.75	.84	.80	.81	.78	.75	.66	.59	.74	.79	.80	.86	.65	.81	.82
	Variance Accounted for	.63	.60	.70	.60	.60	.56	.57	.55	.54	.62	.55	.69	.70	.54	.66	.65
	Reliability estimate	.71	.67	.78	.67	.61	.61	.62	.59	.57	.70	.59	.78	.78	.58	.74	.74

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
FRRE3	Reading Books on History	.76	.81	.81	.75	.76	.17	.81	.83	.71	.80	.85	.77
FRRE4	Reading Books on Art	.86	.81	.70	.75	.84	.83	.84	.76	.75	.86	.85	.70
FRRE5	Reading Books on Science	.85	.72	.79	.77	.76	.83	.73	.83	.77	.80	.84	.76
	Variance Accounted for	.68	.61	.59	.57	.62	.47	.63	.65	.55	.67	.72	.56
	Reliability estimate	.76	.69	.65	.62	.69	.44	.71	.73	.60	.76	.80	.61

Teacher readership (Literature)

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
FRRE6	Reading Novels	.64	.50	.63	.75	.61	.64	.64	.68	.67	.66	.60	.49	.66	.70	.61	.59
FRRE7	Reading Poems	.78	.76	.81	.76	.81	.84	.80	.77	.77	.66	.84	.63	.79	.84	.78	.81
FRRE8	Reading Plays	.72	.75	.72	.76	.79	.77	.75	.63	.68	.76	.80	.69	.75	.79	.77	.65
FRRE9	Reading Children's Books	.63	.60	.64	.68	.48	.61	.57	.68	.51	.73	.69	.67	.38	.64	.66	.50
	Variance Accounted for	.48	.44	.49	.55	.47	.52	.48	.48	.45	.49	.55	.39	.44	.55	.50	.42
	Reliability estimate	.64	.58	.66	.72	.72	.69	.64	.63	.59	.66	.72	.48	.58	.73	.66	.53

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
FRRE6	Reading Novels	.64	.61	.59	.71	.63	.70	.66	.59	.61	.68	.53	.67
FRRE7	Reading Poems	.75	.83	.77	.72	.75	.70	.78	.74	.69	.81	.84	.78
FRRE8	Reading Plays	.81	.78	.72	.66	.82	.70	.75	.68	.66	.77	.75	.75
FRRE9	Reading Children's Books	.63	.63	.61	.48	.62	.69	.64	.57	.67	.64	.55	.75
	Variance Accounted for	.51	.52	.46	.42	.50	.49	.50	.42	.43	.53	.46	.54
	Reliability estimate	.68	.69	.61	.54	.67	.65	.66	.55	.56	.70	.61	.72

Teacher readership (Professional reading)

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
FRRE1	Articles on Teaching	.90	.84	.90	.90	.89	.91	.88	.83	.87	.91	.90	.87	.88	.90	.90	.86
FRRE2	Articles on Reading	.90	.84	.90	.90	.89	.91	.88	.83	.87	.91	.90	.87	.88	.90	.90	.86
	Variance Accounted for	.80	.71	.82	.81	.79	.83	.78	.69	.76	.83	.80	.75	.78	.81	.80	.74
	Reliability estimate	.75	.60	.77	.76	.68	.80	.72	.56	.69	.80	.75	.67	.72	.76	.76	.65

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
FRRE1	Articles on Teaching	.89	.88	.89	.83	.88	.85	.85	.90	.87	.89	.91	.87
FRRE2	Articles on Reading	.89	.88	.89	.83	.88	.85	.85	.90	.87	.89	.91	.87
	Variance Accounted for	.80	.78	.80	.68	.77	.72	.72	.81	.76	.80	.82	.76
	Reliability estimate	.74	.72	.75	.53	.71	.60	.60	.76	.68	.74	.79	.68

Reading materials in school

Sch Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ACSLIBC	No. of Books in School Lib.	.89	.93	.87		.86	.85	.91	.92	.87	.92	.86	.87	.91	.81	.96	.86
ACSLIBA	No. of Books Added	.89	.93	.87		.86	.85	.91	.92	.87	.92	.86	.87	.91	.81	.96	.86
	Variance Accounted for	.79	.86	.76		.73	.73	.82	.84	.75	.85	.74	.75	.82	.65	.91	.73
	Reliability estimate	.74	.84	.68		.64	.62	.78	.81	.67	.83	.65	.68	.78	.47	.91	.63

Sch Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ACSLIBC	No. of Books in School Lib.	.89	.85	.92	.83	.80	.86	.82	.96	.93	.87	.80	.82
ACSLIBA	No. of Books Added	.89	.85	.92	.83	.80	.86	.82	.96	.93	.87	.80	.82
	Variance Accounted for	.79	.72	.84	.69	.64	.74	.68	.92	.87	.75	.63	.68
	Reliability estimate	.74	.61	.81	.54	.44	.64	.52	.92	.85	.67	.42	.52

Note: Blank cells in a construct for a particular country indicate that there was no variance for these variables in this country. Therefore no loadings could be computed.

Home literacy interaction

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
SPRHTL	People Read to Stds in Test Lang.	.62	.64	.69	.51	.40	.41	.67	.56	.61	.41	.82	.34	.47	.76	.54	.67
SREATL	Frequency Reading at Home	.78	.76	.82	.69	.82	.84	.73	.79	.78	.72	.84	.79	.76	.77	.77	.75
SASKRE	Freq. Asked at Home on Reading	.63	.58	.53	.56	.55	.61	.60	.61	.64	.58	.43	.71	.55	.66	.59	.68
SALOUF	Frequency Reading Aloud at Home	.62	.64	.64	.52	.76	.79	.59	.64	.67	.57	.24	.75	.67	.56	.62	.62
	Variance Accounted for	.44	.44	.46	.32	.43	.47	.42	.43	.46	.34	.40	.45	.39	.48	.40	.46
	Reliability estimate	.58	.57	.61	.31	.56	.62	.55	.55	.61	.34	.51	.59	.48	.64	.50	.61

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
SPRHTL	People Read to Stds in Test Lang.	.55	.68	.51	.65	.81	.59	.52	.57	.52	.71	.45	.77
SREATL	Frequency Reading at Home	.82	.81	.79	.79	.83	.80	.75	.81	.79	.74	.95	.80
SASKRE	Freq. Asked at Home on Reading	.42	.58	.61	.52	.54	.74	.66	.61	.63	.68	.51	.65
SALOUF	Frequency Reading Aloud at Home	.80	.58	.70	.65	.38	.67	.66	.71	.68	.59	.95	.45
	Variance Accounted for	.45	.45	.44	.43	.45	.50	.43	.46	.44	.47	.57	.46
	Reliability estimate	.59	.59	.57	.56	.59	.66	.55	.61	.57	.62	.74	.61

Read aloud

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
SMAT2	Material Read Aloud Newspaper	.63	.61	.58	.64	.66	.67	.52	.66	.55	.55	.68	.55	.69	.40	.60	.59
SMAT3	Mat. Read Aloud Magazine	.61	.64	.70	.68	.64	.66	.65	.60	.48	.71	.62	.60	.62	.31	.67	.63
SMAT6	Mat. Read Aloud Comics	.57	.37	.62	.44	.52	.55	.52	.54	.47	.54	.39	.50	.63	.49	.48	.56
SMAT7	Mat. Read Aloud Letters	.57	.51	.62	.35	.53	.46	.57	.44	.50	.49	.58	.57	.56	.56	.55	.65
SMAT8	Mat. Rd.Ald. Words on TV Screen	.62	.51	.58	.55	.60	.58	.54	.54	.69	.51	.56	.61	.70	.60	.50	.67
	Variance Accounted for	.36	.29	.39	.30	.35	.34	.32	.32	.30	.32	.33	.32	.41	.23	.32	.38
	Reliability estimate	.56	.38	.60	.42	.54	.52	.46	.46	.41	.47	.50	.47	.65	.18	.46	.60

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
SMAT2	Material Read Aloud Newspaper	.62	.68	.62	.64	.63	.60	.60	.59	.64	.66	.06	.50
SMAT3	Mat. Read Aloud Magazine	.53	.65	.66	.69	.71	.66	.65	.61	.67	.62	.66	.63
SMAT6	Mat. Read Aloud Comics	.55	.45	.57	.46	.69	.56	.52	.63	.51	.61	-.79	.57
SMAT7	Mat. Read Aloud Letters	.50	.56	.54	.50	.38	.51	.64	.51	.37	.58	.05	.62
SMAT8	Mat. Rd. Ald Words on TV Screen	.60	.61	.63	.58	.44	.54	.65	.59	.51	.58	.05	.59
	Variance Accounted for	.31	.36	.38	.34	.34	.33	.38	.35	.30	.37	.21	.34
	Reliability estimate	.46	.55	.60	.50	.51	.50*	.60	.53	.43	.58	.08	.52

This construct could not be analyzed with PCA and listwise deletion because too many observations had to be dropped. Then the correlation matrix became singular and could not be factor analyzed. Therefore pairwise deletion was used to produce the correlation matrix for the construct. The above is the result of the factor analysis that followed the computing of the correlation matrix with pairwise deletion.

Community resources

Sch Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
AVRE1	Public Library	.72	.80	.78	.81	.63	.88	.80	.61	.60	.68	.77	.77	.56	.70	.74	.71
AVRE2	Bookstore	.84	.83	.75	.88	.87	.96	.83	.84	.88	.81	.72	.93	.82	.89	.84	.84
AVRE3	Secondary Level Schools	.81	.85	.76	.86	.92	.96	.88	.87	.82	.59	.58	.91	.79	.76	.68	.85
AVRE4	Higher Educ. Institutions	.70	.53	.73	.70	.62	.44	.53	.58	.74	.65	.49	.44	.81	.83	.70	.77
	Explained Variance	.59	.58	.57	.66	.59	.70	.60	.54	.58	.47	.42	.62	.57	.63	.55	.64
	Reliability estimate	.77	.76	.75	.83	.77	.86	.77	.72	.76	.63	.54	.80	.75	.81	.73	.81

Sch Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
AVRE1	Public Library		.92	.42	.75		.75	.70	.72		.70	.73	.86
AVRE2	Bookstore		.87	.84	.79		.86	.85	.76		.82	.67	.88
AVRE3	Secondary Level Schools		.88	.86	.82		.83	.83	.86		.77	.65	.83
AVRE4	Higher Educ. Institutions		.57	.64	.72		.56	.60	.61		.63	.76	.76
	Explained Variance		.68	.51	.60		.58	.56	.55		.54	.49	.70
	Reliability estimate		.84	.68	.77		.76		.73		.71	.66	.85

Note: Blank cells in a construct for a particular country indicate that there was no variance for these variables in this country. Therefore no loadings could be computed.

Voluntary reading

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
BOOKF	Books for Fun	.55	.63	.55	.59	.53	.45	.59	.55	.51	.58	.67	.54	.60	.57	.53	.66
COMIF	Comics	.63	.68	.66	.65	.63	.45	.63	.65	.56	.67	.68	.69	.62	.67	.63	.69
MAGAF	Magazines	.70	.64	.72	.74	.67	.70	.65	.70	.70	.69	.67	.75	.63	.78	.65	.61
NEWSF	Newspapers	.66	.52	.58	.62	.65	.68	.60	.63	.67	.57	.51	.67	.65	.66	.61	.56
DIRCF	Directions	.53	.56	.49	.58	.52	.58	.53	.599	.64	.57	.54	.55	.47	.63	.58	.53
	Variance Accounted for	.38	.37	.37	.41	.37	.34	.36	.39	.38	.38	.38	.42	.36	.44	.36	.38
	Reliability estimate	.59	.57	.57	.64	.57	.52	.56	.61	.60	.60	.60	.65	.55	.69	.56	.59

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
BOOKF	Books for Fun	.34	.58	.56	.69	.54	.53	.66	.51	.57	.62	.49	.65
COMIF	Comics	.56	.48	.52	.70	.70	.66	.71	.50	.67	.72	.54	.75
MAGAF	Magazines	.62	.64	.66	.70	.72	.72	.61	.66	.74	.70	.70	.76
NEWSF	Newspapers	.65	.65	.71	.56	.59	.65	.55	.69	.62	.69	.66	.68
DIRCF	Directions	.59	.71	.47	.59	.42	.58	.53	.58	.57	.58	.49	.64
	Variance Accounted for	.32	.35	.35	.42	.36	.40	.38	.35	.41	.44	.34	.49
	Reliability estimate	.46	.53	.53	.66	.56	.62	.59	.54	.63	.68	.52	.74

Principal engagement

TQ	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
TEVALU	Eval. by School Principal	.62	.57	.31	.57	.70	.70	.53	.59	.30	.52	.42	.47	.54	-.05	.68	.71
TPRIN1	Discuss Standards	.75	.77	.70	.81	.71	.59	.76	.70	.51	.76	.64	.60	.70	.69	.69	.82
TPRIN2	Ask for Results	.66	.66	.67	.63	.66	.44	.59	.59	.51	.68	.79	.74	.50	.54	.64	.84
TPRIN3	Suggest Methods	.76	.74	.68	.79	.45	.21	.72	.64	.57	.84	.76	.12	.78	.82	.73	.80
TPRIN4	Encourage Teacher Content	.56	.62	.05	.26	.53	.65	.60	.54	.65	.54	.59	.52	.32	.76	.46	.56
TPRIN5	Teacher Development	.61	.72	.42	.50	.53	.70	.63	.47	.61	.66	.60	.31	.52	.75	.64	.38
TPRIN6	Suggest Content	.72	.66	.74	.67	.38	.28	.52	.62	.58	.81	.76	.50	.63	.58	.63	.62
	Variance Accounted for	.45	.46	.32	.40	.33	.30	.39	.36	.30	.49	.44	.25	.34	.41	.41	.48
	Reliability estimate	.80	.81	.64	.75	.67	.60	.74	.70	.60	.82	.79	.50	.68	.76	.76	.65

TQ	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
TEVALU	Eval. by School Principal	.74	.58	.47	.55		.36	.75	.55	.19	.57	.35	.72
TPRIN1	Discuss Standards	.66	.56	.73	.72		.52	.77	.59	.76	.82	.68	.79
TPRIN2	Ask for Results	.48	.59	.77	.70		.59	.72	.33	.61	.77	.62	.73
TPRIN3	Suggest Methods	.74	.62	.82	.74		.69	.71	.45	.76	.84	.70	.78
TPRIN4	Encourage Teacher Content	.54	.62	.42	.71		.41	.49	.64	.63	.49	.42	.74
TPRIN5	Teacher Development	.73	.53	.42	.75		.32	.53	.55	.70	.66	.37	.70
TPRIN6	Suggest Content	.55	.58	.77	.72		.77	.76	.44	.67	.78	.72	.71
	Variance Accounted for	.41	.34	.42	.49		.30	.47	.27	.41	.51	.33	.55
	Reliability estimate	.76	.68	.77	.83		.61	.81	.55	.76	.84	.66	.86

Note: Blank cells in a construct for a particular country indicate that there was no variance for these variables in this country. Therefore no loadings could be computed.

Staff meetings

T Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
TSTAM1	Curriculum Content	.79	.73	.83	.76	.92	.77	.76	.84	.84	.70	.86	.68	.72	.63	.72	.88
TSTAM2	Presentation of Subject	.79	.81	.78	.80	.92	.70	.77	.83	.84	.64	.91	.78	.72	.82	.84	.87
TSTAM3	Development of Teachers	.70	.69	.44	.70	-.02	.69	.75	.23	.56	.76	.60	.26	.72	.74	.76	.13
	Variance Accounted for	.57	.55	.50	.57	.57	.52	.58	.48	.57	.49	.64	.38	.52	.54	.60	.52
	Reliability estimate	.63	.60	.50	.63	.62	.54	.63	.46	.63	.48	.72	.17	.53	.57	.67	.53

T Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
TSTAM1	Curriculum Content	.66	.78	.84	.74	.83	.79	.76	.77	.82	.84	.78	.79
TSTAM2	Presentation of Subject	.81	.72	.46	.76	.85	.78	.79	.55	.85	.77	.77	.82
TSTAM3	Development of Teachers	.80	.58	.83	.65	.70	.42	.59	.73	.46	.73	.76	.66
	Variance Accounted for	.57	.48	.53	.52	.63	.47	.52	.47	.54	.61	.59	.58
	Reliability estimate	.63	.47	.56	.53	.71	.43	.54	.44	.57	.69	.65	.64

Reading in class

St Q	Variable	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ASTEXTF	Textbooks	.68	.66	.67	.57	.57	.64	.68	.66	.67	.59	.61	.61	.69	.57	.67	.56
ASSTORF	Storybooks	.53	.57	.54	.49	.56	.53	.67	.66	.63	.46	.64	.60	.55	.62	.58	.62
ASWORKF	Workbooks	.67	.64	.62	.55	.65	.72	.67	.58	.59	.55	.59	.59	.62	.64	.64	.64
ASEXERF	Exercises	.71	.68	.64	.65	.56	.70	.66	.74	.70	.66	.58	.51	.67	.65	.69	.69
ASINFOF	Looking up Information	.53	.57	.47	.60	.52	.45	.60	.39	.41	.57	.53	.55	.45	.57	.47	.53
	Variance Accounted for	.39	.39	.35	.33	.33	.38	.43	.38	.37	.32	.35	.33	.36	.38	.38	.37
	Reliability estimate	.62	.61	.54	.49	.48	.60	.67	.60	.58	.48	.54	.49	.56	.58	.59	.58

St Q	Variable	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ASTEXTF	Textbooks	.61	.68	.68	.58	.69	.69	.67	.70	.62	.58	.70	.76
ASSTORF	Storybooks	.53	.63	.57	.34	.63	.63	.63	.51	.62	.66	.56	.69
ASWORKF	Workbooks	.61	.63	.69	.68	.73	.63	.64	.55	.67	.65	.66	.72
ASEXERF	Exercises	.59	.67	.72	.65	.74	.71	.69	.71	.63	.68	.69	.72
ASINFOF	Looking up Information	.36	.55	.43	.57	.41	.56	.44	.59	.51	.61	.46	.65
	Variance Accounted for	.30	.40	.39	.33	.42	.42	.38	.38	.38	.40	.39	.50
	Reliability estimate	.42	.63	.61	.50	.66	.65	.60	.59	.59	.63	.60	.75

Literature Emphasis

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	INS	IRL	ITA
ATACT05	Silent reading in library	.56	.60	.52	.57	.50	.49	.56	.48	.70	.73	.51	.58	.56	.60	.58	.51
ATACT07	Listen to students in groups	.55	.51	.60	.48	.41	.70	.53	.50	.57	.51	.54	.60	.44	.59	.46	.48
ATACT09	Discussion of books	.70	.67	.63	.68	.61	.43	.65	.54	.62	.68	.74	.68	.64	.61	.78	.69
ATACT12	Learning library skills	.65	.67	.57	.60	.56	.63	.57	.66	.62	.65[1]	.59	.61	.51	.76	.67	.70
ATACT13	Reading plays or dramas	.57	.37	.67	.59	.45	.60	.54	.49	.47	.71	.65	.56	.32	.53	.45	.47
ATACT19	Read students writing	.58	.50	.56	.52	.56	.63	.54	.71	.37	.45	.60	.51	.41	.52	.48	.45
ATACT27	Read other subject areas	.34	.40	.39	.66	.40	.37	.31	.24	.08	.67	.30	.19	.37	.52	.42	.22
ATENST4	Encourage stds./hold discussions	.43	.38	.45	.42	.52	-.03	.63	.33	.48	.26	-.07	.21	.40	.31	.61	.45
	Variance accounted for	.31	.28	.31	.33	.26	.28	.30	.26	.27	.36	.29	.27	.22	.32	.32	.27
	Reliability estimate	.68	.63	.68	.70	.59	.63	.67	.60	.62	.75	.65	.62	.48	.70	.70	.61

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATACT05	Silent reading in library	.11	.52	.58	.64	.47	.62	.48	.59	.63	.56	.34	.64
ATACT07	Listen to students in groups	-.23	.55	.42	.52	.59	.54	.62	.52	.38	.44	.64	.65
ATACT09	Discussion of books	.69	.78	.65	.75	.57	.40	.57	.70	.75	.67	.65	.54
ATACT12	Learning library skills	.49	.64	.58	.68	.61	.66	.55	.59	.62	.70	.53	.72
ATACT13	Reading plays or dramas	.22	.58	.04	.54	.59	.63	.58	.42	.15	.55	.60	.52
ATACT19	Read students writing	.58	.38	.35	.58	.56	.56	.54	.58	.55	.62	.70	.64
ATACT27	Read other subject areas	-.09	.18	.42	.50	.38	.39	.28	.23	.21	.31	.11	.52
ATENST4	Encourage stds./hold discussions	.73	.56	.40	.56	.51	.08	.35	.51	.58	.15	.59	.37
	Variance accounted for	.21	.30	.22	.36	.29	.27	.26	.28	.27	.28	.30	.34
	Reliability estimate	.47	.67	.48	.75	.65	.61	.59	.64	.62	.64	.67	.72

Story reading aloud

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	INS	IRL	ITA
ATACT08	Listen to students in class	.77	.82	.84	.71	.85	.80	.81	.79	.82	.53	.77	.79	.89	.70	.80	.63
ATENST3	Encourage stds./read attr. stories	.54	.52	.25	.09	.53	.72	.68	.58	.49	.49	.35	.68	.77	.23	.66	.49
ATSTR09	Strategies/read aloud to children	.75	.80	.85	.45	.89	.89	.78	.78	.79	.45	.67	.72	.89	.71	.78	.73
ATVIE06	Views/read to children every day	.69	.70	.78	.64	.76	.58	.60	.64	.66	.77	.62	.56	.08	.67	.77	.56
ATVIE21	Views/read self-written texts	.40	.29	.19	.54	.02	-.27	.33	.39	-.18	.19	.08	.36	.37	-.18	.37	.13
	Variance accounted for	.42	.43	.43	.28	.47	.47	.44	.42	.40	.27	.31	.41	.46	.30	.48	.30
	Reliability estimate	.65	.67	.66	.37	.72	.72	.68	.66	.63	.33	.44	.64	.71	.43	.73	.42

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATACT08	Listen to students in class	.86	.83	.82	.76	.73	.83	.79	.81	.75	.76	.86	.85
ATENST3	Encourage stds./read attr. stories	.61	.57	.30	.36	.58	.50	.39	.46	.67	.59	.70	.39
ATSTR09	Strategies/read aloud to children	.88	.80	.78	.76	.75	.60	.77	.86	.81	.47	.84	.76
ATVIE06	Views/read to children every day	.64	.53	.81	.29	.59	.45	.70	.64	.52	.56	.69	.39
ATVIE21	Views/read self-written texts	.52	.47	.43	.49	.31	.27	.26	.28	.15	.28	.30	.05
	Variance accounted for	.51	.43	.44	.32	.37	.32	.38	.42	.39	.31	.50	.32
	Reliability estimate	.76	.67	.68	.47	.58	.46	.60	.65	.61	.44	.75	.47

Active teaching of comprehension

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRE	ITA
ATACT10	Rdg.act./systematic vocabulary	.59	.27	.19	.48	.24	.24	.39	.60	.41	.28	.46	.37	.57	.55	.54	.54
ATARE04	Rdg.aspects/literary appreciation	.57	.43	.59	.42	.32	.26	.69	.38	.58	.55	.48	.59	.46	.52	.62	.41
ATMETH2	Methods/tch.-made vocab. tests	.64	.69	.33	.61	.59	.58	.55	.39	.32	.63	.47	.55	.46	.42	.60	.50
ATMETH4	Methods/formal tests of compreh.	.53	.33	.16	.42	.37	.44	.32	.50	.34	.41	.41	.40	.41	.39	.38	.44
ATSTR01	Strategies/introduce background	.59	.55	.57	.16	.58	.54	.41	.62	.35	.33	.72	.23	.63	.53	.41	.56
ATSTR02	Strategies/ask children to describe	.61	.45	.67	.59	.64	.65	.54	.72	.46	.61	.68	.70	.72	.64	.63	.56
ATSTR07	Strategies/show how to understnd.	.62	.64	.70	.48	.67	-.24	.59	.56	.66	.56	.63	.65	.58	.62	.57	.65
ATSTR08	Strategies/compare stories	.68	.53	.77	.62	.62	.61	.63	.66	.64	.57	.67	.51	.67	.62	.52	.62
ATSTR13	Strategies/use self-prepared mats.	.50	.54	.59	.58	.36	.65	.46	.60	.28	.50	.53	.44	.44	.49	.56	.48
	Variance accounted for	.35	.26	.30	.25	.26	.25	.27	.32	.22	.26	.33	.26	.31	.29	.30	.29
	Reliability estimate	.77	.64	.71	.63	.65	.62	.67	.74	.56	.64	.74	.65	.72	.69	.70	.69

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATACT10	Rdg.act./systematic vocabulary	.39	.38	.56	.38	.51	.57	.43	.56	.55	.39	.37	.58
ATARE04	Rdg.aspects/literary appreciation	.54	.54	.43	.58	.57	.40	.56	.50	.69	.62	.55	.73
ATMETH2	Methods/tch.-made vocab. tests	.38	.42	.40	.55	.63	.54	.48	.53	.41	.44	.40	.65
ATMETH4	Methods/formal tests of compreh.	.42	.37	.32	.30	.39	.53	.37	.24	.30	.05	.29	.69
ATSTR01	Strategies/introduce background	.54	.51	.37	.64	.54	.58	.50	.44	.47	.61	.47	.57
ATSTR02	Strategies/ask children to describe	.56	.68	.64	.21	.63	.56	.64	.74	.61	.61	.54	.62
ATSTR07	Strategies/show how to understnd.	.61	.63	.66	.43	.56	.61	.56	.67	.62	.60	.64	.59
ATSTR08	Strategies/compare stories	.72	.74	.60	.64	.62	.54	.68	.66	.54	.59	.63	.43
ATSTR13	Strategies/use self-prepared mats.	.30	.45	.23	.61	.60	.62	.56	.23	.23	.58	.56	.51
	Variance accounted for	.26	.29	.24	.25	.32	.31	.29	.29	.26	.28	.26	.36
	Reliability estimate	.64	.70	.60	.63	.73	.72	.69	.69	.65	.68	.64	.78

Comprehension through graded materials

	Variables	INT	BEL (Fr)	CAN (BC)	CYP	DEN	FIN	FRA	GER (E)	GER (W)	GRE	HNK	HUN	ICE	IND	IRL	ITA
ATVIE01	Views/expect accuracy	.59	.27	.58	.46	.51	.32	.44	.58	.48	.50	.40	.30	.65	.54	.65	.53
ATVIE04	Views/should enjoy reading	.42	.56	.23	.31	.34	-.34	.48	.33	.44	.07	.59	.48	.41	.47	.41	.33
ATVIE09	Views/must understand why rdg.	.66	.38	.65	.63	.69	.60	.36	.64	.44	.58	.42	.64	.66	.78	.66	.86
ATVIE13	Views/insuff.comprehension skills	.52	.48	.16	.37	.33	.49	.54	.13	.24	.53	.52	.06	.41	.03	.41	.37
ATVIE16	Views/no new bk until finish last	.49	.41	.54	.26	.40	.44	.58	.48	.37	.49	.33	.24	.33	-.11	.33	.28
ATVIE22	Views/understand what they read	.69	.67	.68	.77	.70	.77	.56	.66	.75	.57	.70	.80	.65	.74	.65	.81
ATVIE23	Views/choose own books	.52	.47	-.04	.53	.21	.43	.41	.35	.52	.40	.26	.17	.39	.01	.39	.14
	Variance accounted for	.32	.23	.23	.26	.24	.26	.24	.24	.24	.23	.23	.21	.27	.24	.27	.29
	Reliability estimate	.64	.44	.43	.51	.47	.52	.47	.46	.46	.43	.45	.36	.54	.47	.54	.59

	Variables	NET	NZL	NOR	POR	SIN	SLO	SPA	SWE	SWI	T/T	USA	VEN
ATVIE01	Views/expect accuracy	.26	.26	.57	.55	.58	.62	.36	.22	.34	.36	.50	.31
ATVIE04	Views/should enjoy reading	-.45	.21	.45	.63	.23	.49	.59	.46	.49	.42	.27	.53
ATVIE09	Views/must understand why rdg.	.58	.66	.59	.76	.65	.76	.52	.73	.74	.76	.62	.59
ATVIE13	Views/insuff.comprehension skills	.39	.45	.56	.46	.51	.42	.16	.33	.45	.60	.48	.63
ATVIE16	Views/no new bk until finish last	.67	.39	.44	.35	.63	.36	.47	.41	.51	.35	.53	.40
ATVIE22	Views/understand what they read	.67	.81	.65	.76	.79	.69	.69	.84	.72	.73	.76	.61
ATVIE23	Views/choose own books	-.32	.43	.39	.24	.29	-.10	.38	.32	.46	.37	.19	.53
	Variance accounted for	.25	.25	.28	.32	.31	.28	.23	.27	.30	.29	.26	.28
	Reliability estimate	.50	.50	.57	.65	.63	.58	.45	.54	.61	.59	.53	.56

UNITAB_B.LST

Univariates for the construct: (BSFSFF) Reading Activities I: Fiction and Faction 17:41 Wednesday, September 9, 1992 1

OBS	ID_CNTRY	N	MEAN	STD
1	80	2124	-0.32096	0.80751
2	110	3051	0.33632	1.08682
3	155	4077	-0.37999	0.85515
4	200	1113	0.26628	0.95363
5	215	3277	-0.56193	0.72019
6	275	1196	-0.41713	0.67915
7	280	2267	-0.20294	0.83042
8	290	1815	-0.43667	0.68869
9	295	4035	-0.47674	0.70291
10	305	3105	0.27034	0.98772
11	345	2889	-0.35187	0.81951
12	350	2925	0.39754	0.86270
13	355	3193	-0.61837	0.73917
14	380	2797	-0.20540	0.85967
15	390	2301	0.16193	0.97933
16	565	3259	-0.63816	0.68917
17	570	2586	-0.46547	0.77669
18	585	0	.	.
19	590	1880	-0.65369	0.71971
20	625	9677	1.13366	1.03856
21	635	2815	0.00927	0.89818
22	690	4762	0.15193	0.90155
23	715	7193	0.01712	0.92460
24	755	2990	-0.47083	0.79074
25	760	5490	-0.41596	0.73094
26	780	2525	0.40279	0.88548
27	795	1974	0.70690	1.11996
28	840	3100	-0.16576	0.99264
29	850	2128	0.20941	1.09723
30	856	2838	-0.08252	0.84405
31	870	1814	0.63694	0.92957

OBS	ID_CNTRY	N	MEAN	STD
1	80	2282	-0.58259	0.80544
2	110	3834	0.59286	0.94972
3	155	4263	-0.04457	1.02287
4	200	1212	-0.11672	0.87426
5	215	3435	-0.33300	0.80202
6	275	1243	-0.09455	0.93248
7	280	2296	-0.12626	0.84372
8	290	1798	-0.41479	0.79016
9	295	3998	-0.37871	0.78017
10	305	3289	-0.16470	0.89907
11	345	2952	-0.34744	0.85031
12	350	3107	0.09200	0.79722
13	355	3434	-0.25819	1.01847
14	380	3110	0.05042	0.96922
15	390	2756	-0.06083	0.96073
16	565	3340	-0.37190	0.89909
17	570	2761	0.20394	1.02295
18	585	0	.	.
19	590	2005	-0.25396	0.93963
20	625	9693	0.55822	1.01115
21	635	3142	-0.51126	0.87022
22	690	4741	0.06214	0.93458
23	715	7754	-0.41867	0.85250
24	755	3214	-0.29858	0.91551
25	760	5891	-0.26757	0.86464
26	780	2680	0.13000	0.92548
27	795	2533	0.78707	1.07630
28	840	3154	0.01785	1.02337
29	850	2421	-0.19977	0.99724
30	856	2980	-0.07826	0.86562
31	870	2159	0.77241	0.86699

Univariates for the construct: (BSFSSH) Reading Activities III: School & Homework 17:41 Wednesday, September 9, 1992 3

OBS	ID_CNTRY	N	MEAN	STD
1	80	2231	-0.79099	0.95213
2	110	3822	0.34701	0.89595
3	155	4183	-0.05922	0.96415
4	200	1220	0.10454	0.90503
5	215	3418	-0.30445	0.97161
6	275	1226	-0.15043	0.96551
7	280	2303	-0.03542	1.00831
8	290	1807	-0.37136	0.91040
9	295	4030	-0.45414	0.89615
10	305	3539	0.29231	0.88505
11	345	2956	-0.26188	0.99864
12	350	3128	-0.30761	0.85485
13	355	3305	-0.11731	0.95729
14	380	2904	0.25640	0.98419
15	390	2884	0.57635	0.95181
16	565	3371	-0.60115	0.98751
17	570	2449	-0.37834	0.95858
18	585	0	.	.
19	590	2015	-0.12659	0.97129
20	625	9694	0.59949	0.88091
21	635	2993	-0.54068	0.90036
22	690	4693	0.14875	0.72033
23	715	7620	-0.06017	0.98830
24	755	3172	-0.10936	0.97189
25	760	5867	-0.31134	0.98160
26	780	2717	0.23337	0.83489
27	795	2326	0.56182	0.84237
28	840	3038	0.31376	0.96680
29	850	2632	-0.35871	0.90562
30	856	2954	-0.14416	0.87723
31	870	2085	0.57632	0.77557

Univariates for the construct: (BSFSRR) Reading Activities IV: Recreation
Reading 17:41 Wednesday, September 9, 1992 4

OBS	ID_CNTRY	N	MEAN	STD
1	80	2387	-0.23011	1.02111
2	110	4235	-1.08722	0.93412
3	155	4354	-0.29109	1.00697
4	200	1284	0.34791	0.91332
5	215	3571	-0.07608	0.93476
6	275	1274	0.79871	0.77009
7	280	2397	-0.13054	0.94291
8	290	1804	-0.28008	0.89396
9	295	4024	-0.17265	0.96635
10	305	3491	0.20113	0.98448
11	345	3039	-0.32704	0.95495
12	350	3170	0.35078	0.82615
13	355	3560	0.15819	0.97662
14	380	3271	0.07804	0.90385
15	390	2913	0.17891	0.98807
16	565	3401	-0.45910	1.01608
17	570	2810	-0.18730	0.92406
18	585	0	.	.
19	590	2051	0.27146	0.90683
20	625	9701	0.03526	0.94268
21	635	3154	-0.15113	0.97915
22	690	4776	0.32962	0.84897
23	715	8016	-0.08920	1.04934
24	755	3291	0.45450	0.90721
25	760	6064	-0.02101	0.98118
26	780	2726	0.05022	0.87597
27	795	2676	0.29174	0.91001
28	840	3194	-0.09208	1.00785
29	850	2695	0.21325	0.99880
30	856	3116	0.02783	0.91770
31	870	2293	-0.71308	0.95809

OBS	ID_CNTRY	N	MEAN	STD
1	80	2286	-0.30344	0.93943
2	110	3872	-0.18253	0.84381
3	155	4237	-0.30924	0.95787
4	200	1208	0.28344	1.07784
5	215	3497	-0.18395	0.96478
6	275	1219	-0.02379	0.98785
7	280	2332	-0.19281	0.97375
8	290	1801	0.09731	1.02977
9	295	4015	-0.16698	0.99766
10	305	3362	0.09628	0.99658
11	345	2994	-0.20104	0.91583
12	350	3115	0.56941	1.05293
13	355	3393	-0.06953	0.96581
14	380	3065	-0.08480	1.00138
15	390	2801	-0.07788	0.98150
16	565	3366	-0.21437	0.91098
17	570	2737	-0.15859	0.95825
18	585	0	.	.
19	590	1977	0.00363	1.00668
20	625	9703	0.42776	0.86463
21	635	3011	0.10709	0.95845
22	690	4776	0.08261	1.02529
23	715	7669	-0.24940	0.97768
24	755	3166	-0.00074	1.05511
25	760	5834	-0.09359	1.00338
26	780	2675	0.13047	0.88232
27	795	2424	0.26226	1.06906
28	840	3177	0.06027	1.03971
29	850	2448	0.41015	1.09993
30	856	3028	0.26542	1.05321
31	870	2117	0.19967	0.89523

Univariates for the construct: (BSFSNP) Reading Activities VI: News & Politics 17:41 Wednesday, September 9, 1992 6

OBS	ID_CNTRY	N	MEAN	STD
1	80	2409	-0.67690	0.85166
2	110	4351	-0.04666	0.86991
3	155	4364	-0.47863	0.85450
4	200	1278	0.40366	0.93388
5	215	3615	-0.10267	1.01628
6	275	1267	0.12505	0.95164
7	280	2414	-0.50589	0.89914
8	290	1804	-0.15044	1.01940
9	295	4024	-0.23876	0.98612
10	305	3522	0.26555	1.00807
11	345	3068	0.50695	0.95744
12	350	3200	0.19942	0.95575
13	355	3590	-0.32189	0.98116
14	380	3292	-0.16576	0.95917
15	390	2923	-0.59454	0.85645
16	565	3414	-0.39582	0.94306
17	570	2858	-0.27765	0.91496
18	585	0	.	.
19	590	2072	0.17205	0.95729
20	625	9703	0.32097	0.83231
21	635	3187	-0.07979	0.94131
22	690	4782	0.46181	0.91355
23	715	8002	-0.60990	0.85086
24	755	3333	0.10732	0.99445
25	760	6144	-0.12741	1.02821
26	780	2728	0.60780	0.79793
27	795	2757	0.27151	0.93570
28	840	3207	0.02071	0.96691
29	850	2759	0.07149	0.99712
30	856	3115	-0.30941	0.95252
31	870	2368	0.24582	0.92019

unitabc.lst

Univariates for the construct: (ACFSRM) Reading Materials in School
20:07 Wednesday, September 9, 1992 1

OBS	NAME	MEAN	N	STD
1	BEF01	-0.17499	93	0.18050
2	CAC01	0.46688	130	0.49365
3	CYP01	.	0	.
4	DEN01	0.66129	153	0.55133
5	FIN01	-0.11350	51	0.11081
6	FRA01	-0.17133	103	0.11146
7	GER01	-0.19999	105	0.06483
8	GER02	-0.22973	41	0.07117
9	GRC01	-0.23113	122	0.08984
10	HKO01	0.08740	67	0.31801
11	HUN01	0.27041	134	0.40743
12	ICE01	0.10773	149	0.21852
13	INS01	-0.08944	161	0.28121
14	IRE01	-0.10930	95	0.23058
15	ITA01	-0.17726	71	0.10022
16	NET01	-0.23369	82	0.04732
17	NEZ01	0.19621	157	0.28578
18	NOR01	-0.04246	171	0.26882
19	POR01	-0.24857	127	0.05962
20	SIN01	0.65006	205	0.59319
21	SPA01	-0.02139	270	0.24476
22	SWE01	0.04987	105	0.35868
23	SWI01	-0.11122	162	0.18260
24	TRI01	-0.19820	104	0.10912
25	USA01	0.40804	144	0.49940
26	VEN01	-0.12822	79	0.29965
27	YOU01	0.31914	137	0.35326

Univariates for the construct: (ACFSCR) Community Resources
20:07 Wednesday, September 9, 1992 2

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.03189	141	0.39644
2	CAC01	0.14597	140	0.26043
3	CYP01	0.12545	177	0.28097
4	DEN01	0.03569	158	0.30241
5	FIN01	0.02792	70	0.36675
6	FRA01	-0.13865	120	0.44368
7	GER01	-0.02731	136	0.34498
8	GER02	-0.07296	95	0.30661
9	GRC01	0.04131	165	0.29813
10	HKO01	-0.02051	113	0.32577
11	HUN01	0.10560	141	0.24938
12	ICE01	0.00070	161	0.42955
13	INS01	-0.62993	136	0.61066
14	IRE01	0.00165	116	0.33186
15	ITA01	0.08535	128	0.37936
16	NET01	.	0	.
17	NEZ01	0.14437	168	0.27952
18	NOR01	0.03754	171	0.30483
19	POR01	-0.47572	101	0.60964
20	SIN01	.	0	.
21	SPA01	-0.05114	264	0.44816
22	SWE01	0.08360	116	0.29226
23	SWI01	.	0	.
24	TRI01	-0.20040	150	0.43167
25	USA01	0.17490	161	0.18895
26	VEN01	-0.14775	104	0.56913
27	YOU01	-0.09816	139	0.35106

UNITABS.lst

Univariates for the construct: (ASFSHL) Home Literacy Interaction
23:05 Wednesday, September 9, 1992 1

OBS	_NAME_	MEAN	N	STD
1	BEF01	0.02603	1618	0.91202
2	CAC01	0.27951	1530	1.00520
3	CYP01	0.15814	1000	0.84517
4	DEN01	0.31905	2443	0.92658
5	FIN01	-0.18929	1005	0.89643
6	FRA01	-0.15108	876	0.86887
7	GER01	-0.18802	1529	0.82562
8	GER02	0.09456	1039	0.90343
9	GRC01	0.06864	1947	0.81266
10	HKO01	-0.11050	1128	0.81307
11	HUN01	0.00636	2605	0.87896
12	ICE01	0.01183	3015	0.90798
13	INS01	-0.33079	1068	0.84115
14	IRE01	0.31273	1698	0.88688
15	ITA01	0.28876	1358	0.96170
16	NET01	-0.29836	948	0.88139
17	NEZ01	0.11935	1703	0.94434
18	NOR01	0.16485	1387	0.88330
19	POR01	0.46208	1782	0.93802
20	SIN01	0.38832	2650	0.99001
21	SPA01	-0.05297	4503	0.94756
22	SWE01	-0.01580	3388	0.89703
23	SWI01	-0.17444	2371	0.85684
24	TRI01	0.68548	2293	1.03444
25	USA01	-0.33013	3809	0.62587
26	VEN01	0.41787	2200	0.98727
27	YOU01	0.27097	2541	0.92632

Univariates for the construct: (ASFSRA) Read Aloud
23:05 Wednesday, September 9, 1992 2

OBS	_NAME_	MEAN	N	STD
1	BEF01	0.17153	1646	0.90856
2	CAC01	0.18348	1582	1.09081
3	CYP01	-0.13723	1041	0.89241
4	DEN01	0.34163	2495	1.03249
5	FIN01	0.08254	1030	0.93850
6	FRA01	-0.23392	913	0.84940
7	GER01	-0.20797	1751	0.80377
8	GER02	0.18228	1143	0.97272
9	GRC01	.	0	.
10	HKO01	-0.27375	1152	0.84172
11	HUN01	-0.11498	2669	0.91261
12	ICE01	-0.06973	3106	1.01678
13	INS01	-0.32373	1162	0.69753
14	IRE01	0.09001	1738	0.96747
15	ITA01	-0.02733	1394	0.96803
16	NET01	0.10748	975	0.89173
17	NEZ01	0.13676	1737	1.03363
18	NOR01	0.03489	1464	0.99394
19	POR01	0.14620	1806	0.99577
20	SIN01	.	0	.
21	SPA01	.	0	.
22	SWE01	0.10712	3448	0.94773
23	SWI01	0.05750	2517	0.88450
24	TRI01	0.35063	2391	1.10785
25	USA01	-0.73494	4064	0.35818
26	VEN01	0.05593	2380	0.97302
27	YOU01	-0.07955	2770	0.93367

Univariates for the construct: (ASFSVR) Voluntary Reading
23:05 Wednesday, September 9, 1992 3

OBS	NAME	MEAN	N	STD
1	BEF01	0.04330	2535	0.98349
2	CAC01	-0.29180	2478	0.93205
3	CYP01	0.20170	1411	0.96591
4	DEN01	-0.03690	3286	0.96092
5	FIN01	0.58844	1463	0.84972
6	FRA01	0.01595	1796	0.93402
7	GER01	-0.38004	2052	0.89589
8	GER02	-0.16181	1403	0.95849
9	GRC01	0.14706	3301	0.94807
10	HKO01	-0.03424	3160	0.87068
11	HUN01	0.21420	2844	0.89779
12	ICE01	-0.12117	3752	0.97043
13	INS01	-0.17935	2974	1.04806
14	IRE01	-0.15754	2602	0.91856
15	ITA01	-0.20629	2064	0.94009
16	NET01	-0.27514	1634	0.90180
17	NEZ01	-0.19229	2896	0.88514
18	NOR01	0.06221	2308	0.91575
19	POR01	0.11020	2636	1.04910
20	SIN01	0.12646	7215	0.93224
21	SPA01	-0.20873	7958	1.01041
22	SWE01	0.17769	4146	0.93101
23	SWI01	-0.11493	3217	0.97721
24	TRI01	-0.02939	3476	1.04414
25	USA01	-0.20931	5844	0.88583
26	VEN01	0.29628	3450	1.19416
27	YOU01	0.26452	3139	0.95455

Univariates for the construct: (ASFSRC) Reading in Class
23:05 Wednesday, September 9, 1992 4

OBS	NAME	MEAN	N	STD
1	BEF01	-0.57994	2514	0.99334
2	CAC01	-0.36101	2440	0.95950
3	CYP01	0.58624	1411	0.68052
4	DEN01	0.10094	3262	0.78586
5	FIN01	0.32189	1460	0.72127
6	FRA01	-0.17937	1783	1.06801
7	GER01	-0.62260	2190	0.91815
8	GER02	-0.31346	1529	0.88256
9	GRC01	0.66524	3244	0.69534
10	HKO01	-0.12789	3210	0.82901
11	HUN01	0.15191	2800	0.63663
12	ICE01	-0.41516	3732	0.96791
13	INS01	-0.12645	2953	0.89505
14	IRE01	0.38362	2566	0.85628
15	ITA01	-0.02107	2050	0.99797
16	NET01	-0.64670	1648	0.83429
17	NEZ01	-0.28624	2856	1.04798
18	NOR01	-0.10578	2263	0.96479
19	POR01	0.44440	2624	0.81770
20	SIN01	0.72225	7222	0.80190
21	SPA01	0.10984	7981	1.07337
22	SWE01	-0.16069	4080	0.90498
23	SWI01	-0.57103	3209	0.90049
24	TRI01	0.29307	3472	0.89506
25	USA01	0.30021	5936	0.94102
26	VEN01	-0.01151	3443	1.14842
27	YOU01	0.14121	3140	0.87392

UNITABT

Univariates for the construct: (ATFSCI) Comprehension Instruction
11:31 Thursday, September 10, 1992 1

OBS	NAME	MEAN	N	STD
1	BEF01	-0.40455	133	0.30739
2	CAC01	0.22734	142	0.35412
3	CYP01	0.51150	270	0.29331
4	DEN01	-0.26573	197	0.27952
5	FIN01	-0.14027	64	0.29909
6	FRA01	-0.42532	112	0.28130
7	GER01	-0.16772	129	0.25640
8	GER02	0.01531	87	0.24238
9	GRC01	0.44370	154	0.36316
10	HKO01	-0.20666	143	0.31517
11	HUN01	0.47352	135	0.25610
12	ICE01	-0.49292	247	0.30875
13	INS01	0.05738	161	0.32166
14	IRE01	-0.17886	108	0.38656
15	ITA01	0.36003	122	0.38164
16	NET01	-0.29391	88	0.30043
17	NEZ01	0.10445	161	0.35795
18	NOR01	-0.13348	169	0.33549
19	POR01	0.33030	120	0.33423
20	SIN01	-0.01096	197	0.36723
21	SPA01	-0.11324	224	0.39133
22	SWE01	-0.18019	217	0.33339
23	SWI01	-0.20356	211	0.28144
24	TRI01	0.08657	189	0.35699
25	USA01	0.14579	281	0.35306
26	VEN01	0.16020	111	0.43028
27	YOU01	0.17455	122	0.22064

Univariates for the construct: (ATFSAL) Assessment of Low Order Skills 11:31 Thursday, September 10, 1992 2

OBS	NAME	MEAN	N	STD
1	BEF01	-0.06803	133	0.37921
2	CAC01	0.05504	138	0.36249
3	CYP01	0.19627	242	0.25073
4	DEN01	-0.04704	159	0.38349
5	FIN01	-0.20302	68	0.32626
6	FRA01	-0.18194	114	0.42591
7	GER01	-0.15297	143	0.43020
8	GER02	-0.05777	98	0.33938
9	GRC01	0.26635	164	0.20907
10	HKO01	-0.04175	150	0.34649
11	HUN01	0.07908	131	0.31243
12	ICE01	-0.10327	212	0.40137
13	INS01	0.28954	162	0.20308
14	IRE01	-0.04805	108	0.34758
15	ITA01	0.20713	134	0.27471
16	NET01	-0.49100	91	0.39621
17	NEZ01	-0.16592	154	0.41821
18	NOR01	-0.15212	149	0.40288
19	POR01	0.35196	131	0.19208
20	SIN01	-0.02484	187	0.37479
21	SPA01	0.14804	204	0.28583
22	SWE01	-0.64191	199	0.62506
23	SWI01	-0.39829	196	0.47094
24	TRI01	0.34370	201	0.17727
25	USA01	0.26760	281	0.24018
26	VEN01	0.15954	106	0.33800
27	YOU01	0.27700	126	0.16566

Univariates for the construct: (ATFSHD) High Demand and Structure 11:31 Thursday, September 10, 1992 3

OBS	NAME	MEAN	N	STD
1	BEF01	-0.10414	137	0.35346
2	CAC01	-0.50944	147	0.32673
3	CYP01	0.15783	301	0.28143
4	DEN01	-0.31833	198	0.31711
5	FIN01	-0.20901	68	0.42497
6	FRA01	-0.22091	125	0.32813
7	GER01	-0.17726	145	0.31791
8	GER02	0.02286	98	0.28520
9	GRC01	0.37683	166	0.27345
10	HKO01	0.35455	146	0.20087
11	HUN01	0.50007	141	0.29463
12	ICE01	-0.05179	260	0.33619
13	INS01	0.54543	171	0.24016
14	IRE01	-0.04881	116	0.27043
15	ITA01	0.29978	144	0.30315
16	NET01	-0.07851	92	0.21637
17	NEZ01	-0.33769	171	0.30659
18	NOR01	-0.18891	174	0.30497
19	POR01	0.35143	148	0.30525
20	SIN01	-0.04542	203	0.29599
21	SPA01	0.28671	270	0.29769
22	SWE01	-0.48014	232	0.33897
23	SWI01	-0.26363	225	0.26079
24	TRI01	0.17888	209	0.31928
25	USA01	-0.26436	292	0.33668
26	VEN01	0.41335	129	0.30290
27	YOU01	-0.17227	133	0.31777

Univariates for the construct: (ATFSER) Encouragement to Read
11:31 Thursday, September 10, 1992 4

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.02205	146	0.45359
2	CAC01	0.19536	145	0.29534
3	CYP01	0.25354	209	0.28448
4	DEN01	-0.15535	202	0.38484
5	FIN01	-0.08037	68	0.42806
6	FRA01	-0.03039	129	0.46405
7	GER01	-0.43186	145	0.45221
8	GER02	-0.31719	97	0.44692
9	GRC01	0.12735	174	0.35711
10	HKO01	-0.19635	156	0.45660
11	HUN01	0.09610	142	0.34398
12	ICE01	-0.01847	272	0.40975
13	INS01	0.16999	170	0.32211
14	IRE01	0.09562	116	0.39309
15	ITA01	0.07472	146	0.35763
16	NET01	-0.16699	96	0.45513
17	NEZ01	0.11170	175	0.34192
18	NOR01	-0.17555	177	0.40463
19	POR01	0.12309	144	0.37044
20	SIN01	0.20602	204	0.32631
21	SPA01	0.17110	299	0.34930
22	SWE01	-0.02313	227	0.36931
23	SWI01	-0.38398	224	0.41459
24	TRI01	0.10553	208	0.36715
25	USA01	0.15534	298	0.32541
26	VEN01	0.08879	126	0.41022
27	YOU01	0.09668	136	0.36738

OBS	NAME	MEAN	N	STD
1	BEF01	-0.16106	129	0.47867
2	CAC01	0.26708	137	0.31226
3	CYP01	-0.08078	269	0.40618
4	DEN01	0.24505	190	0.33166
5	FIN01	0.23649	70	0.28696
6	FRA01	-0.24726	116	0.43568
7	GER01	-0.10349	143	0.36141
8	GER02	-0.01044	99	0.36084
9	GRC01	-0.09181	163	0.44411
10	HKO01	-0.20166	150	0.44334
11	HUN01	0.32620	138	0.20908
12	ICE01	-0.01305	253	0.44464
13	INS01	0.19611	162	0.32014
14	IRE01	-0.27427	115	0.41702
15	ITA01	0.25620	127	0.28309
16	NET01	-0.23647	93	0.45478
17	NEZ01	0.12889	167	0.34115
18	NOR01	0.02587	164	0.33886
19	POR01	0.05071	136	0.41016
20	SIN01	-0.18371	200	0.43265
21	SPA01	-0.08121	257	0.42339
22	SWE01	-0.15912	209	0.34413
23	SWI01	-0.23847	205	0.38221
24	TRI01	0.05379	176	0.40780
25	USA01	0.02122	288	0.39723
26	VEN01	0.00673	110	0.50715
27	YOU01	0.26411	119	0.26688

Univariates for the construct: (ATFSAE) General Emphasis on Assessment 11:31 Thursday, September 10, 1992 6

OBS	NAME	MEAN	N	STD
1	BEF01	0.04570	144	0.33936
2	CAC01	-0.62191	134	0.47515
3	CYP01	0.18819	303	0.29005
4	DEN01	-0.25327	188	0.41498
5	FIN01	-0.11279	64	0.27248
6	FRA01	0.06492	117	0.30750
7	GER01	-0.41237	139	0.36811
8	GER02	-0.25516	95	0.35531
9	GRC01	0.23171	171	0.31284
10	HKO01	-0.02240	142	0.29006
11	HUN01	0.31157	142	0.19293
12	ICE01	-0.18411	240	0.33362
13	INS01	0.20220	163	0.26762
14	IRE01	0.19341	113	0.33028
15	ITA01	0.27351	140	0.27057
16	NET01	-0.27625	93	0.34852
17	NEZ01	-0.40193	161	0.52478
18	NOR01	0.09339	155	0.33576
19	POR01	0.42081	141	0.20679
20	SIN01	0.17992	201	0.25125
21	SPA01	0.14606	271	0.33388
22	SWE01	-0.25170	213	0.44610
23	SWI01	-0.28422	208	0.36375
24	TRI01	0.24253	198	0.28951
25	USA01	0.15246	292	0.33303
26	VEN01	0.12717	116	0.38845
27	YOU01	0.19929	131	0.23517

Univariates for the construct: (ATFSPT) Phonics Teaching
11:31 Thursday, September 10, 1992 7

OBS	NAME	MEAN	N	STD
1	BEF01	-0.22622	132	0.41119
2	CAC01	0.15991	138	0.37436
3	CYP01	-0.30947	277	0.49505
4	DEN01	0.03548	174	0.31819
5	FIN01	-0.10878	68	0.29189
6	FRA01	-0.35168	113	0.44866
7	GER01	0.00283	135	0.31413
8	GER02	0.15996	89	0.27241
9	GRC01	-0.13506	156	0.47004
10	HKO01	0.21703	151	0.32568
11	HUN01	0.06794	135	0.38838
12	ICE01	-0.36255	241	0.45697
13	INS01	0.27187	165	0.19687
14	IRE01	0.23492	118	0.29918
15	ITA01	0.13298	125	0.37046
16	NET01	-0.21148	86	0.44680
17	NEZ01	0.05574	163	0.39235
18	NOR01	-0.19042	167	0.38881
19	POR01	0.13869	124	0.46011
20	SIN01	0.20059	199	0.21998
21	SPA01	0.14432	270	0.33475
22	SWE01	-0.49063	203	0.39929
23	SWI01	-0.22428	211	0.36504
24	TRI01	0.35219	210	0.21820
25	USA01	0.14426	286	0.39808
26	VEN01	0.07211	113	0.38226
27	YOU01	0.14212	124	0.25459

Univariates for the construct: (ATFSTE) Teacher Readership (Expository)
11:31 Thursday, September 10, 1992 8

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.18686	145	0.38991
2	CAC01	0.19643	146	0.44467
3	CYP01	-0.02558	284	0.36259
4	DEN01	-0.15384	199	0.38304
5	FIN01	0.02552	66	0.37031
6	FRA01	-0.07311	131	0.36259
7	GER01	0.00083	144	0.34380
8	GER02	-0.09406	99	0.31250
9	GRC01	0.05405	168	0.36429
10	HKO01	0.05902	144	0.38522
11	HUN01	0.19898	136	0.32730
12	ICE01	-0.26243	256	0.38327
13	INS01	0.54295	157	0.32553
14	IRE01	-0.19788	110	0.40272
15	ITA01	-0.03782	135	0.40408
16	NET01	0.18742	86	0.44793
17	NEZ01	-0.03949	166	0.37358
18	NOR01	-0.27643	173	0.32780
19	POR01	-0.04896	137	0.38923
20	SIN01	-0.23862	195	0.34019
21	SPA01	-0.09444	253	0.40344
22	SWE01	-0.14549	225	0.38960
23	SWI01	-0.15179	219	0.35959
24	TRI01	0.20134	197	0.39870
25	USA01	-0.09851	288	0.41716
26	VEN01	0.13231	114	0.42760
27	YOU01	0.47108	133	0.27032

Univariates for the construct: (ATFSTL) Teacher Readership (Literature)
11:31 Thursday, September 10, 1992 9

OBS	NAME	MEAN	N	STD
1	BEF01	-0.25473	146	0.37567
2	CAC01	0.20644	144	0.36323
3	CYP01	-0.07367	294	0.44823
4	DEN01	-0.01168	201	0.36365
5	FIN01	-0.01767	68	0.40472
6	FRA01	-0.29900	130	0.40347
7	GER01	-0.11897	144	0.38862
8	GER02	-0.11580	99	0.35680
9	GRC01	-0.12353	168	0.41765
10	HKO01	-0.12381	148	0.43164
11	HUN01	0.27739	138	0.26116
12	ICE01	0.03992	260	0.34070
13	INS01	0.29385	161	0.42542
14	IRE01	-0.07111	111	0.44513
15	ITA01	0.19423	140	0.37379
16	NET01	-0.21347	85	0.39089
17	NEZ01	0.18549	166	0.40823
18	NOR01	0.04307	172	0.37040
19	POR01	-0.06049	136	0.40964
20	SIN01	-0.15260	200	0.41151
21	SPA01	-0.16950	259	0.45840
22	SWE01	0.03417	223	0.35584
23	SWI01	-0.08508	221	0.36867
24	TRI01	0.26812	215	0.38852
25	USA01	0.13534	295	0.36874
26	VEN01	0.07210	121	0.51288
27	YOU01	0.10844	131	0.35526

Univariates for the construct: (ATFSTP) Teacher Readership (Professional)
11:31 Thursday, September 10, 1992 10

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.24220	148	0.36182
2	CAC01	0.15139	146	0.35166
3	CYP01	-0.16546	305	0.39923
4	DEN01	0.11739	206	0.35305
5	FIN01	0.08513	69	0.38277
6	FRA01	-0.10886	133	0.42285
7	GER01	-0.42969	144	0.36807
8	GER02	-0.19531	99	0.36774
9	GRC01	-0.17986	171	0.44714
10	HKO01	-0.14533	153	0.46256
11	HUN01	0.34001	143	0.22862
12	ICE01	-0.08947	271	0.31886
13	INS01	0.25467	159	0.41976
14	IRE01	-0.21735	117	0.40729
15	ITA01	0.26880	142	0.33843
16	NET01	-0.23228	92	0.40230
17	NEZ01	0.08735	170	0.35108
18	NOR01	0.15705	179	0.30644
19	POR01	0.05442	147	0.43606
20	SIN01	-0.24046	205	0.38864
21	SPA01	0.09393	290	0.37848
22	SWE01	0.24908	230	0.28531
23	SWI01	-0.22867	222	0.36320
24	TRI01	0.13330	219	0.39091
25	USA01	0.09487	295	0.37483
26	VEN01	0.26156	121	0.44422
27	YOU01	0.15914	136	0.29713

Univariates for the construct: (ATFSPE) Principal Engagement
11:31 Thursday, September 10, 1992 11

OBS	_NAME_	MEAN	N	STD
1	BEF01	0.04260	133	0.43736
2	CAC01	-0.02846	147	0.30154
3	CYP01	0.44745	289	0.26329
4	DEN01	-0.49192	191	0.25945
5	FIN01	-0.13411	67	0.28185
6	FRA01	-0.20477	100	0.40119
7	GER01	-0.30863	127	0.32532
8	GER02	-0.08904	94	0.34965
9	GRC01	-0.16872	159	0.44248
10	HKO01	0.09862	149	0.41809
11	HUN01	0.15441	139	0.21662
12	ICE01	-0.24886	256	0.29227
13	INS01	0.49353	169	0.21769
14	IRE01	-0.13183	109	0.42052
15	ITA01	0.05229	123	0.40090
16	NET01	-0.12538	85	0.40576
17	NEZ01	0.13314	163	0.33704
18	NOR01	0.00284	157	0.39412
19	POR01	0.00137	129	0.47152
20	SIN01	0.27168	205	0.26859
21	SPA01	.	0	.
22	SWE01	-0.35444	223	0.25945
23	SWI01	-0.30271	217	0.34831
24	TRI01	0.18941	190	0.43534
25	USA01	0.21702	289	0.31994
26	VEN01	0.15349	121	0.47419
27	YOU01	0.23442	133	0.29614

Univariates for the construct: (ATFSSM) Staff Meetings
11:31 Thursday, September 10, 1992 12

OBS	_NAME_	MEAN	N	STD
1	BEF01	2.11260	102	0.87450
2	CAC01	2.19630	143	0.52400
3	CYP01	1.91712	298	0.52577
4	DEN01	2.76140	203	0.46916
5	FIN01	1.76391	67	0.64126
6	FRA01	2.16894	108	0.88007
7	GER01	2.09496	141	0.65257
8	GER02	1.78037	94	0.61049
9	GRC01	1.96746	158	0.76969
10	HKO01	1.44239	146	1.03093
11	HUN01	0.99467	137	0.67006
12	ICE01	2.25100	264	0.50390
13	INS01	0.09425	165	0.70562
14	IRE01	1.92842	106	0.84440
15	ITA01	1.60498	138	0.71598
16	NET01	2.10479	93	0.41616
17	NEZ01	1.82327	171	0.56228
18	NOR01	2.25265	168	0.39897
19	POR01	1.77037	142	0.69164
20	SIN01	1.89004	204	0.54205
21	SPA01	1.93694	271	0.73313
22	SWE01	2.37183	224	0.49671
23	SWI01	2.45262	191	0.65637
24	TRI01	1.82538	206	0.80967
25	USA01	2.05929	290	0.56122
26	VEN01	1.89869	113	0.95038
27	YOU01	1.75749	134	0.50702

Univariates for the construct: (ATFSSR) Story Reading Aloud to Children 11:31 Thursday, September 10, 1992 13

OBS	NAME	MEAN	N	STD
1	BEF01	-0.16163	140	0.40707
2	CAC01	0.50334	144	0.23604
3	CYP01	-0.09195	300	0.29066
4	DEN01	0.06153	193	0.34402
5	FIN01	0.05713	67	0.39734
6	FRA01	-0.49796	123	0.40731
7	GER01	-0.28288	137	0.36269
8	GER02	-0.26321	93	0.38848
9	GRC01	-0.08158	164	0.31318
10	HKO01	-0.26675	142	0.34674
11	HUN01	-0.34939	135	0.35528
12	ICE01	0.26035	265	0.33380
13	INS01	-0.17266	156	0.32232
14	IRE01	-0.01727	116	0.43990
15	ITA01	-0.07907	141	0.31890
16	NET01	-0.09023	92	0.38624
17	NEZ01	0.44660	170	0.25198
18	NOR01	0.39236	176	0.27907
19	POR01	0.17360	143	0.30979
20	SIN01	0.07620	202	0.35977
21	SPA01	0.07030	289	0.39908
22	SWE01	0.36639	223	0.30295
23	SWI01	-0.10792	216	0.32946
24	TRI01	-0.12885	193	0.31655
25	USA01	0.34247	287	0.35175
26	VEN01	-0.02135	116	0.31361
27	YOU01	-0.08623	133	0.30564

Univariates for the construct: (ATFSLE) Literature Emphasis
11:31 Thursday, September 10, 1992 14

OBS	_NAME_	MEAN	N	STD
1	BEF01	-0.20985	137	0.39250
2	CAC01	0.48222	140	0.37021
3	CYP01	0.08156	278	0.37077
4	DEN01	-0.32970	188	0.32333
5	FIN01	-0.02164	66	0.34120
6	FRA01	-0.16441	111	0.36091
7	GER01	-0.29323	119	0.32629
8	GER02	-0.20407	86	0.32496
9	GRC01	-0.19445	161	0.38600
10	HKO01	-0.19966	142	0.36393
11	HUN01	-0.00868	134	0.30185
12	ICE01	-0.13050	242	0.30373
13	INS01	0.43473	154	0.39880
14	IRE01	-0.11748	104	0.43449
15	ITA01	-0.24200	128	0.33387
16	NET01	-0.02410	94	0.25913
17	NEZ01	0.51805	166	0.31858
18	NOR01	-0.09746	162	0.31592
19	POR01	0.05688	125	0.46854
20	SIN01	0.13514	201	0.36950
21	SPA01	0.00665	241	0.40514
22	SWE01	0.10520	220	0.34474
23	SWI01	-0.26996	215	0.35079
24	TRI01	0.06766	177	0.36656
25	USA01	0.24734	277	0.35665
26	VEN01	0.11693	109	0.46062
27	YOU01	0.15392	126	0.29174

Univariates for the construct: (ATFSAT) Active Teaching of Comprehension
11:31 Thursday, September 10, 1992 15

OBS	NAME	MEAN	N	STD
1	BEF01	-0.16869	139	0.32485
2	CAC01	0.18221	118	0.32171
3	CYP01	0.30089	275	0.31354
4	DEN01	-0.24360	187	0.31292
5	FIN01	-0.21235	63	0.27425
6	FRA01	-0.12692	118	0.34591
7	GER01	-0.29425	140	0.26931
8	GER02	-0.05863	92	0.35726
9	GRC01	0.38193	165	0.32086
10	HKO01	-0.07792	147	0.38320
11	HUN01	0.13724	133	0.27310
12	ICE01	-0.46874	235	0.36116
13	INS01	0.35448	163	0.25739
14	IRE01	-0.11131	103	0.40118
15	ITA01	0.28445	127	0.36408
16	NET01	-0.47346	87	0.28685
17	NEZ01	-0.02282	150	0.35993
18	NOR01	-0.25712	162	0.30727
19	POR01	0.45348	122	0.30576
20	SIN01	0.08684	195	0.35884
21	SPA01	-0.05385	260	0.39457
22	SWE01	-0.34741	200	0.35698
23	SWI01	-0.29573	205	0.31262
24	TRI01	0.33611	175	0.30840
25	USA01	0.22663	272	0.30930
26	VEN01	0.30258	106	0.42167
27	YOU01	0.17865	120	0.26999

OBS	NAME	MEAN	N	STD
1	BEF01	-0.08918	145	0.38510
2	CAC01	-0.41228	145	0.30980
3	CYP01	0.33231	285	0.30891
4	DEN01	-0.45654	196	0.32942
5	FIN01	-0.21021	66	0.38941
6	FRA01	-0.06378	125	0.35849
7	GER01	0.01213	145	0.32465
8	GER02	0.17413	98	0.32860
9	GRC01	0.33987	167	0.32432
10	HKO01	0.34826	148	0.21593
11	HUN01	0.09373	141	0.28356
12	ICE01	-0.09237	268	0.38557
13	INS01	0.24526	167	0.28721
14	IRE01	-0.16719	114	0.40123
15	ITA01	0.22604	144	0.33347
16	NET01	-0.02289	92	0.26535
17	NEZ01	-0.26541	171	0.29946
18	NOR01	-0.37881	168	0.37516
19	POR01	0.40743	149	0.33740
20	SIN01	-0.01679	200	0.41580
21	SPA01	0.24974	291	0.30729
22	SWE01	-0.19547	218	0.37255
23	SWI01	-0.11389	224	0.36232
24	TRI01	-0.04149	210	0.40072
25	USA01	-0.30635	285	0.33774
26	VEN01	0.48849	120	0.35206
27	YOU01	-0.06699	127	0.38870

IEA-Reading Literacy Study:

Population A

Booklet A3

Student Questionnaire

International Coordinating Center
Hamburg 1990

Identification:	target pop.	
	country	
	language	
	stratum	
	school	
	class	
	teacher	
	student	
	seq. no.	

Reading Literacy:

Student Questionnaire Population A

Directions:

Please answer all of the following questions as best as you can. If you need help, ask me (NRC: this is read by the test administrator/teacher).

A. Questions about yourself and your home (Questions 1 to 18)

1 How old are you?

_____ years and _____ months

2 Are you a boy or a girl?
(Circle one only)

Boy..... 1
Girl..... 2

3 How often do you speak *[language of test]* at home?
(Circle one only)

Always 1
Almost always 2
Sometimes..... 3
Hardly ever..... 4
Never..... 5

4 How often do you eat each of the following meals?
(Circle one number on each line)

	Never	1 or 2 times a week	3 or 4 times a week	Every day
Breakfast	1	2	3	4
Lunch	1	2	3	4
Evening meal	1	2	3	4

5 Do you get a daily newspaper at home?
(Circle one only)

No 1
Yes..... 2

6 On a school day, about how many hours do you usually watch TV or video outside of school hours?
(Circle one only)

I do not watch TV or video 1
Up to 1 hour..... 2
Between 1 and 2 hours 3
Between 2 and 3 hours 4
Between 3 and 4 hours 5
Between 4 and 5 hours 6
More than 5 hours..... 7

7 About how many books are there in your home?
(Do not count newspapers or magazines; circle one only)

None 1
1 - 10 2
11 - 50..... 3
51 - 100 4
101 - 200..... 5
More than 200.... 6

8 Check how many of the following things can be found in your home.
(Be sure to check the corresponding box if you have none. Check all that are found in your home. Check ✓ one box per line only)
[NRCs: Fill in the items.]

		None	One or more
a)	Item 1	☐	☐
b)	Item 2	☐	☐
c)	Item 3	☐	☐
d)	Item 4	☐	☐
e)	Item 5	☐	☐
f)	Item 6	☐	☐
g)	Item 7	☐	☐
h)	Item 8	☐	☐
i)	Item 9	☐	☐
j)	Item 10	☐	☐
		(1)	(2)

9 Please say whether you have the following things or not.
(Check all that you have; check one box per line only)
[NRCs: Fill in the items.]

		None	One or more
a)	Item 1	☐	☐
b)	Item 2	☐	☐
c)	Item 3	☐	☐
d)	Item 4	☐	☐
e)	Item 5	☐	☐
f)	Item 6	☐	☐
g)	Item 7	☐	☐
h)	Item 8	☐	☐
		(1)	(2)

10 How often do people at home read to you in *[language of test]*?
(Circle one only)

Never..... 1
1 or 2 times a week 2
3 or 4 times a week 3
Nearly every day..... 4

11 Is there any other place outside of school and your home where someone reads to you in *[language of test]*?
(Circle one only)

No 1
Yes..... 2

12 How often do people at home read to you in another language?
(Circle one only)

Never..... 1
1 or 2 times a week 2
3 or 4 times a week 3
Nearly every day..... 4

13 Is there any other place outside of school and your home where someone reads to you in another language?
(Circle one only)

No 1
Yes..... 2

B. Questions about your reading (Questions 14 to 18)

14 How often do you borrow books from a school or public library?
(Circle one only)

Never..... 1
Hardly ever..... 2
Once a month..... 3
Once a week..... 4
More than once a week 5

15 How good are you at reading?
(Circle one only)

Not very good 1
Average..... 2
Good..... 3
Very good..... 4

16 Which do you think are the three most important ways to become a good reader?
(Choose three only)

		MOST IMPORTANT
a)	Liking it	2
b)	Having lots of time to read	2
c)	Being able to concentrate well	2
d)	Knowing how to sound out words	2
e)	Learning the meaning of lots of words	2
f)	Having many good books around	2
g)	Having a lively imagination	2
h)	Having lots of reading for homework	2
i)	Having lots of drill (practice) at hard things	2
j)	Having lots of written exercises	2
k)	Being told how to do it	2

17 How often do you read in *[language of test]* to someone at home?
(Circle one only)

Never..... 1
1 or 2 times a week 2
3 or 4 times a week 3
Nearly every day..... 4

18 How often do your parents or other people at home ask you what you have been reading?
(Circle one only)

Never..... 1
1 or 2 times a week 2
3 or 4 times a week 3
Nearly every day..... 4

C. Questions about your reading homework and your classroom work (Questions 19 to 24)

19 How often do you get reading homework?

Never..... 1
1 or 2 times a week 2
3 or 4 times a week 3
Every day..... 4

20 About how much time do you spend on your reading homework?
(Circle one only)

None 1
Up to a quarter of an hour 2
Up to half an hour 3
More than half an hour 4

21 How often are you asked questions in class about your reading homework?
(Circle one only)

I do not get reading homework 1
Always 2
Most of the time 3
Sometimes 4
Hardly ever 5
Never 6

22 How often are you helped with your reading homework?
(Circle one only)

I do not get reading homework 1
I rarely get help 2
I sometimes get help 3
I get help most of the time 4

23 If you don't finish the reading work you get to do by yourself in class, are you expected to finish it in your own time?
(Circle one only)

Always 1
Most of the time 2
Sometimes 3
Hardly ever 4
Never 5
I do not get reading work to do by myself in class 6

24 How often are you given written work about the reading work that you have been given?
(Circle one only)

Always 1
Most of the time 2
Sometimes 3
Hardly ever 4
Never 5

D. Reading for Enjoyment (Questions 25 to 37)

25 Did you read a book for fun last week? (☞ *see accompanying note*)
(Circle only one)

No 1
Yes..... 2

(If 'Yes', write in the title or author of the book)

Book title/author:_____

26 How often do you read books for fun?
(Circle only one)

Almost never..... 1
About once a month..... 2
About once a week 3
Almost every day..... 4

27 Did you read a comic book last week?
(Circle only one)

No 1
Yes..... 2

(If 'Yes', write in the title or the person in the story)

Comic book title/person:_____

28 How often do you read comic books?
(Circle only one)

Almost never..... 1
About once a month..... 2
About once a week 3
Almost every day..... 4

29 Did you read a magazine last week?
(Circle only one)

No 1
Yes..... 2

(If 'Yes', write in the title of the magazine or the topic you read about)

Magazine title/topic:_____

30 How often do you read magazines?
(Circle only one)

Almost never..... 1
About once a month..... 2
About once a week 3
Almost every day..... 4

31 Did you read a newspaper last week?
(Circle only one)

No 1
Yes..... 2

(If 'Yes', write in the name of the newspaper)

Newspaper name:_____

32 How often do you read newspapers?
(Circle only one)

Almost never..... 1
About once a month..... 2
About once a week 3
Almost every day..... 4

33 How often do you read written directions or instructions to do something you enjoy?
(You might read them to put a toy together, follow a recipe, to use a tool or to do something else. Circle only one.)

Almost never..... 1
About once a month..... 2
About once a week 3
Almost every day..... 4

34 Do you read aloud at home?
(Circle only one)

No 1
Yes..... 2

35 How often do you read aloud to someone at home?
(Circle only one)

Never. I do not read aloud to someone at home..... 1
Less than 1 time per week..... 2
1 to 3 times per week..... 3
Nearly every day..... 4

36 To whom do you read aloud at home?
(Circle only one)

No one. I do not read aloud at home..... 1
Parents..... 2
Brother or sister..... 3
Other person..... 4

37 What do you read aloud at home?
(You may circle more than one)

Nothing. I do not read aloud at home 2
Newspaper..... 2
Magazine..... 2
Book..... 2
Textbook..... 2
Comic book..... 2
Letters..... 2
Words on television screens..... 2

E. Reading in School (Questions 38 to 43)

38 In school, how often do you read textbooks in reading or language class?
(Circle only one)

Almost never 1
About once a month..... 2
About once a week 3
Almost every day..... 4

39 How often do you read story books in addition to your textbooks in reading or language class?
(Circle only one)

Almost never 1
About once a month..... 2
About once a week 3
Almost every day..... 4

40 How often do you use workbooks or practice exercises in reading or language class?
(Circle only one)

Almost never 1
About once a month..... 2
About once a week 3
Almost every day..... 4

41 In school, how often do you read textbooks or practice exercises in science, geography or environmental studies?
(Circle only one)

Almost never 1
About once a month..... 2
About once a week 3
Almost every day..... 4

42 How often do you look up information in books like encyclopedias, dictionaries, manuals or maps for schoolwork?
(Circle only one)

Almost never 1
About once a month..... 2
About once a week 3
Almost every day..... 4

43 Do you have a favourite book? If so, fill in the title below.

_____

Thank you very much for your cooperation

Reading Literacy:

Student Questionnaire Population A Accompanying Notes

Questions 1 (Age):
Age may be collected in a way which is most appropriate for each country, but must be reported in years and months.

Questions 8 and 9 (Possessions):
It is for NRCs to decide on the items they will use. From the univariates in the pilot testing it is easy to see which items were not discriminating (i.e. one category contains more than 90% of the responses) and these should be changed. There should be ten items for Question 8 and eight items for Question 9 (see also Table 2 in 'Univariates for Background Questionnaires' - doc. ref. RL/ALL/90.436).

Question 14 (Borrow books from a library):
This means any library (e.g. school, town, mobile library).

Question 19-22 (Reading Homework):
Reading homework includes any reading work set to be done at home to improve reading or language skills in the language of the test.

Question 24 (Written Work):
Written work includes answering set questions, writing book reports, completing work sheets - in fact, any kind of written response.

Questions 25 - 43:
These questions were formerly in the Reading Activity Inventory. The Student Questionnaire is considerably shorter and administration time has been reduced by combining these two instruments. A number of NRCs felt that this step would be desirable.

Question 25:
Students need not have completed a book in one week.

IEA-Reading Literacy Study:

Population A
Teacher Questionnaire

International Coordinating Center
Hamburg 1990

Identification:	target pop.	
	country	
	language	
	stratum	
	school	
	class	
	teacher	
	seq. no.	

Reading Literacy:

Teacher Questionnaire

Population A

✎ Directions:

The following questionnaire is part of an international study of Reading Literacy and attempts to identify differences in reading instruction. It is recognized that teachers are likely to respond quite differently.
Please answer all questions in such a way as to reflect most clearly your teaching practice. Most questions require you to circle your selected response. Others require you to write in a number. Where it is appropriate to enter '0' in the answer, please do so. Do not leave blanks.
We thank you for your effort.

A. The first set of questions has to do with you and your educational training (Questions 1 to 9)

1 Your sex:
(Circle one number only)

Male..... 1
Female..... 2

2 Is your mother-tongue the same as *[language of the test]*?
(Circle one number only)

No..... 1
Yes..... 2

3 How many years of primary and secondary school education did you have altogether?
(If you have had no school education, please enter '0'.
Do not include pre-compulsory education e.g. Kindergarten. Also do not count grade repetition years.)
[NRCs: see accompanying note.]

_____ years (or full-time years equivalent to nearest whole number) ___

4 How many years of all your education was pre-service teacher-training?
(If you have had no pre-service teacher training, please enter '0'.)
[NRCs: see accompanying note.]

_____ years ___

5 How many years of post-secondary education did you have?
(If you have had no post-secondary education, please enter '0'.
Do not include pre-service teacher training and grade repetition.)
[NRCs: see accompanying note.]

_____ years (or full-time years equivalent to nearest whole number) ___

6 Approximately how many hours have you devoted to the further study of the teaching of reading after your initial teacher qualifications?
(Circle one number only)

None	1
Less than 10 hours	2
10 to 29 hours	3
30 to 49 hours	4
50 to 100 hours	5
More than 100 hours	6

___

7 How many times have you been to in- or on-service teacher training course in reading in the last three years?
(Circle one number only)

None	1
Once	2
Twice	3
Three times	4
Four or more times	5

___

8 About how often do you read each of the following?
(Do not include reading for preparation of class lessons.
Circle one number per line only)

		never or almost never	once a year	about once a term	about once a month	about once a week or more
a)	Articles on teaching	1	2	3	4	5
b)	Articles on reading	1	2	3	4	5
c)	Books on history or politics	1	2	3	4	5
d)	Books on the arts	1	2	3	4	5
e)	Books on science	1	2	3	4	5
f)	Novels or short stories	1	2	3	4	5
g)	Poems	1	2	3	4	5
h)	Plays	1	2	3	4	5
i)	Books for children	1	2	3	4	5

9 By the end of this school year how many years will you have been teaching altogether?

_____ years (or full-time years equivalent)

B. This set of questions has to do with your class being tested (Questions 10 to 19)

10 How long have you been teaching the class being tested?
(Circle one number only)

Less than half a year..... 1
Between half a year and one year..... 2
Between one year and two years..... 3
Between two years and three years..... 4
More than three years 5

11 Is the class tested a multi-grade class?
(Circle one number only)

No 1
Yes..... 2

12 What is the total number of students and the total number of *[grade level being tested]* students in this class?
[NRCs: see accompanying note.]

_____ total students _____ *[grade level being tested]* students

13 How many *[grade level being tested]* students in this class do not have *[language of test]* as their first language?
(If none, please enter '0'.)

_____ students

14 As a general rule, how often do you meet with parents of the students in the class tested?
(Circle one number only)

Never..... 1
Once a year 2
Once a term..... 3
Once a month or more..... 4

15 How many students in this class need remedial help in reading?
(If none, please enter '0')

_____ students

16 How many students in this class receive remedial help in reading?
(If none, please enter '0')

_____ students

17 What is the number of hours and minutes of total instructional time excluding breaks for this class in a typical week? *(For all subject areas) (see explanation in the accompanying notes)*

_____ hours and _____ minutes per week

18 How much class time per school week do you typically devote to the teaching and learning of *[language of test]* including reading, writing, speaking, literature, listening, and other language skills for this class?

_____ hours and _____ minutes per week

19 How much class time per school week do you typically devote to the teaching and practice of reading in *[language of test]* for this class?

_____ hours and _____ minutes per week

C. The following set of questions has to do with your teaching (Questions 20 to 34)

20 How often are your students typically involved in the following reading activities?
(Circle one number per line only)

	Reading Activities	Frequency			
		almost never	about once a month	about 1 or 2 times a week	almost every day
a)	Learning letter-sound relationships and/or phonics	1	2	3	4
b)	Word-attack skills (e.g. prediction)	1	2	3	4
c)	Silent reading in class	1	2	3	4
d)	Answering reading comprehension exercises in writing	1	2	3	4
e)	Independent silent reading in a library	1	2	3	4
f)	Listening to students reading aloud to a whole class	1	2	3	4
g)	Listening to students reading aloud to small groups or pairs	1	2	3	4
h)	Listening to teachers reading stories aloud	1	2	3	4
i)	Discussion of books read by students	1	2	3	4
j)	Learning new vocabulary systematically *(e.g. from lists)*	1	2	3	4
k)	Learning new vocabulary from texts	1	2	3	4
l)	Learning library skills	1	2	3	4
m)	Reading plays or dramas	1	2	3	4
n)	Playing reading games (e.g. forming sentences from jumbled words)	1	2	3	4
o)	Dramatizing stories	1	2	3	4
p)	Drawing in response to reading	1	2	3	4
q)	Orally summarizing their reading	1	2	3	4
r)	Relating experiences to reading	1	2	3	4
s)	Reading other students' writing	1	2	3	4
t)	Making predictions during reading	1	2	3	4
u)	Diagramming story content	1	2	3	4
v)	Looking for the theme or message	1	2	3	4
w)	Making generalizations and inferences	1	2	3	4
x)	Studying the style or structure of a text	1	2	3	4
y)	Comparing pictures and stories	1	2	3	4
z)	Student leading discussion about passage	1	2	3	4
aa)	Reading in other subject areas	1	2	3	4
bb)	Writing in response to reading	1	2	3	4

21 In a normal reading instruction period how many reading textbooks are available for each student in the class tested?
(Circle only the most appropriate answer)

None	1
1 book for about 5 or more students	2
1 book for each 2 students	3
1 book for each student	4
2 books for each student	5
3 or more books for each student	6

22 Please rank only five of the following aims of reading instruction in order of the importance you attach to each of them.
(Place '1' next to the most important and so on to '5' for the least important. Use all 5 ranks once only.)

	Aims	Importance
a)	Developing skill in reading aloud	
b)	Developing a lasting interest in reading	
c)	Improving students' reading comprehension	
d)	Developing students' research and study skills	
e)	Extending students' vocabulary	
f)	Developing students' critical thinking	
g)	Expanding students' world views	
h)	Deepening students' emotional development	
i)	Improving word-attack skills	
j)	Increasing speed of reading	
k)	Expanding students' reading choice	
l)	Making reading enjoyable	

23 How often do you use the following instructional strategies when teaching reading?
(Circle one number per line only)

Instructional Strategies		Frequency			
		almost never	about once a month	about 1 or 2 times a week	almost every day
a)	Introduce the background of a passage before reading it	1	2	3	4
b)	Ask children to describe their strategy for understanding	1	2	3	4
c)	Encourage parents to be involved with the reading program	1	2	3	4
d)	Maintaining a graded sequence of text difficulty	1	2	3	4
e)	Ask questions to assess text comprehension	1	2	3	4
f)	Ask questions to deepen understanding	1	2	3	4
g)	Show children how to understand a text	1	2	3	4
h)	Compare stories, poems, fables and tales	1	2	3	4
i)	Read aloud to children	1	2	3	4
j)	Encourage parents to read to children	1	2	3	4
k)	Encourage the children to read more	1	2	3	4
l)	Encourage children to use the library more	1	2	3	4
m)	Use materials you have prepared yourself	1	2	3	4

24 Do you divide the students in this class into groups for reading instruction?

No 1
Yes.....2

If you answered 'No' to this question, go straight to Question 27.

25 What type of grouping do you use most often?
(Circle one only)

Do not use grouping..... 1
Age groups 2
Ability groups..... 3
Interest groups..... 4
Other *(please specify)* 5

26 How many groups do you typically form?

_____ groups

27 How frequently did you teach in your class this year how to read each of the following kinds of text?
(Circle one number per line only)

	Frequency					
	almost never	3 or 4 times a year	about once a month	at least once a week	nearly every day	
a) Narration: texts that tell a story or give the order in which things happen	1	2	3	4	5	___
b) Exposition: texts that describe things or people or explain how things work or why things happened	1	2	3	4	5	___
c) Documents: tables, charts, diagrams, lists, maps	1	2	3	4	5	___

28 Below you will find a number of statements about which we ask you to give your views with respect to issues in reading instruction.
(Please mark for each statement your degree of agreement/disagreement by circling the appropriate number. Circle one number in each line.)

		strongly disagree	disagree	uncertain	agree	strongly agree	
1.	When my pupils read to me, I expect them to read every word accurately.	1	2	3	4	5	
2.	Teachers should keep careful records of every child's reading progress.	1	2	3	4	5	
3.	Children should not be encouraged to read a word they don't know.	1	2	3	4	5	
4.	All children should enjoy reading.	1	2	3	4	5	
5.	Most of what a child reads should be assessed.	1	2	3	4	5	
6.	Every day children should be read to by the teacher from a story book.	1	2	3	4	5	
7.	Reading aloud by children to a class is a waste of time.	1	2	3	4	5	
8.	Most children improve their reading best by extensive reading on their own.	1	2	3	4	5	
9.	Children should always understand why they are reading.	1	2	3	4	5	
10.	Teachers should always group children, according to their reading ability.	1	2	3	4	5	
11.	9-year-olds should not have access to books they will read in the next year at school.	1	2	3	4	5	
12.	Class sets of graded reading material should be used as the basis for the reading programme.	1	2	3	4	5	
13.	Children who can't understand what they read haven't been taught proper comprehension skills.	1	2	3	4	5	
14.	Every mistake a child makes in reading aloud should be corrected at once.	1	2	3	4	5	
15.	All children's comprehension assignments should be marked carefully to provide them with feedback.	1	2	3	4	5	
16.	Children should not start a new book until they have finished the last.	1	2	3	4	5	
17.	Parents should be actively encouraged to help their children with reading.	1	2	3	4	5	
18.	Children should learn most of their new words from lessons designed to enhance their vocabulary.	1	2	3	4	5	
19.	Reading learning materials should be carefully sequenced in terms of language structures and vocabulary.	1	2	3	4	5	
20.	Children should take a book home to read every day.	1	2	3	4	5	
21.	Children should be encouraged to read texts they have written.	1	2	3	4	5	
22.	Children should always understand what they are reading.	1	2	3	4	5	
23.	Children should always choose their own books to read.	1	2	3	4	5	
24.	A word recognition test is sufficient for assessing children's reading levels.	1	2	3	4	5	
25.	Teachers should carefully follow the sequence of the textbook	1	2	3	4	5	
26.	Children should undertake research projects to improve their reading	1	2	3	4	5	

29 What do you regularly do (i.e. at least once a week) to encourage your students to read outside school?
(You may circle more than one number)

a)	Suggest books to students to read	2
b)	Suggest newspaper articles to students to read.....	2
c)	Read attractive stories to students.....	2
d)	Hold discussions about books	2
e)	Other	2

30 How often do you use the following methods to discover your students' needs in reading?
(Circle one number per line only)

		never or almost never	once a year	about once a term	about once a month	about once a week or more
a)	Listening to students' reading	1	2	3	4	5
b)	Teacher-made vocabulary tests	1	2	3	4	5
c)	Exercises in workbooks and textbooks	1	2	3	4	5
d)	Standardized or formal tests of comprehension	1	2	3	4	5
e)	Knowledge of students' reading interests	1	2	3	4	5
f)	Comments from other teachers	1	2	3	4	5
g)	Informal observation	1	2	3	4	5
h)	Interviews	1	2	3	4	5
i)	Tests in workbooks and text-books	1	2	3	4	5

31 How often do you assess these aspects of reading with all or most of your class?
(Circle one number per line only)

		never or almost never	once a year	about once a term	about once a month	about once a week or more
a)	Word recognition	1	2	3	4	5
b)	Vocabulary	1	2	3	4	5
c)	Text comprehension	1	2	3	4	5
d)	Literary appreciation	1	2	3	4	5
e)	Use of background knowledge	1	2	3	4	5
f)	Sentence understanding	1	2	3	4	5
g)	Phonic skills	1	2	3	4	5
h)	Reading study skills	1	2	3	4	5
i)	Amount of reading	1	2	3	4	5
j)	Decoding	1	2	3	4	5

32 How often do you use these assessment methods?
(Circle one number per line only)

		never or almost never	once or twice a year	about once a term	about once a month	about once a week or more
a)	Multiple-choice questions of reading	1	2	3	4	5
b)	Listening to students reading aloud	1	2	3	4	5
c)	Records of student interests	1	2	3	4	5
d)	Oral discussions	1	2	3	4	5
e)	Oral questions on material read	1	2	3	4	5
f)	Written open-ended questions on material read	1	2	3	4	5

33 Do you assign homework in reading to the class tested?
(Circle one number only)

No 1
Yes 2

If you answered 'No' to Question 33, go straight to Question 36

34 How often do you ask children to read something at home as part of your reading/language program?
(Circle one number only)

Never 1
Less than once a week 2
1 or 2 times a week 3
3 or 4 times a week 4
More than 4 times a week 5

35 About how many minutes do you expect an average student to spend on reading homework when you assign it?
(Circle one number only)

None 1
Up to 10 minutes 2
11 - 20 minutes 3
21 - 30 minutes 4
31 - 40 minutes 5
41 - 50 minutes 6
More than 50 minutes 7

D. The following set of questions has to do with your classroom library (Questions 36 to 39)

36 Do you have a classroom library (i.e. a small book or magazine corner in your classroom)?
(Circle one number only)

No 1
Yes..... 2

⇨ If you answered 'No' to Question 36, please go to Question 40.

37 About how many books with different titles does your classroom library contain?
(Circle one number only)

Less than 20..... 1
21-40 2
41-60 3
61-80 4
81-100 5
More than 100.... 6

38 About how many different titles of magazines and/or newspapers do you have in your classroom library?
(Circle one number only)

None..... 1
1-5 2
6-10..... 3
11-15 4
16-20 5
More than 20..... 6

39 Can your students borrow books from the classroom library to take home?
(Circle one number only)

No 1
Yes..... 2

E. The following set of questions is about your school library (Questions 40 to 42)

40 Do you have a school library in your school?
(Circle one number only)

No 1
Yes..... 2

➪ If you answered 'No' to Question 40, then go to Question 43.

41 How often do your students visit the school library as a class?
(Circle one number only)

Hardly ever 1
Once a month 2
Once a week 3
More than once a week 4

42 Can your students borrow books from the school library to take home?
(Circle one number only)

No 1
Yes..... 2

F. The last set of questions are to do with school organisation (Questions 43 to 46)

43 Is your work as a teacher evaluated by the school principal (or deputy school principal)?

No 1
Yes.....2

44 Does the school principal (or deputy principal)...
(Check the appropriate answers.)

		No	Yes
a)	discuss with you explicit achievement standards for the subject that you teach	1	2
b)	ask for evaluation results or progress of your students in reading	1	2
c)	make suggestions about the choice of instructional methods in reading	1	2
d)	encourage contacts among teachers	1	2
e)	initiate activities directed at the professional development of teachers	1	2
f)	make suggestions about the content that must be covered in reading	1	2

45 How often do you have staff meetings at your school?
(Check one only.)

Never.....1
Once a year.....2
Once a term.....3
Monthly 4
Weekly 5

46 If you have staff meetings, please indicate how often the following items occur as subjects of discussion during staff meetings.

		all staff meetings	most staff meetings	some staff meetings	not in any staff meetings
a)	curriculum content	1	2	3	4
b)	the way the subject matter is presented	1	2	3	4
c)	professional development of teachers	1	2	3	4
d)	issues of 'pastoral care' (e.g. student problems, guidance, welfare)	1	2	3	4
e)	organizational issues (e.g. school climate, co-ordination of work among teachers, the way decision-making procedures are conducted)	1	2	3	4
f)	other topics (e.g. purely administrative tasks, leisure and social activities)	1	2	3	4

Thank you very much for your cooperation

Reading Literacy:

Teacher Questionnaire Population A Accompanying Notes

Question 3:
The intention of this question is to collect years of school education and exclude pre-compulsory education. NRCs should ask the question in such a way that it is absolutely clear what is meant. In the pilot study some systems with 12 years of primary and secondary education were submitting this variable with an average of 14 to 17 years of primary education! Please ensure that this does not happen for the main study. NRCs must supply valid range information.

Questions 4 and 5:
The intention of these questions is to capture all post-secondary education and to be able to distinguish between teacher training and other forms of post-secondary (e.g. university) education.

Question 5:
The intention is to capture the duration of teacher training.

Question 6:
The intention is to capture systematic pursuit of studying after teacher training.

Question 7:
NRCs must check the accuracy of replies in terms of how their system works.

Question 12:
Note that it is important to collect the total number of students and the number of students at the grade level being tested. Note that in most cases the number will be the same. Only where the class is multi-grade will the numbers be different.

Questions 12 and 13:
We are interested in the percentage in the grade level being tested.

Question 14:
'Parents' include guardians and could be any parent.

Question 17:
Note that this is a check question against Question 14 from the School Questionnaire. 'Total instructional time' here includes all subject areas.

Question 20:
Note that there are minor changes to several options and some new options have been added.

Question 20 b:
'Word-attack skills' means vocabulary decoding.

Question 22i:
In this case 'word attack' equals vocabulary decoding.

Question 33-35:
'Reading homework' includes any work assignment related to reading and language programmes.

Question 37:
Derived variable: classroom library books per student.

Questions 43-46:
These are the OECD questions and are concerned with school leadership.

IEA-Reading Literacy Study:

Population A
School Questionnaire

International Coordinating Center
Hamburg 1990

Identification:	target pop.	
	country	
	language	
	stratum	
	school	
	class	
	seq. no.	

Reading Literacy:

School Questionnaire

Population A

Directions:

The following questionnaire is part of an international study of Reading Literacy. The questions asked attempt to gather information which captures the wide range of experience and practice which is likely to exist across countries.
Please answer all questions in such a way as to reflect most accurately the situation in your school. It is important that all questions are answered. Most questions require you to circle your selected response. Others require you to write in a number. Where it is appropriate to enter '0' in the answer, please do so. Do not leave it blank. All information will be treated in the strictest confidence..

1 How many years will you have been a school principal by the end of this school year?

in your total career _____ years

in your present school _____ years

2 What is the total enrollment of full-time students in your school?
(If there are no boys or no girls, please enter '0')

Number of boys _____
Number of girls _____

3 What is the total enrollment of full-time *[grade level being tested]* students in your school? (☞ *see note*)
(If there are no boys or girls, please enter '0')

Number of boys _____
Number of girls _____

4 Is your school... (☞ *see note*)

A state school..... 1
A private school..... 2

5 What is the type of community served by your school? (☞ *see note*)
(Circle one only)

A village or rural community 1
A small town community..... 2
A large town community..... 3
A city of 1 million or more..... 4

6 Please indicate the availability of the following resources in relation to your school?
(Circle one number on each line)

	Not readily available	Available in neighbouring town or city (less than 2 hours of normal one way travel time in)	Available locally (within 30 minutes of normal one way travel time in)
Public Library	1	2	3
Bookstore/book department store	1	2	3
Secondary level schools	1	2	3
A higher education institution	1	2	3

7 What is the degree of parent co-operation with the school in terms of support for the schools educational principles or goals (compared with other schools you know)?
(Circle one only)

Much below average 1
Below average..... 2
Average 3
Above average..... 4
Much above average 5

8 Which of the following resources and activities are there in your school?
(Circle one number on each line)

	No	Yes	
School library	1	2	___
Reading room for students	1	2	___
Student/school newspaper or magazine	1	2	___
Teacher (Professional) library	1	2	___

☞ The following questions are about the school library.
If you have indicated in Question 8 that your school does not have a library, please go straight to Question 12.

9 Approximately how many books with different titles does your school library contain?
(Exclude magazines and periodicals.)

_____ different titles ___

10 Approximately how many books with different titles were added to your school library in the last year?
(Exclude magazines and periodicals.)

_____ books with different titles ___

11 Can *[grade level being tested]* students in your school borrow books from the school library to take home?
(Circle one only)

No 1
Yes ..... 2 ___

12 How many full-time (or full-time equivalent) teaching teachers are there in your school? (☞ *see note*)
(Exclude non-teaching principal and administrators. If there are no male or no female special teachers, please enter '0'. For full-time equivalent, add the number of part-time teachers. For example, two half-time teachers equal one full-time equivalent teacher. Or, three third-time teachers equal one full-time equivalent teacher. Round to nearest whole number.)

_____ male teachers ___
_____ female teachers ___

13 How many full-time (or full-time equivalent to the nearest whole number) special teachers are there in your school (i.e. for remedial teaching, counselling, guidance. Exclude, for example, librarians and physical education teachers)?
(If there are no male or no female special teachers, please enter '0'.)

_____ male teachers
_____ female teachers

14 What is the number of hours and minutes of total instructional time *[see explanation in the accompanying notes]* excluding breaks *[for the class being tested]* in a typical week in your school for all subject areas? (☞ *see note*)

_____ hours and _____ minutes per week

15 How many weeks per year is your school open for the class being tested? (☞ *see note*)

_____ weeks per year

16 How many days of instruction were lost in the last school year (due to accidents, floods, strikes, festivals, staff days, etc.?)
(If there were no days lost, please enter '0')

_____ days lost in the last school year

17 On an average day, approximately what percentage of students are absent from school? (☞ *see note*)

_____ percent

18 Does your school have any special programs or teacher initiatives for reading outside the normal classroom activities (e.g. enrichment programs, or special programs for the disadvantaged)?
(You may circle more than one)

Extra class lessons in reading..... 2
Extra individual tuition at school..... 2
Special remedial reading courses..... 2
Other..... 2

19 Does your school sponsor any informal initiatives to encourage reading (e.g. book clubs)?

No 1
Yes..... 2

20 Do you have a program for the improvement of reading instruction (teaching and learning of reading) in your school? (☞ *see note*)

No 1
Yes..... 2

21 What are the two most serious problems which you experience in providing for the teaching and learning of reading in your school? (☞ *see note*)
(Circle two problems only)

Insufficient specialized staff 2
Insufficient time 2
Students' lack of interest 2
Insufficient classroom material 2
Other..... 2
None - there are no serious problems 2

22 Please rank the following activities in order of importance in your work as a school principal.
('1' is the most important activity, '8' is the least important activity, 'NA'= not applicable. Do not assign equal rankings.)

		rank of importance
a)	representing the school at official meetings	
b)	evaluation of staff	
c)	contacts with local community (e.g. parents, community organizations, local industry)	
d)	discussing educational objectives with the teaching staff	
e)	administrative tasks concerning the functioning of the school (e.g. regulations, disciplinary duties, school budget, timetable)	
f)	using records of pupils' progress	
g)	taking care of issues of 'pastoral care' (e.g. student problems, guidance, welfare)	
h)	activities aimed at the professional development of teachers	

23 How often do you systematically evaluate the pedagogical work of the teachers at your school?
(Check one only.)

Never.....1
Less than once a year.....2
About once a year.....3
More than once a year.....4

24 What procedures do you use to gather information for your evaluation?
(You may choose more than one alternative.)

	Yes
I do not systematically evaluate	2
Interviews	2
Written or oral self reports by teachers	2
Observational data on teachers' classroom work	2
Student ratings of teachers' performance	2
Other forms of systematic evaluation	2

Thank you very much for your cooperation

Reading Literacy:

School Questionnaire Population A Accompanying Notes

Question 3:

Two derived variables will be produced:

1. Total enrolment;
2. Sex of school.

Question 4:

State/Private Schools:

The purpose of this question is to compare state and private schools. At issue here is control of the school, rather than its funding. NRCs may have to substitute an appropriate term for 'state' in their own countries.

Question 5:

The basic aim of this question is to have a rural-urban scale for where schools are situated. The relative size of the types of community mentioned in the queston will vary from country to country. It is up to each country to phrase the options in such a way that the dimension from rural to large urban is clear. The result must be a four-point scale. If a dichotomy of urban/rural is to be made it will be category 1 for rural and categories 2, 3, and 4 for urban.

Question 12:

Principals will have to convert part-time to full-time teacher equivalent rounded to the nearest whole number. NRCs to make sure that this is clear to the respondents.

Two derived variables will be produced:

1. Total number of full-time teaching staff;
2. Pupil teacher ratio.

Question 13:

The intention is to capture remedial and/or specialist resources for teaching programmes excluding librarians and physical education teachers. Two derived variables will be produced:

1. Total number of special teachers;
2. Ratio of pupils to special teachers.

Question 14:
In the pilot test some countries did not make clear that the question targeted total instructional time in a typical week. NRCs need to specify valid ranges.

Question 15:
Note in the pilot study values ranged from 0 - 52 weeks. NRCs need to specify valid ranges.

Question 16:
Maximum values in the pilot study seemed very unlikely in some countries (e.g. '91'). NRCs to specify valid ranges.

Question 17:
In the pilot study some maximum values were not likely (98 percent and 165 percent). NRCs to supply valid ranges.

Questions 20 and 21:
The teaching and learning of reading includes any program aimed at the improvement of *[language of test]* skills.

Questions 22-24:
These are the OECD questions and are concerned with school leadership.

Appendix B

Population B Questionnaires

Student, Teacher, and School

IEA-Reading Literacy Study:

Population B

Booklet B3

Student Questionnaire

International Coordinating Center
Hamburg 1990

Identification:	target pop.	
	country	
	language	
	stratum	
	school	
	class	
	teacher	
	student	
	seq. no.	

Reading Literacy:

Student Questionnaire Population B

✎ Directions:

Please answer all of the following questions as best as you can. If you need help, ask me (NRC: this is read by the test administrator/teacher).

A. Questions about yourself and your home (Questions 1 to 15)

1 How old are you?

_____ years and _____ months

2 Are you a boy or a girl?
(Circle one only)

Boy..... 1
Girl..... 2

3 How often do you speak *[language of test]* at home?
(Circle one only)

Always 1
Almost always 2
Sometimes..... 3
Hardly ever..... 4
Never..... 5

4 How many years of education did your father or male person responsible for your education receive? (☞ *see note*)

_____ years

5 How many years of education did your mother or female person responsible of your education receive? (☞ *see note*)

_____ years

6 How often do you eat each of the following meals?
(Circle one number on each line)

	Never	1 or 2 times a week	3 or 4 times a week	Every day
Breakfast	1	2	3	4
Lunch	1	2	3	4
Evening meal	1	2	3	4

7 Do you have a job or regular (almost every day) family responsibilities (e.g. helping with housework or family business) before or after school?
(Circle one only)

No 1
Yes..... 2

8 How much time does your job or regular family responsibility take each day?
(Circle one only)

I do not have a job or any family responsibilities..... 1
Up to half an hour 2
Between half an hour and one hour..... 3
Between one and two hours 4
More than two hours 5

9 How many years of further education do you expect to complete after this year?
(Circle one only)

0 years	1
1 or 2 years	2
3 or 4 years	3
5 or 6 years	4
7 or 8 years	5
9 or 10 years	6
More than 10 years	7

10 Do you get a daily newspaper at home?
(Circle one only)

No	1
Yes	2

11 On a school day, about how many hours do you usually watch TV or video outside of school hours?
(Circle one only)

I do not watch TV or video	1
Up to 1 hour	2
Between 1 and 2 hours	3
Between 2 and 3 hours	4
Between 3 and 4 hours	5
Between 4 and 5 hours	6
More than 5 hours	7

12 About how many books are there in your home?
(Do not count newspapers or magazines; circle one only)

None	1
1 - 10	2
11 - 50	3
51 - 100	4
101 - 200	5
More than 200	6

13 Check how many of the following things can be found in your home?
(Be sure to check the corresponding box if you have none. Check all that you have in your home. Check ✓ one box per line only)
[NRCs: Fill in the items.]

		None	One or more
a)	Item 1	☐	☐
b)	Item 2	☐	☐
c)	Item 3	☐	☐
d)	Item 4	☐	☐
e)	Item 5	☐	☐
f)	Item 6	☐	☐
g)	Item 7	☐	☐
h)	Item 8	☐	☐
i)	Item 9	☐	☐
j)	Item 10	☐	☐
		(1)	(2)

14 Please say whether you have the following things or not
(Check all that you have; check one box per line only)
[NRCs: Fill in the items.]

		None	One or more
a)	Item 1	☐	☐
b)	Item 2	☐	☐
c)	Item 3	☐	☐
d)	Item 4	☐	☐
e)	Item 5	☐	☐
f)	Item 6	☐	☐
g)	Item 7	☐	☐
h)	Item 8	☐	☐
		(1)	(2)

15 How often do your parents or other people at home ask you about what you are reading?
(Circle one only)

Never 1
1 or 2 times a week 2
3 or 4 times a week 3
Nearly every day 4

B. Questions about your homework (Questions 16 to 21)

16 How often are you given homework?
(Circle one only)

Never 1
1 or 2 days per week 2
3 or 4 days per week 3
5 days or more per week 4

17 About how much time do you spend on homework when it is given?
(Circle one only)

None 1
Up to 1 hour 2
Between 1 hour and 2 hours 3
Between 2 hours and 3 hours 4
Between 3 hours and 4 hours 5
Between 4 hours and 5 hours 6
More than 5 hours 7

18 How often are you given *[language of test]* homework?
(Circle one only)

Never 1
1 or 2 days per week 2
3 or 4 days per week 3
5 days or more per week 4

19 About how much time do you spend on *[language of test]* homework when you get it?
(Circle one only)

I do not get *[language of test]* homework 1
None 2
Up to a quarter of an hour 3
Up to half an hour 4
Up to one hour 5
More than one hour 6

20 How often are you given reading to do at home by your *[language of test]* teacher?
(Circle one only)

Never 1
1 or 2 days per week..... 2
3 or 4 days per week..... 3
5 days or more per week..... 4

21 Are you asked questions in class about your *[language of test]* homework?
(Circle one only)

I do not get *[language of test]* homework..... 1
Always..... 2
Most of the time..... 3
Sometimes..... 4
Hardly ever..... 5
Never..... 6

C. Questions about your reading (Questions 22 to 24)

22 How often do you borrow books from a school or public library?
(Circle one only)

Never..... 1
Hardly ever..... 2
Once a month..... 3
Once a week..... 4
More than once a week 5

23 How good are you at reading?
(Circle one only)

Not very good 1
Average..... 2
Good..... 3
Very good..... 4

24 Which do you think are the three most important ways to become a good reader?
(Choose three only)

		MOST IMPORTANT
a)	Liking it	2
b)	Having lots of time to read	2
c)	Being able to concentrate well	2
d)	Knowing how to sound out words	2
e)	Learning the meaning of lots of words	2
f)	Having many good books around	2
g)	Having a lively imagination	2
h)	Having lots of reading for homework	2
i)	Having lots of drill (practice) at hard things	2
j)	Having lots of written exercises	2
k)	Being told how to do it	2

D. Questions about your school (Questions 25 and 26)

25 How much time do you usually spend per week on reading silently in your (language or literature) class?
(Write your answer in hours and minutes; if you do not read silently in class, enter '0 hours and 0 minutes')

_____ hours and _____ minutes

26 How often are you given written work about what you have been reading in your language or literature class?
(Circle one only)

Never..... 1
Hardly ever..... 2
Sometimes..... 3
Most of the time 4
Always 5

Questions about your reading activities

E. School and homework

	How often do you read books and textbooks in the following areas in school and for homework?	almost never	less than once a month	one or two times a month	about once a week	two or three times a week	almost every day	
27	Stories/Literature/Fiction in *(language of test)*	1	2	3	4	5	6	
28	Science/Mathematics	1	2	3	4	5	6	
29	Foreign Language	1	2	3	4	5	6	
30	History/Geography/ Civics/Economics	1	2	3	4	5	6	
31	Vocational-Technical/Home Economics	1	2	3	4	5	6	
32	Reference/Directory/ Dictionary or Encyclopedia, in all subjects	1	2	3	4	5	6	
33	Workbooks and Exercises	1	2	3	4	5	6	

F. Personal Interest and Leisure

a) Books

	How often do you read these materials for personal interest and leisure time?	almost never	less than once a month	one or two times a month	about once a week	two or three times a week	almost every day	
34	Mystery/Spy	1	2	3	4	5	6	
35	Romance	1	2	3	4	5	6	
36	Sports/Recreation/Health	1	2	3	4	5	6	
37	Adventure	1	2	3	4	5	6	
38	Science Fiction	1	2	3	4	5	6	
39	Horror	1	2	3	4	5	6	
40	Poetry	1	2	3	4	5	6	
41	History/Politics	1	2	3	4	5	6	
42	Humour	1	2	3	4	5	6	
43	Science/Technology	1	2	3	4	5	6	
44	Travel	1	2	3	4	5	6	
45	Wildlife and Nature	1	2	3	4	5	6	
46	Music	1	2	3	4	5	6	
47	Classical Literature	1	2	3	4	5	6	
48	Biography/Autobiography	1	2	3	4	5	6	

F. Personal Interest and Leisure (ctd.)

b) Magazines

	How often do you read these topics in magazines?	almost never	less than once a month	one or two times a month	about once a week	two or three times a week	almost every day	
49	Sports/Recreation/Health	1	2	3	4	5	6	
50	Movies/TV	1	2	3	4	5	6	
51	Romance/Fashion	1	2	3	4	5	6	
52	News/Politics	1	2	3	4	5	6	
53	Music	1	2	3	4	5	6	
54	Wildlife and Nature	1	2	3	4	5	6	
55	Computers/Technology	1	2	3	4	5	6	
56	Car/Motorcycle	1	2	3	4	5	6	

c) Newspaper Reading

	How often do you read these sections of newspapers?	almost never	less than once a month	one or two times a month	about once a week	two or three times a week	almost every day	
57	Sports/Recreation/Health	1	2	3	4	5	6	
58	Comic Strips/	1	2	3	4	5	6	
59	Movies/TV	1	2	3	4	5	6	
60	News/Politics	1	2	3	4	5	6	
61	Romance/Fashion	1	2	3	4	5	6	
62	Classified Advertisements	1	2	3	4	5	6	
63	Business/Finance	1	2	3	4	5	6	

G. School and Personal Interest

a) How often do you read these types of documents?		almost never	less than once a month	one or two times a month	about once a week	two or three times a week	almost every day	
64	Tables	1	2	3	4	5	6	
65	Charts/Graphs	1	2	3	4	5	6	
66	Maps	1	2	3	4	5	6	
67	Schedules/Timetables	1	2	3	4	5	6	
68	Directories	1	2	3	4	5	6	
69	Forms	1	2	3	4	5	6	
70	Diagrams	1	2	3	4	5	6	
71	Directions/Instructions	1	2	3	4	5	6	

International Option

72 School is a place where...
(Each item in this question says that school is a place where particular things happen to you or you feel in a particular way. We want to see whether you 'definitely agree', 'mostly agree', 'mostly disagree', or 'definitely disagree' with the items.
Please read each item carefully and circle one number in each line which best describes how you feel)

	Definitely disagree	Mostly disagree	Mostly agree	Definitely agree
1) I know how to cope with the work	1	2	3	4
2) Teachers are fair and just	1	2	3	4
3) I really like to go	1	2	3	4
4) Mixing with other people helps me understand myself	1	2	3	4
5) I feel important	1	2	3	4
6) I learn to get along with other people	1	2	3	4
7) Teachers help me to do my best	1	2	3	4
8) People have confidence in me	1	2	3	4
9) Teachers treat me fairly in class	1	2	3	4
10) People come to me for help	1			
11) I feel lonely	1	2	3	4
12) I feel restless	1	2	3	4
13) I know that people think a lot of me	1	2	3	4
14) Teachers give me the marks I deserve	1	2	3	4
15) People look up to me	1	2	3	4
16) I feel depressed	1	2	3	4
17) I know I can reach a satisfactory standard in my work	1	2	3	4
18) I know the sorts of things I can do well	1	2	3	4
19) I learn a lot about myself	1	2	3	4
20) Teachers listen to what I say	1	2	3	4
21) I feel happy	1	2	3	4
22) I find that learning is a lot of fun	1	2	3	4
23) I get enjoyment from being there	1	2	3	4
24) I get satisfaction from the school work I do	1	2	3	4
25) I get to know myself better	1	2	3	4
26) I know I can do well enough to be successful	1	2	3	4
27) I get upset	1	2	3	4
28) I feel great	1	2	3	4
29) I have learnt to accept other people as they are	1	2	3	4

Thank you very much for your cooperation

Reading Literacy:

Student Questionnaire Population B Accompanying Notes

Questions 4 and 5 (Parental Education):

This question has been changed. The problem with the pilot-study question was that primary education has a very different number of years in different countries. Hence, the change to number of years. It is up to each country to collect these data in the most appropriate way. However, each NRC must hand in the valid range of values for his/her country. NRCs must also check their data before sending them in. This might be achieved in some countries by asking for the highest level of education and converting it to years.

Question 13 and 14 (Possessions):

Same remarks as for Population A.

Question 17 (Homework):

Work set by the teacher to be completed outside of normal school time but which could be done anywhere.

Question 18-19 (*[Language of test]* Homework):

Reading homework includes all reading in the language of the test program. This could include reading comprehension from the language or literature program.

Questions 27 - 71:

These questions were formerly in the Reading Activity Inventory. The Student Questionnaire is considerably shorter and administration time has been reduced by combining these two instruments. A number of NRCs felt that this step would be desirable.

Question 72:

The 'Quality of School Life' question has now become an International Option.

IEA-Reading Literacy Study:

Population B
Teacher Questionnaire

International Coordinating Center
Hamburg 1990

Identification:	target pop.	
	country	
	language	
	stratum	
	school	
	class	
	teacher	
	seq. no.	

Reading Literacy:

Teacher Questionnaire

Population B

✎ Directions:

The following questionnaire is part of an international study of Reading Literacy and attempts to identify differences in [language of test] instruction. It is recognized that teachers are likely to respond quite differently.
Please answer all questions in such a way as to reflect most clearly your teaching practice. Most questions require you to circle your selected response. Others require you to write in a number. Where it is appropriate to enter '0' in the answer, please do so. Do not leave blanks.
We thank you for your effort.

A. The first set of questions has to do with you and your educational training (Questions 1 to 7)

1 Your sex:
(Circle one number only)

Male 1
Female 2

2 Is your mother-tongue the same as *[language of the test]*?
(Circle one number only)

No 1
Yes 2

3 How many years of primary and secondary school education did you have altogether?
(If you have had no school education, please enter '0'.
Do not include pre-compulsory education e.g. Kindergarten. Also do not count grade repetition years.)
[NRCs: see accompanying note.]

_____ years (or full-time years equivalent [e.g., two half-years equal one full year] to nearest whole number) ___

4 How many years of all your education was pre-service teacher-training?
(If you have had no pre-service teacher training, please enter '0'.)
[NRCs: see accompanying note.]

_____ years ___

5 How many years of post-secondary education did you have?
(If you have had no post-secondary education, please enter '0'.
Do not include pre-service teacher training and grade repetition.)
[NRCs: see accompanying note.]

_____ years (or full-time years equivalent to nearest whole number) ___

6 By the end of this school year how many years will you have been teaching altogether?

_____ years (or years equivalent) ___

7 About how often do you read each of the following?
(Do not include reading for preparation of class lessons.
Circle one number per line only)

		never or almost never	once a year	about once a term	about once a month	about once a week or more	
a)	Articles on teaching	1	2	3	4	5	___
b)	Articles on reading comprehension	1	2	3	4	5	___
c)	Books on history or politics	1	2	3	4	5	___
d)	Books on the arts	1	2	3	4	5	___
e)	Books on science	1	2	3	4	5	___
f)	Novels or short stories	1	2	3	4	5	___
g)	Poems	1	2	3	4	5	___
h)	Plays	1	2	3	4	5	___
i)	Articles on literature	1	2	3	4	5	___

B. This set of questions has to do with your class being tested (Questions 8 to 12)

8 How many students are enrolled in this class?

_____ students

9 How many students in this class do not have *(language of test)* as their first language?
(If none, please enter '0'.)

_____ students

10 How many students in this class need remedial help in reading?
(If none, please enter '0')

_____ students

11 How many students in this class receive remedial help in reading?
(If none, please enter '0')

_____ students

12 What is the number of hours and minutes of total instructional time *(see explanation in the accompanying notes)* excluding breaks for this class in a typical week in your school? *(For all subject areas)*

_____ hours and _____ minutes per week

C. The following set of questions has to do with your teaching (Questions 13 to 20)

13 How much time per school week is devoted to the teaching of *(language of test)* for this class?
(If [language of test] is not taught, please enter '0'.)

_____ hours and _____ minutes per week

14 What assessment methods do you use most often in the *[language of test]* class?
(Rank order by assigning a '1' to the most frequent, '2' to the next, and '7' to the least frequent.)

a)	Teacher quizzes	_____
b)	Multiple-choice questions	_____
c)	Records of student interests	_____
d)	Oral discussions	_____
e)	Oral discussions on material read	_____
f)	Written open-ended questions on material read	_____
g)	Essays in response to literature	_____

15 How frequently did you teach in your class this year how to understand each of the following kinds of text?
(Circle one number per line only)

	Frequency					
	almost never	3 or 4 times a year	about once a month	at least once a week	nearly every day	
a) Narration: texts that tell a story or give the order in which things happen	1	2	3	4	5	___
b) Exposition: texts that describe things or people or explain how things work or why things happened	1	2	3	4	5	___
c) Documents: tables, charts, diagrams, lists, maps	1	2	3	4	5	___

16 What do you regularly do (i.e. at least once a week) to encourage your students to read outside school?
(Choose the two you do most frequently)

a)	Hold discussions about books	2	___
b)	Suggest titles/authors	2	___
c)	Encourage them to borrow books from the school library	2	___
d)	Give special reading assignments	2	___
e)	Other *(please specify)*	2	___

_____

17 How often are your students typically involved in the following activities?
(Circle one number per line only)

	Activities	Frequency			
		almost never	about once a month	about 1 or 2 times a week	almost every day
a)	Silent reading in class	1	2	3	4
b)	Answering text comprehension questions in writing	1	2	3	4
c)	Independent silent reading in a library	1	2	3	4
d)	Listening to students reading aloud	1	2	3	4
e)	Discussion of books	1	2	3	4
f)	Learning new vocabulary systematically *(e.g. from lists)*	1	2	3	4
g)	Learning new vocabulary from texts	1	2	3	4
h)	Learning library skills	1	2	3	4
i)	Reading plays or dramas	1	2	3	4
j)	Summarizing their reading	1	2	3	4
k)	Relating experiences to reading	1	2	3	4
l)	Reading other students' writing	1	2	3	4
m)	Studying the style or structure of a text	1	2	3	4
n)	Reading in other subject areas	1	2	3	4
o)	Writing in response to reading	1	2	3	4
p)	Participating in a discussion of texts led by students	1	2	3	4
q)	Learning to use illustrations (graphs, diagrams, tables) to understand text	1	2	3	4

18 Please rank only five of the following aims of reading instruction in order of the importance you attach to each of them.
(Place '1' next to the most important and so on to '5' for the least important. Use all 5 ranks once only.)

	Aims	Importance
a)	Developing a lasting interest in reading	_____
b)	Improving students' reading comprehension strategies	_____
c)	Developing students' research and study skills	_____
d)	Extending students' vocabulary	_____
e)	Developing students' critical thinking	_____
f)	Expanding students' world views	_____
g)	Increasing speed of reading	_____
h)	Expanding students' variety of reading choice	_____
i)	Teaching students how to apply study strategies to other subjects	_____
j)	Increasing students' appreciation of literature	_____
k)	Teaching students how to interpret diagrams and graphs	_____

19 How often do you teach or help students to increase their comprehension by...?
(Circle one number on each line)

	Never	Once in a while	Quite often	Most of the time
a) Thinking about what they know already about the same topic	1	2	3	4
b) Remembering other things they have read about the same topic	1	2	3	4
c) Trying to predict what will happen or what information they might learn	1	2	3	4
d) Talking to somebody else about what they will read	1	2	3	4

20 How often do you teach or encourage students to improve their comprehension by using these strategies?
(Circle one number on each line)

	Never	Once in a while	Quite often	Most of the time
a) Picturing in their mind what is happening as they read	1	2	3	4
b) Trying to predict what will come next	1	2	3	4
c) Looking back over what they have read	1	2	3	4
d) Writing down notes or ideas about what they have read	1	2	3	4
e) Comparing what they have read with experiences they have had	1	2	3	4
f) Thinking about similar things they have read	1	2	3	4
g) Talking to somebody else about what they have read	1	2	3	4
h) Writing something of their own on what they have read	1	2	3	4

D. The following set of questions has to do with your school library and class resources (Questions 21 to 24)

21 In a normal *[language of test]* period how many *[language of test]* textbooks are available for each student in the class tested?
(Circle only the most appropriate answer)

None 1
1 book for about 5 or more students 2
1 book for each 2 students 3
1 book for each student 4
2 books for each student 5
3 or more books for each student 6

22 Do you have a school library in your school?
(Circle one number only)

No 1
Yes 2

➪ If you answered 'No' to Question 22, please go to Question 25. Thank you.

23 How often do your students visit the school library as a class?
(Circle one number only)

Hardly ever 1
Once a month 2
Once a week 3
More than once a week 4

24 Can your students borrow books from the school library to take home?
(Circle one number only)

No 1
Yes 2

E. The last set of questions are to do with school organisation (Questions 25 to 28)

25 Is your work as a teacher evaluated by the school principal (or deputy school principal)?

No 1
Yes.....2

26 Does the school principal (or deputy principal)...
(Check the appropriate answer.)

		No	Yes
a)	discuss with you explicit achievement standards for the subject that you teach	1	2
b)	ask for evaluation results or progress of your students in reading	1	2
c)	make suggestions about the choice of instructional methods in reading	1	2
d)	encourage contacts among teachers	1	2
e)	initiate activities directed at the professional development of teachers	1	2
f)	make suggestions about the content that must be covered in reading	1	2

27 How often do you have staff meetings at your school?
(Check one only.)

Never.....1
Once a year.....2
Once a term.....3
Monthly 4
Weekly.....5

28 If you have staff meetings, please indicate how often the following items occur as subjects of discussion during staff meetings.
(Circle one number on each line.)

		all staff meetings	most staff meetings	some staff meetings	not in any staff meetings
a)	curriculum content	1	2	3	4
b)	the way the subject matter is presented	1	2	3	4
c)	professional development of teachers	1	2	3	4
d)	issues of 'pastoral care' (e.g. student problems, guidance, welfare)	1	2	3	4
e)	organizational issues (e.g. school climate, co-ordination of work among teachers, the way decision-making procedures are conducted)	1	2	3	4
f)	other topics (e.g. purely administrative tasks, leisure and social activities)	1	2	3	4

Thank you very much for your cooperation

Reading Literacy:

Teacher Questionnaire Population B Accompanying Notes

Question 3:
NRCs should ask the question in such a way that it is absolutely clear what is meant. In the pilot study some systems with 12 years of primary and secondary education were submitting this variable with an average of 14 to 17 years of primary education! Please ensure that this does not happen for the main study. NRCs must supply in valid range information.

Questions 4-5:
Same type of comment as for Population A. NRCs must check the validity of replies. Also send in valid range information.

Question 12:
Note that this is a check question against Question 15 in the School Questionnaire.
'Total instructional time' here includes both language and/or literature.

Questions 25-28:
These are the OECD questions and are concerned with school leadership.

IEA-Reading Literacy Study:

Population B
School Questionnaire

International Coordinating Center
Hamburg 1990

Identification:	target pop.	
	country	
	language	
	stratum	
	school	
	class	
	seq. no.	

Reading Literacy:

School Questionnaire

Population B

Directions:

The following questionnaire is part of an international study of Reading Literacy. The questions asked attempt to gather information which captures the wide range of experience and practice which is likely to exist across countries.
Please answer all questions in such a way as to reflect most accurately the situation in your school. It is important that all questions are anweredMost questions require you to circle your selected response. Others require you to write in a number. Where it is appropriate to enter '0' in the answer, please do so. Do not leave it blank. All information will be treated in the strictest confidence.

1 How many years will you have been a school principal by the end of this school year?

in your total career _____ years

in your present school _____ years

2 What is the total enrollment of full-time students in your school?
(If there are no boys or no girls, please enter '0')

Number of boys _____
Number of girls _____

3 What is the total enrollment of full-time *[grade level being tested]* students in your school?
(If there are no boys or girls, please enter '0')

Number of boys _____
Number of girls _____

4 Is your school...

A state school..... 1
A private school..... 2

5 What is the type of community served by your school? *(☞ see note)*
(Circle one only)

A village or rural community..... 1
A small town community..... 2
A large town community..... 3
A city of 1 million or more..... 4

6 Please indicate the availability of the following resources in relation to your school?
(Circle one number on each line)

	Not readily available	Available in neighbouring town or city (less than 2 hours of normal one way travel time)	Available locally (within 30 minutes of normal one way travel time)
Public Library	1	2	3
Bookstore/book department store	1	2	3
Other secondary level schools	1	2	3
A higher education institution	1	2	3

7 What is the degree of parent co-operation with the school in terms of support for the schools educational principles (compared with other schools you know)?
(Circle one only)

Much below average..... 1
Below average..... 2
Average..... 3
Above average..... 4
Much above average..... 5

18 Does your school have any special programs or teacher initiatives for reading outside the normal classroom activities (e.g. enrichment programs, or special programs for the disadvantaged)?
(You may circle more than one)

Extra class lessons in reading..... 2
Extra individual tuition at school..... 2
Special remedial reading courses..... 2
Other..... 2

19 Does your school sponsor any informal initiatives to encourage reading (e.g. book clubs)?

No 1
Yes..... 2

20 Do you have a program for the improvement of reading instruction in your school?

No 1
Yes..... 2

21 What are the two most serious problems which you experience in providing for the teaching and learning of reading in your school? (☞ *see note*)
(Circle two problems only)

Insufficient specialized staff..... 2
Insufficient time 2
Students' lack of interest 2
Insufficient classroom material..... 2
Other..... 2
None - there are no serious problems 2

8 Which of the following resources and activities are there in your school?
(Circle one numer on each line)

	No	Yes
School library	1	2
Reading room for students	1	2
Student newspaper or magazine	1	2
Teacher (Professional) library	1	2
Drama Club	1	2
Debating Club	1	2
Literature Club	1	2
Writing Club	1	2

☞ The following questions are about the school library.
If you have indicated in Question 8 that your school does not have a library, please go straight to Question 12.

9 Approximately how many books with different titles does your school library contain?
(Exclude magazines and periodicals.)

_____ different titles

10 Approximately how many books with different titles were added to your school library in the last year?
(Exclude magazines and periodicals.)

_____ books with different titles

11 Can *[grade level being tested]* students in your school borrow books from your school library to take home?
(Circle one only)

No 1
Yes..... 2

12 How many full-time (or full-time equivalent) teaching teachers are there in your school? (☞ *see note*)
(Exclude non-teaching principal and administrators. If there are no male or no female teachers, please enter '0'. For full-time equivalent, add the number of part-time teachers. For example, two half-time equivalent teachers equal one full-time equivalent. Or, three third-time teachers equal one full-time equivalent teacher. Round to nearest whole number.)

_____ male teachers
_____ female teachers

13 How many full-time (or full-time equivalent rounded to the nearest whole number) special teachers are there in your school (i.e. for remedial teaching, counselling, guidance. Exclude, for example, librarians and physical education teachers)?
(If there are no male or no female special teachers, please enter '0').

_____ male teachers
_____ female teachers

14 What is the number of hours and minutes of total instructional time *(see explanation in the accompanying notes)* excluding breaks *(for the class being tested)* in a typical week in your school (for all subject areas)?

_____ hours and _____ minutes per week

15 How many weeks per year is your school open for the class being tested?

_____ weeks per year

16 How many days of instruction were lost in the last school year (due to accidents, floods, strikes, festivals, staff days, etc.?)
(If there were no days lost, please enter '0')

_____ days lost in the last school year

17 On an average day, approximately what percentage of students are absent from school?

_____ percent

22 Please rank the following activities in order of importance in your work as a school principal.
('1' is the most important activity, '8' is the least important activity, 'NA'= not applicable. Do not assign equal rankings.)

		rank of importance
a)	representing the school at official meetings	_____
b)	evaluation of staff	_____
c)	contacts with local community (e.g. parents, community organizations, local industry)	_____
d)	discussing educational objectives with the teaching staff	_____
e)	administrative tasks concerning the functioning of the school (e.g. regulations, disciplinary duties, school budget, timetable)	_____
f)	using records of pupils' progress	_____
g)	taking care of issues of 'pastoral care' (e.g. student problems, guidance, welfare)	_____
h)	activities aimed at the professional development of teachers	_____

23 How often do you systematically evaluate the pedagogical work of the teachers at your school?
(Check one only.)

Never.....1
Less than once a year.....2
About once a year.....3
More than once a year.....4

24 What procedures do you use to gather information for your evaluation?
(You may choose more than one alternative.)

	Yes
I do not systematically evaluate	2
interviews	2
written or oral self reports by teachers	2
observational data on teachers' classroom work	2
student ratings of teachers' performance	2
other forms of systematic evaluation	2

Thank you very much for your cooperation

Reading Literacy:

School Questionnaire Population B Accompanying Notes

Question 3:
Two derived variables will be produced:
1. Total enrolment;
2. Sex of school.

Question 4:
State/Private Schools:
The purpose of this question is to compare state and private schools. At issue here is control of the school, rather than its funding. NRCs may have to substitute an appropriate term for 'state' in their own countries.

Question 5:
The basic aim of this question is to have a rural-urban scale for where schools are situated. The relative size of the types of community mentioned in the queston will vary from country to country. It is up to each country to phrase the options in such a way that the dimension from rural to large urban is clear. The result must be a four-point scale. If a dichotomy of urban/rural is to be made it will be category 1 for rural and categories 2, 3, and 4 for urban.

Question 12:
Principals will have to convert part-time to full-time equivalent rounded to the nearest whole number. NRCs to make sure that this is clear to the respondents.
Two derived variables will be produced:
1. Total number of full-time teaching staff;
2. Pupil teacher ratio.

Question 13:
The intention is to capture remedial and/or specialist resources for teaching programmes. Two derived variables will be produced:
1. Total number of special teachers;
2. Ratio of pupils to special teachers.

Question 14:

In the pilot test some countries did not make clear that the question targeted total instructional time in a typical week. NRCs need to specify valid ranges.

This question refers to the class being tested. If two or more classes are tested within one school then the school principal should put in the figures for the first class.

Question 15:

Note in the pilot study values ranged from 0 - 52 weeks. NRCs need to specify valid ranges.

Question 16:

Maximum values in the pilot study seemed very unlikely in some countries (e.g. '91'). NRCs to specify valid ranges.

Question 17:

In the pilot study some maximum values were not likely (98 percent and 165 percent). NRCs to supply valid ranges.

Questions 20 and 21:

The teaching and learning of reading includes any program aimed at the improvement of *[language of test]* skills.

Questions 22-24:

These are the OECD questions and are concerned with school leadership.

Appendix C

National Case Study Questionnaire

International Reading Literacy Study

National Case Study Questionnaire

Final Version

BELGIUM (FRENCH) • BOTSWANA • CANADA (B.C.) • CANADA (Q.) • CHINA • DENMARK • FEDERAL REPUBLIC OF GERMANY • FINLAND • FIJI • FRANCE • GREECE HONG KONG • HUNGARY • ICELAND • INDONESIA • IRELAND • ITALY • KENYA • KOREA LUXEMBOURG • MAURETANIA • NETHERLANDS • NEW ZEALAND • NIGERIA • NORWAY PHILIPPINES • POLAND • PORTUGAL • SINGAPORE • SLOVENIA • SPAIN • SWEDEN SWITZERLAND • THAILAND • TRINIDAD AND TOBAGO • UNITED STATES OF AMERICA VENEZUELA • ZAMBIA • ZIMBABWE

Reading Literacy:
National Case Study Questionnaire

Instructions to National Research Coordinators

The National Case Study Questionnaire is designed to gather educational, economic, social, political, and cultural data which may help explain variances in school achievement across nations and regions.

It is not expected that a single respondent will actually gather all the data requested. However, an individual at each National Centre should be responsible for seeing that the necessary assistance from experts is arranged for and that the questionnaire is completely filled in. The National Centre might find it useful for an economist, a sociologist, an administrator, and other specialists to gather data about those items within each of their areas of specialization.

Much of the data can be obtained from published national and international sources. However, in some cases, the published sources are badly out-of-date. With the assistance of experts, National Centers should be able to locate data which are more accurate and up-to-date. For some information, we are seeking trend data and, therefore, are asking for data over a period of time.

It would be helpful if, in completing the questionnaire, the National Centre would indicate the source of data for each item and the level of confidence - i.e., high, medium, or low. Where data do not exist or are known to be poor, a "best estimate" should be provided. Best estimates are preferred to no data at all but they should be indicated as such.

When data are not available and no estimate is possible, please do not leave the item blank. Instead, mark it with "No Information Available" or "Not Applicable" or similar indication.

Please use the back of the sheets for any additional information you wish to provide - including comments on or amplifications of your responses. Your attention is called to the last page in which you are asked to comment on the particular aspects of your situation which might help to explain the achievement of the students in your country.

This questionnaire will be completed close to the time of national testing.

1. Structure of the School System

1a. Using whatever sources available (e.g. your government sources, European Council for Cultural Cooperation, *School Systems, A Guide*; or UNESCO, *World Survey*, or a similar source) prepare an up-to-date diagram of the organization of your national school system. Use arrows to indicate points at which critical decisions and transfers are made. Also, mark the points at which Reading Literacy target populations are to be found. The length of the base line should represent 100 per cent of an age group in school and in different school types should be drawn to scale. An example of Australia is given in Appendix I to this questionnaire.

1b. Complete student enrollments by age, grade of full-time status, for each school type. See Table 1b on page 3. Note: these should be adjusted and not be greater than 100%

1c. Complete student enrollments by age, grade and sex, for each school type. See Table 1b on page 3.

1d. What is the total number of years of *compulsory full-time schooling*?

_____

1e. At what age does *compulsory full-time schooling* begin?

_____

1f. At what age does *compulsory full-time schooling* end?

_____

1g. For how many years has education been compulsory?

(a) Up to age 9 _____

(b) Up to age 14. _____

1h. How many years, if any, of institutionalized *pre-compulsory school care and education* are generally available? What proportion of the total age cohort attends each year?

Year of pre-schooling	Age	Proportion of Total Age Cohort	Proportion Attending		Total State	Total Private
			Male	Female		

What kinds of pre-reading experiences do chidlren have in these institutions?

		In most institutions	In a few institutions	Rarely or Never
eg	(a) Are children taught the alphabet?			
	(b) Are they read to regularly?			
	(c) Do they study environmental print?			
	(d) Do they learn the conventions of print as presented in books?			
	(e) Do they write captions on their art work?			
	(f) Do they write simple messages?			
	(g) Are they provided with experences aimed at expanding their oral language?			

Excluded population

When defining the target population to be tested some countries excluded a number of schools/classes for a variety of reasons.

Did you exclude any types of schools/classes?

Yes _____

No _____

If yes, please describe the types of schools/classes and indicate the percentage of the relevant age group of students represented in these excludes schools/classes.

Type of school/class	Population A or B	Percentage of age group

1b. Student Enrollments: Last year data available, Male/Female 198_

Grade in the 8th month of the school year which the modal age of students is	Grade	Elementary/ Primary School	Comprehensive Secondary School		Academic Secondary School		Vocational Secondary School		Other Secondary School		Special Ed Retarded, Handicapped	Total Enrolled		Total Size of Age Cohort		Est. % of students (still in school)
			Full Time	Part Time	Full Time	Part Time	Full Time	Part Time	Full Time	Part Time		Full Time	Part Time	Full Time	Part Time	
4 years																
5 years																
6 years																
7 years																
8 years																
9 years																
10 years																
11 years																
12 years																
13 years																
14 years																
15 years																

1c. Student Enrollments: Last year data available, Male/Female 198_

Grade in the 8th month of the school year in which the modal age of students is	Grade	Elementary/ Primary School	Comprehensive Secondary School		Academic Secondary School		Vocational Secondary School		Other Secondary School		Special Ed Retarded, Handicapped	Total Enrolled		Total Size of Age Cohort		Est. % of students (still in school)	
			Male	Female	Male	Female	Male	Female	Male	Female		Male	Female	Male	Female	Male	Female
4 years																	
5 years																	
6 years																	
7 years																	
8 years																	
9 years																	
10 years																	
11 years																	
12 years																	
13 years																	
14 years																	
15 years																	
16 years																	
17 years																	
18 years																	
19 years																	
20 years																	

2. Curriculum in Reading

2a. Is there a curriculum in reading prescribed by a central or regional authority?

☐ Yes at the central level ☐ No
☐ Yes at the regional level ☐ No

2b. If it is a centrally prescribed curriculum , please send a copy to the ICC in Hamburg.

2c. How much freedom do teachers have to implement the curriculum in respect of:

		a great deal	some	very little	none at all
(I)	Aims and objectives of reading	o	o	o	o
(II)	Selection of reading materials	o	o	o	o
(III)	Methods of reading instruction	o	o	o	o
(IV)	Assessment of reading	o	o	o	o
(V)	Time assigned to reading	o	o	o	o

2d. At what age do most children <u>begin</u> to receive formal instruction in reading?

_____

2e. At what age does formal instruction in reading <u>cease</u> for the majority of pupils?

_____

2f. Describe briefly the main methods used to teach reading in the <u>first year</u> of reading instruction. *(Use back of page if necessary.)*

Please indicate which do not apply, refer especially to:

(I) Relationship with speaking and listening
(II) Role of phonics
(III) Other decoding/word-attack skills
(IV) Kinds of materials
(V) Control over vocabulary
(VI) Role of meaning
(VII) Relationship with writing"

_____

_____

3. The Teachers

3a. What percentage of primary/elementary teachers are female?

_____%

3b. What percentage of secondary teachers have university degrees?

_____%

3c. What percentage of primary/elementary teachers have university degrees or equivalent?
(Bachelor's level or beyond)

_____%

3d. What percentage of primary/elementary teachers have less than one year formal teacher training?

_____%

3e. Please describe briefly current developments in the provision of teacher training.

_____

_____

_____

3f. Give the median annual remuneration from work, before taxes, for each of the following groups in 1980 and the latest year for which data is available. Indicate if figure includes value of pensions and all other prerequisites.

Currency Unit: _____

	1980	Latest year (198)
Elementary/primary school teachers		
Secondary school teachers		
Bus Driver		
Nurses		
Police		
University lecturers		
If currency unit not constant, give index of inflation. 1980 = 100		

3g. How many teachers of reading would a typical child have had in the first 4 years of school?

- o 1
- o 2
- o 3
- o 4
- o 5 or more

4. Instruction Time

4a. How many instructional days are there in a school year in:

Primary school? _____ days
Secondary school? _____ days

4b. What is the average number of classroom (include laboratory and gymnasium) clock hours per day in:

Primary school? _____ hours
Secondary school? _____ hours

4c. On the average, what is the percentage of total daily clock hours per day devoted to instruction in):

(a) Instruction in the Language of Test includes all activities designed to improve children's control over language eg. reading, writing, speaking, listening, spelling, grammar, etc. Exclude reading in Maths, Science and Social Studies etc.

(b) Instruction in Reading includes all activities designed to improve children's reading abilities. Do not count other aspects of language included under (a).

Grade	a) Language Acts Excluding Reading	b) Reading and Book-Related Instruction (in Language of Test)
K		
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		

(K= Kindergarten)

4d. Elaborate if necessary.

5. Social and Economic Conditions

5a. How much money is spent on supporting the public and private schools (1980 and in the latest year for which figures are available)?

Currency Unit: _____

	Total Expenditure 1980	Source %		Latest Year (198-)	
		% Public	% Private	% Public	% Private
Primary/elementary schools Funds for recurrent expenditure (teachers' salaries, materials, interest on debt, etc.)					
for capital expenditure					
Secondary schools Funds for recurrent expenditure (teachers' salaries, materials, interest on debt, etc.)					
for capital expenditure					

5b. What percentage of the Total State Government Budget is spent on formal schooling? (Include only primary, elementary and secondary schools). _____

5c. What is current live birth rate per 1,000 population?

_____ _____ year

5c. What is the percentage of all women aged 20 to 44 who are in employment outside the home?

Full-time employment _____ % _____ year

Part-time employment _____ % _____ year

5d. What is the current labour fource Unemployment rate?

_____ % _____ year

5e. What is the mean Gross Domestic Product (GDP) per capita over the last 5 years?

_____

5f. What is the national average disposable income per capita (i.e., income after payment of direct taxes in 1990)?

(I) Local currency _____

(II) US-Dollars _____

5g. What is the rate of infant mortality?

_____ per 1,000 live births _____ year

5h. What is the number of hospital beds per 1,000 population?

_____ per 1,000 population _____ year

5i. What number of households have telephones?

_____ per 1,000 households _____ year

6. Language and Culture

6a. List the major mother tongue languages spoken in this country (up to 7 languages)

Mother tongue/ (or Ethnicity)	Percent of Population

6b. Name the language(s) recognized for official administrative and instructional use by the state:

Language(s)	Administrative	Instructional
At the national level:		
At the regional level:		

6c. For approximately what percentage of the children enrolled in school is the language of schooling different from the language of the home?

In the primary schools _____ % Grades (1-9)
In the secondary schools _____ % Grades (10-13)

6d. Is there political debate at the national, or regional levels concerning the language of instruction?

Primary schools:	Regional Level:	o	Yes
		o	No
	National Level:	o	Yes
		o	No
Secondary schools:	Regional Level:	o	Yes
		o	No
	National Level:	o	Yes
		o	No

If *yes*, indicate the nature of debate:

6e. Does the public school curriculum contain instruction based on minority group history, language, music, folk-art, etc.?

o Yes
o No

If *yes*, please specify giving examples:

6f. Identify up to five daily newspapers with the largest circulation and supply the following information about them.

Name of Newspaper	Language	Approximate Circulation Numbers

6g. What total circulation of daily newspapers in your country?_____

6h. How many books were published nationally in the last year for which figures are available?

_____ books _____ year

6i. What is the total of number of books held in public libraries in your country - please indicate the year to which the figure applies.

_____ books _____ year

6j. What is the total number of books sold annually in the last year for which figures are available?

_____ books _____ year

7. Final Comments

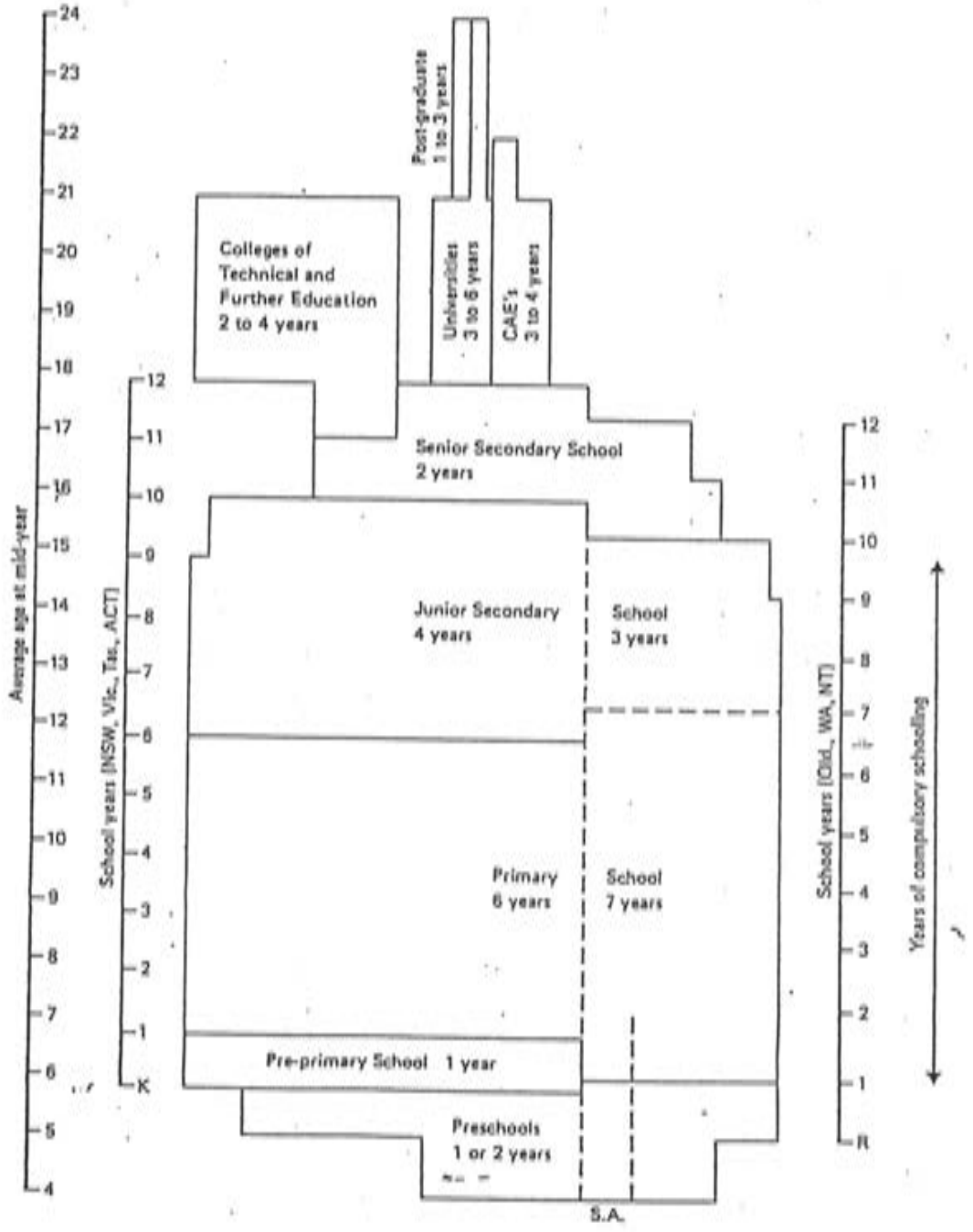
Having read the Student, Teacher, School and National Case Study Questionnaires, what are the special features of your nation's educational system which, in your estimation, may assist the interpretation of the IEA Reading Literacy results for your country with respect to the two questions below.

A The relative level of reading ability in your country?

B The spread of differences between pupils?

APPENDIX 1

The Structure of the Australian Education System



Appendix D

Variable List from the IEA Reading Literacy Study

Appendix D

This appendix is for use by authors having to interpret these variables and by secondary analysts.

The appendix lists, for each background variable in the Student, Teacher, and School Principal Questionnaires for Population A first and then for Population B, the following information.

1. Question number in the questionnaire
2. Variable label
3. Those countries with between 10 and 20 percent missing data for the variable
4. Those countries with more than 20 percent missing data for the variable
5. Special comments given by NRCs about problems with the variables in their country

Authors and analysts should undertake caution when working with variables with more than 20 percent missing data and where comments suggest that the variable should not be used.

A further word of warning concerns the "compliancy effect" for certain variables. When examining the mean values for certain variables it would appear to be the case that "socially desirable" answers may have been given in some countries. When examining the means you will see that it would seem to be unreasonable that some country means, particularly developing countries, are unexpectedly high. It would be unwise to compare national means on these variables.

The country/educational system abbreviations are as follows:

BEL (Fr)	Belgium (French)
BOT	Botswana
CAN (BC)	Canada (Brit. Columbia)
CYP	Cyprus
DEN	Denmark
FIN	Finland
FRA	France
GER (E)	Germany (East)
GER (W)	Germany (West)
GRE	Greece
HK	Hong Kong
HUN	Hungary
ICE	Iceland
IND	Indonesia
IRL	Ireland
ITA	Italy
NET	The Netherlands
NZ	New Zealand
NIG	Nigeria
NOR	Norway
PHI	Philippines
POR	Portugal
SIN	Singapore
SLO	Slovenia
SPA	Spain
SWE	Sweden
SWI	Switzerland
THA	Thailand
T/T	Trinidad & Tobago
USA	United States
VEN	Venezuela
ZIM	Zimbabwe

STUDENT QUESTIONNAIRE
Population A

No		Percent Missing		Special Comments	No
		10 - 20	21+		
1	ASAGEY/M ASAGED				1
2	ASSEX				2
3	ASUSLAN				3
4	ASMEAL1-3 ASMEALD	CAN (2,3) T/T:(2,3) HUN (2,3) ICE (2) IND (1) VEN (1)	IND (2,3) VEN (2,3)	ICE: "Hadegismatur" which is a hot meal. Children might at lunch time eat a sandwich or yoghurt which they might not look at as a meal.	4
5	ASNEWS	SLO		SLO: Probably drop for analysis	5
6	ASTV				6
7	ASBOOKS			GER(W): In analysis, use recoded variable only. GER(W): Ministry of Education of Lower Saxony refused permission to this question (Question not administered and coded to not applicable.)	7
8	ASHOM01-10 ASHOMD	CAN (8,10) CYP (8-10) NOR (2-4,8,9) GER(W) (1-10) POR (4,6) SWI (2,7) T/T (1-10) USA (1-10) VEN (5)	GER(E) (10) IND (1-10) VEN (1-4, 6-10)	GER(E): many pupils did not understand 10. GER(W): Ministry of Education of Lower Saxony refused permission to this question (Question not administered and coded to not applicable.) SLO: HOM01,HOM08,HOM10: missing is likely to mean *not possessing*. CYP: Many pupils did not understand the wording of 8, 9, 10	8
9	ASSTUD1-8 ASSTUDD	T/T(1-8). CAN (5-7) NOR (3,5-7) ICE (6) IND (1,3) USA (1-8) VEN (1,3, 6)	IND (2,4-8) VEN (2,4, 5,7,8)	GER(W): Ministry of Education of Lower Saxony refused permission to this question (Question not administered and coded to not applicable.) SLO: drop STUD7 for analysis.	9
10	ASPRHTL				10
11	ASPRETL				11

12	ASPRHOL	USA		DEN: This var dropped from dataset because of mistranslation.	12
13	ASPREOL	USA	DEN		13
14	ASBORBO				14
15	ASSRATE			CYP: Be careful using this variable. Translation was: *How good do you read and how good do you understand what you are reading?*	15
16	ASWAY01-11	NOR (1-5, 7-11) IND (1-11)	VEN (1-11)		16
17	ASREATL			SWI: If SLANG=26 or 27 then ASREATL=*not applicable* USA: Question missing in instrument but data are on the file.	17
18	ASASKRE				18
19	ASHWKF	BEL VEN		CYP: *reading homework* translated as *language exercises*	19
20	ASHWKT ASHWKD	CAN		CYP: *reading homework* translated as *language exercises*	20
21	ASQUEST	CAN		CYP: *reading homework* translated as *language exercises*	21
22	ASHWKH	CAN		CYP: *reading homework* translated as *language exercises*	22
23	ASHWKO			CYP: *reading homework* translated as *language exercises*	23
24	ASWRITE				24
25	ASBOOKW	VEN			25
26	ASBOOKF	GER(W) GER(E) VEN		GER(E)/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	26
27	ASCOMIW	VEN			27
28	ASCOMIF	GER(W) VEN		GER(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	28
29	ASMAGAW	GER(W) VEN			29

30	ASMAGAF	GER(W) GER(E) VEN		GER(E)/(W):If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	30
31	ASNEWSW	GER(W) VEN			31
32	ASNEWSF	GER(W) GER(E) VEN		GER(E)/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	32
33	ASDIRCF ASVREADD ASSREADD	GER(W) VEN		GER(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	33
34	ASALOUD	GER(W) VEN		FRA: no emphasis on reading aloud	34
35	ASALOUF	GER(E) ITA	SIN GER(W) VEN	POR: Not common to read aloud SIN: Confusion between reading aloud to self and to some one else SWI: Question doesn't make sense	35
36	ASALOUW	BEL CAN NET GER(E) ITA SIN	GER(W) VEN	SIN: Confusion between reading aloud to self and to some one else SLO: unreliable answers	36
37	ASMAT1-8	GER(W) (1-8) GER(E) (1-8) ITA (1-8) SLO (1-8)	VEN (1-8)	FRA: no emphasis on reading aloud NET: children do not read aloud POR: not common to read aloud USA: Instruction incorrect, drop variable from analysis	37
38	ASTEXTF	GER(W) GER(E) VEN		GER(E):/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	38
39	ASSTORF	GER(W) GER(E) VEN		GER(E)/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank. Most NRCs indicated that children had story rather than textbooks. Therefore, great caution should be used when reporting on this variable.	39
40	ASWORKF	GER(W) GER(E) VEN		GER(E)/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	40
41	ASEXERF	GER(W) GER(E) VEN		GER(E)/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank.	41
42	ASINFOF	GER(W) GER(E)	VEN	GER(E)/(W): If students did not know or if they were unable to write down the title of the book read they may have left the whole question blank. Wording refers to *homework* rather than *school work*. Drop from analysis.	42

TEACHER QUESTIONNAIRE
Population A

No		Percent Missing		Special Comments	No
		10 - 20	21+		
1	ATSEX			VEN: General comment: Grade 4 was tested. Officially the teaching of reading stops at the end of Grade 3.	1
2	ATMOTON				2
3	ATSCHED			FIN: Normally min ATSCHED should be *missing*.	3
4	ATTETRA	HNK POR VEN	T/T	CAN: This question was asked in a very different way. Caution to be exercised when reporting. GER: Question was asked in a different way. Variable should not be used for analysis HNK: Some teachers misunderstood pre-service teacher training as years of training before taking a teacher post. ITA: Misunderstanding: T's considered as *teacher training* their pre-service teaching years.	4
5	ATSECED ATEDUCD ATTRAIND			SWE: Formula not "fair" for Swedish teachers; their real teacher education. Q4 is not included. It is suggested that this not be used in analysis for Sweden. GER: Question was asked in a different way. Variable should not be used for analysis	5
6	ATSTUTE		POR		6
7	ATSETRA				7
8	ATFRRE1-9 ATFRRED	POR (3-5,8) SLO VEN (2-5, 7,8)		IND, SLO,T/T,VEN: This variable likely to be subject to *compliance effect*.	8
9	ATYEARS				9
10	ATTECLA			BEL: *Class* was translated as *grade*. FRA: *Class* was translated as *grade*. SWI: Option 5 omitted. In Belgium and France class was translated as grade.	10
11	ATMUGRA				11
12	ATCLSIZ			POR: 25 is legal max in public schools, anything above that must be private schools. T/T: The normal max. is 45, in some cases, a higher value represents a temporary joint class.	12
12	ATGRSIZ			HNK: ATGRSIZ=ATCLSIZ	12

13	ATOTLAN ATOTLAD			ITA: High Maximum value depends on confusion between dialect and other language.	13
14	ATMEPAR				14
15	ATNHELP ATNHELPD	HNK	IND VEN	HNK: Drop variable in analysis. *Training* was translated as *coaching*. The term 'special' was left out, so teachers included almost all their students in "need remedial help in reading". So probably those answers which do not exceed the maximum value could be included in analysis.	15
16	ATRHELP ATRHELPD ATPHELPD		VEN	GER: If ATRHELP=ATCLSIZ then ATRHELP=.I (invalid) ICE: Small schools have no special education teachers.	16
17	ATINSTH/M ATINSTD	HNK GRC IND ITA		NRCs reported that from their knowledge of the data collection there was inconsistent interpretation of these two variables and that they (or any derivation from them) should not be used in the analysis.	17
18	ATTLANH/M ATTLAND ATPLAND	HNK (H) VEN	GER(E)	GER(E): Some teachers were not willing to convert school hours into clock hours. SWE: Max too high; perhaps it reflects some teachers' opinion that they are teaching *Swedish* in every lesson.	18
19	ATTEREH/M ATTERED	CAN GER(W) GRC VEN	IND	GER(W): Integrated curriculum; individual instruction may have led teachers to leave this question blank. SWE: Max too high; perhaps it reflects some teachers' opinion that they are teaching *reading* in every lesson	19
20	ATACT01-28	DEN (2) FRA (2) ICE (2) ITA (1) POR (1,2,20,23) VEN (1,2)		ICE: Word-attack skills are seldom used in Iceland.	20

21	ATTEXBO				21
22	ATAIM01-12	CAN (1-12) IND (1-12)	SLO	FRA: 01 emphasis is on silent reading	22
23	ATSTR01-13	T/T (4) VEN (4)		NRCs felt that the *compliance effect* would be high for these variables and that comparison of means should not be undertaken.	23
24	ATGROUP				24
25	ATGRTYP	CYP NOR GER(W) IND	SLO VEN		25
26	ATGRNUM	BEL CYP CAN IND POR VEN	SLO	CAN: See ATGRTYP FRA: Many schools have only one class (with often fewer than 10 pupils). Variable probably has no meaning. GRC: Normally no grouping	26
27	ATFRTE1-3		SIN (1-3)	SIN: Be careful when using this variable. Possible confusion on part of teacher.	27
28	ATVIE01-26			HNK: Mistranslation of ATVIE05. Variable dropped from the data set.	28
29	ATENST1-5				29
30	ATMETH1-9	POR (7) T/T (6) VEN (4, 6-9)			30
31	ATARE01-10	CYP (10) DEN (10) NOR (10) ICE (10) VEN (4)	VEN (10)	ICE (10): *Decoding* is a term not known to all teachers. The term was not translated and given in English. NOR (10): Problems with technical terms and not understood CYP (10): Problems with technical terms and not understood.	31
32	ATASME1-6	VEN (1,3)		NOR: Problems with technical terms and probably not understood	32
33	ATASHWK				33
34	ATREHWK	NET		NET: Almost all *missings* mean *never*.	34

35	ATMHWK	NET		NET: Dropped one of the options (Q33 was filter). SIN: max value due to teachers not wishing to estimate	35
36	ATCLLIB				36
37	ATBOOH ATBOOKD	POR VEN	SLO SLO	ICE: Teachers might know only own books but not from other teachers. T/T: 50 per cent of schools did not have a school library.	37
38	ATDMAG ATMAGD	POR VEN	SLO	ICE: Many schools do not have a class library.	38
39	ATCLIBR	POR VEN	SLO	ICE: Many schools do not have a class library.	39
40	ATSCHLI				40
41	ATVISIT			ICE: Not all schools have a school library.	41
42	ATSLIBR			ICE: Not all schools have a school library.	42
43	ATEVALU			FRA: principal does not evaluate teachers, so ATPRIN1-6 missing is too low GER(W): Sensitive data! NRCs felt that the *compliance effect* would be big for this variable. Mean comparisons are dubious	43
44	ATPRIN1-6	FRA (4) POR (1-6) T/T (3)	FRA (1,2,3,5,6)	FRA: In small schools the principal is the teacher. Principals have no authority over teachers. GER(W): Sensitive data!	44
45	ATSTAMP				45
46	ATSTAM1-6	BEL (1) FRA (2) SWI (1) VEN (1-4, 6)		FRA: Principals have no authority over teachers. FRA: Many single teacher schools.	46

SCHOOL QUESTIONNAIRE
Population A

No		Percentage Missing 10 - 20	21+	Special Comments	No
1	ACYEART	CAN FRA	VEN		1
1	ACYEARP	CAN	VEN		1
2	ACENRSB	CAN NOR SWI	VEN	NOR: Words *full time* caused problems	2
2	ACENRSG ACENRSD	CAN NOR SWI	VEN		2
3	ACENRGB	CAN	HNK VEN	HNK: Variable dropped from data set because of mistranslation Note: The ICC coded this variable to 0 when it was coded missing by the NRC and when ACENRSG had a non missing code	3
3	ACENRGG	CAN	HNK VEN	HNK: Variable dropped from data set because of mistranslation	3
4	ACTYSCH	VEN		BEL: Private schools are excluded; ACTYSCH=1 HUN: Note: has some private schools now ITA: Private schools are excluded; ACTYSCH=1 SIN: All schools in the Singapore system are state schools; ACTYSCH=1 SWE: ACTYSCH=1	4
5	ACTYCOM ACTYCOMD	VEN		All NRCs say break of 1 vs 2,3,4 ok for rural/urban split. ICE: 1 = country or country village 2 = coastal village or fishing town 3 = town 4 = Reykjavik NET: Less than 30,000 Between 30 000 and 100 000 Between 100 000 and 250 000 More than 250 000 NEZ: City of 1 million or more changed to a city of 50,000 or more NOR: 1 = under 1000 2 = 1000 - 15 000 3 = more than 15 000 4 = city of 1 million or more (doesn't exist) SIN: In view of the small land area the answers are (4) and (3) respectively	5

6	ACAVRE1-4	HNK (3) T/T (3-4) CAN (2,4) FRA (4) IND (1-4) ITA (2) NOR (4) POR (2,3)	HNK (2,4) POR (4) VEN (1-4)	HNK: Schools probably left blank when they meant *No*.	6
7	ACPACO	VEN			7
8	ACSCRE1-4 ACSCRED	CAN (2-4) FRA (2-4) HNK (1-3) NOR (2,3) SWE (3) T/T (2-4) POR (2-4) SWI (2,3) VEN (1)	HNK (4) ITA (2,3) VEN (2-4)	SWE: School Principals responsible for several schools and may not know for a particular school HNK: Schools probably left blank when they meant *No*.	8
	ACPNTEAD ACPSTEAD ACSEXD	DEN		ICE: Many small schools do not have special education ITA: Possible problem. *Full-time* translated as teachers who teach in classes with a longer instructional time for special education.	
9	ACSLIBC	CAN FIN GER (E) SWE ITA SWI T/T	HNK CYP VEN	CYP: None of the primary schools have school libraries. FIN: Some principals do not know the amount and did not want to count. ICE: Not all schools have a library SIN: school library max of 30,000 is OK - new schools; 32 is OK - new schools SWE: School Principals responsible for several schools and may not know for a particular school NRCs felt that for Qs 9 and 10 there were problems of meaning for *title* and *amount*. Treat with caution.	9
10	ACSLIBA ACSLIBD	CAN FIN GER (W) SWI T/T	CYP HNK ITA VEN	CYP: None of the primary schools have school libraries. FIN: Some principals do not know. ICE: Not all schools have a library SWE: School Principals responsible for several schools and may not know for a particular school NRCs felt that for Qs 9 and 10 there were problems of meaning for *title* and *amount*. Treat with caution.	10
11	ACSLIBOR	CAN FRA GER (E) IRL ITA SWI	HNK T/T CYP VEN	ICE: Not all schools have a library CYP: None of the primary schools have school libraries.	11

12	ACNFTTM	CAN	ITA VEN	ITA: *Full-time teachers* has been translated literally to mean teachers who teach in classes with a longer instruction time. Take care when using this variable. HUN: Librarians counted as teachers. SWE: figures slightly too high because *some* principals might have reported number of teachers in his/her "area" (which can include several schools) instead of in the particular school.	12
12	ACNFTTF	CAN ITA	VEN	ITA: *Full-time teachers* has been translated literally to mean teachers who teach in classes with a longer instruction time. Take care when using this variable. SWE:figures slightly too high because *some* principals might have reported number of teachers in his/her "area" (which can include several schools) instead of in the particular school.	12
13	ACNFSTM	CAN	VEN	GER(W): Some school principals omitted this question. There was no special teacher in their school SWE:figures slightly too high because *some* principals might have reported number of teachers in his/her "area" (which can include several schools) instead of in the particular school.	13
13	ACNFSTF	CAN	VEN	GER(W): Some school principals omitted this question. There was no special teacher in their school ITA: Handicapped children attend normal classes. To each of them one teacher is assigned to look after him T/T: possible in private schools SWE:figures slightly too high because *some* principals might have reported number of teachers in his/her "area" (which can include several schools) instead of in the particular school.	13
14	ACINSTH/M ACINSTD	CAN HNK IND	T/T VEN	POR: Value of 52 in private school only	14
15	ACWEOP ACPWEOD	BEL HNK FRA CAN GER (E) IND ITA T/T	GER (W) VEN	FIN: Normally ACWEOP should be 38 GER(E)/(W): 38 to 40 weeks is ok. Some school principals were not willing to estimate. ITA: Question was phrased in an incomplete way. Be careful when using this variable in Italy. NET: 40 weeks is legal maximum BEL: Some school principals answered their school is open the same number of weeks per year as other schools or "the official number of weeks". Others did not calculate and therefore did not answer!	15
16	ACDAYLO	HNK CAN GER (W) GER (E) GRC VEN		BEL: Max high because of teacher strikes. GRC: high values due to buildings being occupied by senior students NOR: Definition problems. Take care when using this variable SWE: ACDAYLO=ACDAYLO - 12 (12 days subtracted because vacation was included)	16

17	ACSTABS	DEN HNK NOR CAN GER (E) GER (W) ITA SWI	T/T VEN	GER: German teachers not prepared to give best estimates (or not willing to estimate)	17
18	ACSPPR1-5	HNK (1-5) CAN (1-5) GER (W) (2-5) GER (E) (2-5) ITA (2-5) T/T (2-5) VEN (1)	CYP (1-5) VEN (2-5)	BEL: If ACSPPR2,3,4 or 5=2 then ACSPPR1=1; else ACSPPR1=2 (ACSPPR1:filter variable was omitted in BEL) CAN: schools left blank for *No* HNK: schools left blank for *No* NET: 4-5 dropped from questionnaire by NRC CYP: Schools left blank for 'No'.	18
19	ACSPONS	CAN VEN			19
20	ACPRIMP	CAN VEN			20
21	ACPROB1-6	NOR (1-6) CAN (1-6) GER (W) (1-5) GER (E) (1-5)	VEN (1-6)	NOR: Distribution odd. No explanation	21
22	ACACTI1-8	CAN (1-8) CYP (1-8) VEN (1-8)		FRA: school principal cannot evaluate or train teachers (ACACT02,08) GRC: Nos. 2,3,4,6,8 are not applicable. Drop from analysis for Greece. SWE: *Using records of pupils' progress* is done by teachers, not by the headmasters CYP: Some items are not applicable to all schools	22
23	ACEVALF	CAN FRA GER (W) T/T VEN		GER(W): sensitive data	23
24	ACPROC1-6	NOR (1-6) CAN (1-6) GER (W) (1-6) VEN (1-6)	FRA (1-6) GRC (1-6)	FRA: school principal cannot evaluate or train teachers (ACACT02,08). GRC: Change *1* to 2; change *1-6* to *1* HNK: The instructions were changed to *Check one only*. Drop from analysis.	24

STUDENT QUESTIONNAIRE
Population B

No		Percent Missing		Special Comments	No
		10 - 20	21+		
1	BSAGEY/M BSAGED	CAN USA VEN			1
2	BSSEX	CAN VEN			2
3	BSUSLAN	VEN			3
4	BSFEDUC	BOT DEN HNK T/T DEN SWI VEN	BEL: CAN FRA GER(E) GER(W) NET NZL NOR SWE	BEL: For this and many other questions on this questionnaire it must be noted that many students were absent or did not answer the questions. DEN: Had problems of transforming Danish question to international code. Take care. GER(E)/(W): High missing values because Lower Saxony didn't allow this question to be asked and in other cases students refused to answer because of data protection laws. HNK: High missing because students did not know the answer (for sure) NET: Pupils don't know. SWE: Pupils don't know.	4
5	BSMEDUC BSEDUC	DEN HNK T/T SWI VEN	BEL CAN FRA NZL NOR SWE GER(W) GER(E) NET	BEL: Some students included pre-school in BSEDUCD DEN: Had problems of transforming Danish question to international code. Take care. GER(E)/(W): High missing values because Lower Saxony didn't allow this question to be asked and in other cases students refused to answer. HNK: High missing because students did not know the answer (for sure) NET: Pupils don't know. SWE: Pupils don't know	5
6	BSMEAL1-3 BSMEALD	BOT (2,3) DEN HNK T/T USA ZIM: (1)	VEN: (1-3) ZIM: (2,3)	ZIM: Many students misunderstood the question HNK: Some students misunderstood the question	6
7	BSJOB		SLO	BOT: If BSTIJOB >1 and <6 then BSJOB=2; if BSTIJOB=1 then BSJOB=1, (BSJOB: Filter variable) GER(W): High missing values because Lower Saxony didn't allow this question to be asked and in other cases students refused to answer. SLO: Missings probably = *no job*.	7

8	BSTUOB	BEL CAN T/T ITA	NET SLO VEN	GER(W): High missing values because Lower Saxony didn't allow this question to be asked and in other cases students refused to answer. NET: *Every day* included in question. POR: Not common to have a job. SLO: Students possibly left blank when meant *0*.	8
9	BSEDUCA	NET			9
10	BSNEWS			GER(W): Many cases *not administered* because Lower Saxony didn't allow this question to be asked.	10
11	BSTV				11
12	BSBOOKS	NET		GER(W): Many cases *not administered* because Lower Saxony didn't allow this question to be asked.	12
13	BSHOM01-10 BSHOMD	BEL (3,5) BOT: (1,3-5,7-10) CYP (9,10) IRE (7-10) NET (4, 6-8,10) NOR (2-4, 7-9) SWE (9) T/T (1,6-8) GER(W): (10) POR: (3-6,10) VEN: (2-4) ZIM: (1,7,8)	VEN: (1,5-10) ZIM: (2,6,9,10)	GER(W): Many cases *not administered* because Lower Saxony didn't allow this question to be asked. ZIM: We discovered, after data entry, that *missing* represented *none* for this question. So all missing here could be coded to *1* for data analyses. However, in BSHOMD the missings are coded to *not available*	13
14	BSSTUD1-8 BSSTUDD	BEL (7) BOT: (1,3-8) CAN (4,5,7) IRE (6) NET (2,3,7,8) NOR (1,3,5-7) SWE T/T (3,4) USA VEN: (4,6) ZIM: (5-7)	VEN: (1-3,5,7-8) ZIM: (1-4)	GER(W): Many cases *not administered* because Lower Saxony didn't allow this question to be asked. ZIM: We discovered, after data entry, that *missing* represented *none* for this question. So all missing here could be coded to *1* for data analyses. However, in BSHOMD the missings are coded to *not available*	14
15	BSASKRE	VEN			15
16	BSHWKF	VEN			16
17	BSHWKT				17

18	BSHWKTF	BEL CAN NET			18
19	BSHWKTT	VEN		HNK: Mistranslation; dropped from data set.	19
20	BSHWKRF BSHWKD BSHWKRD	BEL DEN HNK USA VEN		HNK: Some students were not willing to give a rough estimate.	20
21	BSHWKQ	VEN			21
22	BSBORBO	VEN			22
23	BSSRATE	VEN			23
24	BSWAY01-11	CAN (all)	VEN: (all)		24
25	BSSREH/M	BEL CAN ITA		PHI: Mispunched; dropped from data set. VEN: Mistranslation; dropped from data set.	25
26	BSWRITE	GER(W) VEN			26
27	BSSHSTO	BEL CAN	VEN		27
28	BSSHSCI	BEL CAN	VEN		28
29	BSSHFLA	BEL CAN	VEN		29
30	BSSHHS	BEL CAN T/T ZIM	VEN	ZIM: believes that all missing variables could be coded to 1 = *almost never* for analyses.	30
31	BSSHVOC	BEL CAN IRE NZL T/T ZIM	VEN	ZIM: believes that all missing variables could be coded to 1 = *almost never* for analyses.	31
32	BSSHREF	BEL CAN ZIM	VEN	- -	32

33	BSSHWOR BSSREADD	BEL CAN USA SLO	T/T VEN		33
34	BSBOOMY	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	34
35	BSBOORO	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	35
36	BSBOOSP	BEL CAN	VEN		36
37	BSBOOAV	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	37
38	BSBOOSF	BEL CAN	VEN		38
39	BSBOOHO	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	39
40	BSBOOPO	BEL CAN ITA ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	40
41	BSBOOHI	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	41
42	BSBOOHU	BEL BOT CAN T/T ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	42
43	BSBOOSC	BEL CAN T/T ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	43
44	BSBOOTR	BEL CAN T/T ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	44
45	BSBOONA	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	45
46	BSBOOMU	BEL CAN ZIM	VEN	ZIM: believes that all missing could be recoded to 1 = *almost never* for analysis purposes.	46

63	BSNEWBU BSNEWSD	BEL CAN USA	VEN		63
64	BSDOCTA	BEL CAN	VEN		64
65	BSDOCGR	BEL CAN	VEN		65
66	BSDOCMA	BEL CAN	VEN		66
67	BSDOCTI	BEL CAN	VEN		67
68	BSDOCDI	BEL CAN NET	VEN		68
69	BSDOCPO	BEL CAN GRC	VEN		69
70	BSDOCD	BEL CAN NET GRC	VEN		70
71	BSDOCIN BSDOCD	BEL CAN	VEN		71
72	BSSPL01-29	BEL BOT (12,16) CAN (1-9, 11-29) CYP (28) IRE:(16,17, 25-28) NET (all) NOR (13, 23,25,28) T/T (9,11, 12,16) GRC: (1-29) ITA: (3-29) ZIM: (1-29)	CAN (10) NZL: (10) NOR (10) T/T: (10) VEN: (1-29)	HNK: Mistranslation for BSSPL10; dropped from data set. NZL: Misprint in Q10 ZIM: Questions and items could have been fuzzy for some of these booklets. It is suggested that these items *not* be used in any analyses or reporting.	72

TEACHER QUESTIONNAIRE
Population B

No	Variable	Percent Missing 10 - 20	Percent Missing 21+	Special Comments	No
1	BTSEX				1
2	BTMOTON				2
3	BTSCHED	ITA			3
4	BTTETRA	BOT T/T	POR VEN	CAN: This question was asked in a very different way. Caution to be exercised when reporting. ITA: Misunderstanding; teachers considered as *teacher training* their pre-service teaching years	4
5	BTSECED BTEDUCD BTTRAIND			SWI: not reliable since it is not clear whether teacher training is part of sceondary education.	5
6	BTYEARS				6
7	BTFRRE1-9 BTFRRED	BOT:4,5 T/T: 4 BOT: (4,5) VEN: (3-5,8) T/T		BOT,CYP,HUN,PHI,SLO,THA,T/T,VEN,ZIM: Beware of *compliance effect.*	7
8	BTCLSIZ				8
9	BTOTLAN BTOTLAD			ITA: Confusion between dialect and other language; too high max number of students	9
10	BTNHELP BTNHELPD	POR	VEN	ZIM: Beware of max value which seems too high. Some schools stream their pupils and so sometimes the whole class is made up of slow learners.	10
11	BTRHELP BTRHELPD BTPHELPD	VEN	SLO T/T	SIN: State schools attend remedial classes even if not necessary ZIM: Some schools have streamed classes and all slow learners are put in one class	11

12	BTINSTH/M BTINSTD	FIN BEL HNK T/T BEL HNK T/T FIN GRC HKO POR SWE T/T	CAN VEN	BEL: Some teachers did not calculate and therefore did not answer. Others gave fanciful answers! CAN: Many teachers did not consider all subjects FIN: Many mother-tongue teachers do not know how many hours their students study other subjects. GER(E): Some teachers were not willing to convert school hours into clock hours HNK: Some teachers did not include all the subjects. Hence some answers were below the minimum value of 20. SWE: Teachers didn't know	12
13	BTTLANH/M BTTLAND BTPLAND	HNK VEN HNK HNK T/T HKO HUN: (M)	HUN: (H) THA		13
14	BTASME1-7	CYP: (1-7)	T/T:1-7	GER(E)/(W): 02; Multiple choice tests are not normally used in German schools. Some teachers omitted this item CYP: Some teachers did not rank the questions properly.	14
15	BTPRTE1-3				15
16	BTENCO1-5	T/T:3			16
17	BTACT01-17	BOT: 15, 13 BOT: (5,13) NET: (1-5) POR: (1,3,4,5,7-9,10-17) T/T: (8) VEN: (3,5-7,9,11-17)	T/T: 3,5, 8,9,14,16 VEN: (8)		17
18	BTAIM01-11	SLO: 1,4	T/T: 1-11 SLO: 2,4, 5-11		18
19	BTSTRA1-4	VEN: (4)		USA: Questions missing	19
20	BTSTRB1-8	VEN: (3,7)			20
21	BTTEXBO	CYP T/T VEN		HUN: Variable gang-punched; value was estimated.	21
22	BTSCHLI			NOR: Interpretation problem for library	22

23	BTVISIT	BOT CYP T/T BOT CYP POR		ICE: Not all schools have a library CYP: Missing probably means *never*	23
24	BTSLIBOR	BOT CYP BOT POR		ICE: Not all schools have a library CYP: Missing should be *No*	24
25	BTEVALU	T/T VEN		FRA: In France principal does not evaluate teacher NOR: Evaluation unclear term to teachers	25
26	BTPRIN1-6	T/T: 4,6 POR: (6) VEN: (1, 3-6)		FRA: Missing data is 2% too low SWE: Printing mistake in Q26a may have resulted in slight bias to results.	26
27	BTSTAMF				27
28	BTSTAM1-6	BEL: (1-3) T/T: 3 VEN: (1-6) FRA: (3)			28

SCHOOL QUESTIONNAIRE
Population B

No	Variable	Percent Missing		Special Comments	No
		10 - 20	21+		
1	BCYEART	BOT CAN NET THA ZIM	T/T VEN		1
1	BCYEARP	BOT CAN NET THA ZIM	T/T VEN		1
2	BCENRSB	BOT CAN NET NOR THA T/T ZIM	VEN		2
2	BCENRSO BCENRSD	BOT CAN NET NOR THA T/T ZIM	VEN		2
3	BCENRGB	BOT CAN IRE NET NOR THA T/T ZIM	VEN	HNK: Mistranslation. Drop variable from data set.	3
3	BCENRGG BCPNTEAD BCPSTEAD	BOT CAN IRE NET NOR THA T/T ZIM	VEN	HNK: Mistranslation. Drop variable from data set SIN: probably unusable = primary and secondary enrollment for pupils recorded but only secondary teachers. T/T: There is a scarcity of special teachers.	3
4	BCTYSCH	CAN THA VEN ZIM	BOT T/T	BEL: BCTYSCH=1 (Private schools were excluded) ITA: BCTYSCH=1 (Private schools were excluded) SIN: All schools in Singapore systems are state schools; BCTYSCH=1	4

5	BCTYCOM BCTYCOMD	BOT CAN THA VEN ZIM	T/T	ICE: 1=country or country village 2=coast village or fishing town 3=town 4=Reykjavik NET: up to 30,000 between 3,000 and 100,000 between 100,000 and 250,000 over 250,000 NZL: City of 1 million or more changed to a city of 50,000 or more SIN: In view of the small land area the answers are (4) and (3) respectively ZIM: Majority of schools are rural.	5
6	BCAVRE1-4	CAN HNK: (1-4) IRE (3,4) ITA (3) POR THA (1) ZIM: (1)	BOT: (1-4) THA (2-4) T/T VEN: (1-4) ZIM: (2-4)	HNK: Schools probably left blank when they meant *No*. ZIM (2,4): believes that missing could be recoded to 1.	6
7	BCSEXD				7
7	BCPACO	BOT CAN THA T/T VEN ZIM			7
8	BCSCRE1-8 BCSCRED	BEL: 6-8 CAN FRA: (5) HNK: (2,3,7) ICE (7) IRE: (2-6,8) ITA (2,3,5) NOR: (2,3 5-8) POR (2,4,5) SWE:(3, 5-8) SWI (6-8) THA (1-4) T/T: (1) VEN: (1) ZIM: (1,2,5,6)	BOT FRA: (3, 6-8) HNK: (4,5,6,8) IRE (7) ITA (6-8) POR(6-8) THA (5-8) T/T: (2-8) VEN: (2-8) ZIM: (3,5,7,8)	HNK: Schools probably left blank when they meant *No*. ICE: Student book clubs are not common in Icelandic schools. ZIM: believes that, for analyses, missing could be recoded to 1.	8
9	BCSLIBC	BEL HNK NET POR SWI	BOT CAN THA T/T VEN ZIM	ICE: Max ok; school and county library are one institution.	9

10	BCSLIBA BCSLIBD	BEL FRA GER(E) GER(W) HNK NOR SWI	BOT CAN IRE ITA NET THA T/T VEN ZIM		10
11	BCSLIBOR	CAN GER(E) IRE NET T/T	BOT IRE THA VEN ZIM		11
12	BCNFTTM	BOT CAN ITA THA T/T ZIM	NET VEN	GER: Maximum value ok. ITA: *Full-time teachers* translated literally as teachers who teach in classes with a longer instruction time. Take care when using variable.	12
12	BCNFTTF	BOT CAN ITA THA T/T ZIM	NET VEN	GER: Maximum value ok ITA: *Full-time teachers* translated literally as teachers who teach in classes with a longer instruction time. Take care when using variable.	12
13	BCNFSTM	BOT CAN GER(W) THA T/T ZIM	NET VEN	GER(W): Question sometimes omitted (no special teachers in school). ITA: Handicapped children attend normal classes; one teacher is assigned to look after each one -- high number of *special teachers*. POR: Max. not full time ZIM: Misinterpretation: subject-special teachers were called *special teachers* in some cases. Drop variable from analysis.	13
13	BCNFSTF	BOT CAN GER(W) THA T/T ZIM	NET VEN SLO	ITA: Handicapped children attend normal classes; one teacher is assigned to look after each one -- high number of *special teachers*. POR: Max. not full time ZIM: Misinterpretation: subject-special teachers were called *special teachers* in some cases. Drop variable from analysis.	13
14	BCINSTH/M BCINSTD	BOT CAN FRA HNK NET ZIM	THA T/T VEN	HUN: If BCINSTH >26 then BCINSTH=BCINSTH multiplied by 0.75 (for recoding school periods into hours and minutes)	14
15	BCWEOP BCPWEOD	GER(E) HNK THA T/T ZIM	BOT CAN GER(W) ITA VEN	FIN: Normally BCWEOP should be 38 GER(E)/W): Some school principals were not willing to estimate (38 - 40 is ok) ITA: Question was phrased in an incomplete way: *Number of weeks the school was open* without specifying *for the class being tested*. Take care when using this variable.	15

16	BCDAYLO	BEL BOT CAN GER(E) GER(W) HNK ICE NET PHI ZIM	BOT FRA THA T/T VEN	BEL: Max. high because of teacher strike. Some school principals did not answer; we think they hesitated to include the long period of strikes in their answer. GER(E)/(W): Some school principals were not willing to estimate. HNK: Some principals misunderstood the question and included all the holidays. Some answers exceed max. value. ICE: In the very small schools there is a high proportion of principals who change school. A new principal might not know this. SWE: BCDAYLO=BCDAYLO - 12 (12 days subtracted because vacation was included) ZIM: Max. high because of teacher strike	16
17	BCSTABS	BEL CAN DEN GER(E) GER(W) HNK NET NOR POR SWI ZIM	BOT THA T/T VEN	GER(E)/(W): Some school principals were not willing to estimate. HNK: Some principals were not willing to estimate.	17
18	BCSPPR1-5	GER(E)(2-5) GER(W) (2-5) NET (1-3) PHI (2-5)	BOT BEL CAN: (1-5) CYP (2-5) HNK IRE (1-5) NET (4,5) SIN (2-5) SWI (2-5) THA (1-5) T/T VEN: (1-5) ZIM: (1-5)	Some countries did not print the first option. If nothing checked, the ICC recoded to *No special programs*. ZIM:believes that missing had 'no special program' CYP: Missing means 'no special program'	18
19	BCSPONS	CAN THA ZIM	BOT T/T VEN		19
20	BCPRIMP	CAN THA ZIM	BOT T/T VEN		20
21	BCPROB1-6	CAN:(6) NET (6) NOR SWI (1-5) THA (6) T/T (6) ZIM: (1-6)	BOT CAN:(1-5) NET (1-5) THA (1-5) T/T (1-5) VEN: (1-6)		21
22	BCACTI1-8	CAN NET SWI THA T/T ZIM: (1-8)	BOT CYP VEN: (1-8)	NOR: Interpretation problem: only most important checked.	22

23	BCEVALF	CAN GER(W) NET SWI THA T/T ZIM	BOT VEN	GER(W): Sensitive data	23
24	BCPROC1-6	CAN GER(W) NOR POR THA ZIM	BOT GRE NET SWI T/T VEN: (1-6)		24