DOCUMENT RESUME

ED 133 707 CS 003 148

AUTHOR	Postlethwaite, T. Neville
TITLE	A Selection from the Overall Findings of the I.E.A. Study in Science, Reading Comprehension, Literature, French as a Foreign Language, English as a Foreign Language and Civic Education. IIEP Occasional Papers No. 30.
INSTITUTION	United Nations Educational, Scientific, and Cultural Organization, Paris (France). International Inst. for Educational Planning.
PUB DATE	76
NOTE	46p.
AVAILABLE FROM	International Institute for Educational Planning, 7-9 Rue Eugene Delacroix, 75016 Paris, France (4.00 francs)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Academic Achievement; Civics; Comparative Analysis; *Comparative Education; *Cross Cultural Studies; Developing Nations; *Educational Research; Elementary Secondary Education; English (Second Language); French; Literature; *Performance Factors; Reading Achievement; Science Education
IDENTIFIERS	IEA; *International Assn Evaluation Educ Achievement

ABSTRACT

The International Association for the Evaluation of Educational Achievement (IEA) conducted a six-subject survey to examine achievement in 23 nations with 15 different languages of instruction. The technique used was cross-sectional, with sampling at three different educational levels. In addition to achievement testing in science, reading comprehension, literature, French, English, and civic education, extensive information on input and process variables was collected from students and school personnel. This report discusses selected findings from the analyses of between-country differences in achievement as well as from within-country analyses of the relationships between school outcomes and the input and process variables. (AA)

IIEP Occasional Papers No. 30

A SELECTION FROM THE OVERALL FINDINGS OF THE IEA STUDY IN SCIENCE, READING COMPREHENSION, LITERATURE, FRENCH AS A FOREIGN LANGUAGE, ENGLISH AS A FOREIGN LANGUAGE AND CIVIC EDUCATION

T. Neville Postlethwaite

Unesco: International Institute for Educational Planning

IIEP Occasional Papers

The studies in this series include papers contributed by the Institute's staff, visiting fellows, interns and consultants. Some of the studies have originally been prepared as part of the training programme of the Institute; others have previously appeared as working papers for the Institute's seminars and symposia. All of them, in the Institute's views, are of sufficient interest to merit being re-issued and distributed on a wider scale.

By their very nature these papers are less formal and have not been given the full editorial processing customary for IIEP 'official publications'.

Printed in France by the
International Institute for Educational Planning
7-9, rue Eugène-Delacroix, 75016 Paris
November 1973

CONTENTS

INTRODUCTION

In the OECD paper "A Framework for Educational Indicators to Guide Government Decisions" (OECD, 1973)1/, it is suggested that the International Association for the Evaluation of Educational Achievement (IEA) study produces indicators of knowledge and skills transmitted in the achievement scores in a series of subject areas in twenty-three countries. This is true in that the cognitive outcomes of schooling as measured by IEA cover not only the means and ranges of knowledge, but also of skills and attitudes. However, IEA has also attempted, through data collection and analysis in a series of countries, to identify those factors (indicators) which are inputs to schools and educational processes in schools which in some senses influence, or at least are associated with, the various levels of the output of the schools. Indeed, IEA's major thrust has been to identify the factors associated with variation between students and also between schools within a population within a country.

This paper presents some of the main findings of the most recent IEA study (The Six-Subject Study)(2-7). First of all, let us recall the data collection of the IEA study. Table 1 presents the different subjects taken at the various population levels in each of the countries participating in the study.

The definitions of the three populations are as follows:

Population I - all students in full-time schooling aged 10:00 - 10:11 at the time of testing.

Population II - all students in full-time schooling aged 14:00 - 14:11.

Population IV - all students in the terminal year in full-time secondary education programmes which were either pre-university programmes or programmes of the same length. (The interpretation of this definition of this group varied as well as the percentage of an age-group in this target population.)

The above definitions held true for Science, Reading Comprehension, Literature and Civic Education. In the case of English and French as foreign languages, a further condition was that the students should be currently studying the language and have studied it for at least two years. Indeed, when examining the home background distributions of the students taking English and French, it was clear that where English was part of the curriculum it was studied by more or less all of the students in full-time schooling, whereas French tended to be restricted to smaller segments of the population.

1/ OECD, A framework for educational indicators to guide government decisions, OECD, Paris, April 1973.

Table 1. Subject areas and populations tested by country

	Science				Reading Comprehension				Literature			French				English			Civic Education			
	I	II	III[a]	IV	I	II	III[a]	IV	II	III[a]	IV	I	II	III[a]	IV	II	III[a]	IV	I	II	III[a]	IV
Australia		x		x																		
Belgium (French)	x	x		x	x	x		x	x		x					x		x				
Belgium (Flemish)	x	x		x	x	x		x	x		x					x		x				
Chile	x	x		x	x	x		x	x		x				x			x				
England	x	x	x	x	x	x	x	x	x	x	x		x	x	x							
Fed. Rep. of Germany	x	x		x	x	x		x								x		x	x	x		x
Finland	x	x	x	x	x	x	x	x	x	x	x					x	x	x		x	x	x
France				x																		
Hungary	x	x		x	x	x		x									x	x				
India	x	x		x	x	x		x														
Iran	x	x		x	x	x		x	x		x									x		x
Ireland																				x		x
Israel					x	x		x								x		x	x	x		
Italy	x	x		x	x	x		x	x		x					x		x	x	x		
Japan	x	x																				
Netherlands	x	x		x	x	x		x					x		x	x		x	x	x		x
New Zealand		x		x		x		x	x		x		x		x					x		x
Poland																						
Rumania													x		x							
Scotland	x	x		x	x	x		x					x		x							
Sweden	x	x	x	x	x	x	x	x	x	x	x			x	x	x	x	x			x	x
Thailand	x	x		x												x	x	x				
U. S. A.	x	x		x	x	x		x	x		x	x	x		x					x		x

a) Population III was not an internationally defined population. Each participation research centre was able to define a population somewhere between Population II and IV to test with Population II for national purposes. The analyses were national and not international.

Probability samples of schools and students within schools were drawn for each level for each subject (or group of subjects) within each country. The way in which the complex samples were drawn, together with the resultant standard errors of sampling and Design Effects (DEFFS) for selected variables have been shown in detail elsewhere. 1/

Three to five years of development work were required to construct the outcome (i.e. performance) measures in each subject area at each level. A list of the major outcomes classified by cognitive and affective outcome is given in Appendix I. The resultant tests had good reliabilities and were judged by the subject-matter panels in the various countries to be appropriate for testing what was, in general, meant to have been learned in school up to that point of testing. The test formats used included multiple choice, open-ended and file-in formats. For foreign languages tape recorders were used for Listening Comprehension and Speaking. Both mark-sense cards and punched cards were used for data recording.

A series of questions was also given to the students, teachers and school principals in questionnaire booklets. The variables (representing inputs to and processed in the school) for any one subject at any one level numbered approximately some 200 to 500. The sheer amount of information collected (over 150 million pieces) makes this study one of the largest ever in the field of Social Science. As an example, Appendix II lists the student, teacher and school variables associated with French. These are the variables used in every country. However, each country had the opportunity to ask further questions which were of specific interest for its own system. Hence, although this paper will present the findings of the international analyses, each National Centre will be publishing a national report using a fuller set of data (three examples of national reports are those of Australia 2/, Japan 3/ and Sweden 4/). A forthcoming issue of the Comparative Education Review will contain a series of articles from most National Centres, each concerning a special national analysis undertaken on a current national problem in education. 5/

1/Peaker, G.F., An empirical study of education in twenty-one countries : A technical report, International Studies in Evaluation, VIII, Almqvist & Wiksell, Stockholm (in press).

2/Rosier, Malcolm J., Science achievement in Australian secondary schools, Australian Council for Educational Research, Melbourne, Australia, 1973.

3/Shimada, S. et al., International survey of science education : Report of Japanese National Commission of IEA, Vol. I, Results of Japanese Survey, NIER, Tokyo, March 1973.

4/Husén, T. et al., Naturorienteranda ämnen, Almqvist & Wiksell, Stockholm, 1973.

5/Postlethwaite, T. Neville (ed.), Comparative Education Review, Vol. XVIII, No. 2, June 1974 (in press)

The aims of the research were to identify those factors accounting for differences between countries, between schools and between students. The technique used was a cross-sectional survey, albeit at three different levels. This, therefore, implies that this is a description of education as it is an not as it might be. It also means that there are no direct measures of past events - surrogate measures had to be used. A longitudinal study would have produced such measures, but for various reasons the IEA felt that the time and money required for such a study were beyond its reach. The input and process variables data were collected from students and school personnel. The limitations of questionnaire procedures are that, on the one hand, the responses may not be accurate and, on the other hand, may provide only an incomplete and often indirect indication of the phenomenon in which one is interested. These constraints are important to note and bear in mind when interpreting the results.

Before passing to the results, it should be mentioned that the fact that this study has been conducted in twenty-three nations with fifteen different languages of instruction is in itself something of a tour de force. In spite of translation problems and the difficulties of planning and executing such a project, not to mention the research skills required by so many people in so many countries, it has been possible to measure outcomes in a reliable way and to make meaningful comparisons. There was a certain variation in the quality of the research competencies possessed by the participating centres and the standard form of research proved to be very effective in raising the standards to the highest common factor. It also allowed each nation's educational system to be seen in an international context.

I. SELECTED FINDINGS

1. Between countries

With the limited number of observations of country mean scores it is not possible to undertake between-country multivariate analysis. However, differences in means and distributions on single variables are of interest as well as bivariate relationships. IEA has always stressed that it is not conducting a 'cognitive olympics' and that great care should be exercised when comparing country mean scores. First, although international tests have been constructed and each country was happy to use them they were not 100 per cent valid for each country and some countries may well stress some objectives which were not tested by the tests. Secondly, the proportion of an age-group in school varies from country to country even at the same population level, hence making comparison difficult. This being said, some of the test scores between countries are of interest. In what follows, four different modes of presentation have been used. For Reading Comprehension, the pass-fail rate on particular items is given. For Literature, a profile of four different aspects of literature is given. In Science, the increment in General Science performance from one population level to another is presented. For the other subjects, means and standard deviations are presented.

Reading

Perhaps the most dramatic finding is the very large difference in performance between developed and developing nations. If we take Reading, we find that differences among developed countries are fairly modest. However, with complete consistency, the three developing countries fall far below those that have a relatively high level of economic development and a long-standing tradition of universal education.

To exemplify this somewhat, consider one selected passage from the battery of reading tests for 10-year-olds:

Population I (Passage and items)

One of the most interesting birds I have seen is the Indian Tailor Bird. It is a small olive green bird that does not look at all unusual, yet it has a most unusual way of making its nest. The birds work together in pairs. First they find a leaf, the right size, and make holes along the edges with their beaks. Through these holes they thread grass. One bird pushes the thread from the outside, while the other bird sits in the nest and pushes it back until the edges of the leaf are sewn together to make a kind of bag, still hanging on the tree, in which the Tailor Bird lays its eggs.

A selection from the overall findings of an IEA study

What does the Tailor Bird use in place of thread?

- grass
- string
- spider web
- thorns

The Tailor Birds are interesting because they:

- are small and olive green in colour
- live in pairs
- make their nests in a special way
- fly very fast

The Tailor Bird got the name because it:

- is a small bird
- looks unusual
- can sew
- has a beak shaped like a needle

The Tailor Birds make their nests:

- from leaves
- in a hole in a tree
- in the tall grass
- with a lining of grass

The person who wrote about Tailor Birds was trying to:

- give you some new information
- tell you a story
- get you to share his feelings
- keep you guessing on how the story will come out.

Clearly, the little paragraph shown here is at a level that presents difficulty to a substantial fraction of 10-year-olds in every country. Considering all fourteen countries and all five items, the typical failure percentage runs between 35 per cent and 40 per cent (Table 2). The median percentage goes as low as 26 per cent in Finland, but is 48 per cent in Chile, 58 per cent in India and 65 per cent in Iran. When one considers that random marking would be expected to give only 75 per cent of error on these four-choice items, it becomes clear that even this passage is pushing the limit of competence of most 10-year-olds in these three countries.

Similar results were obtained for Populations II and IV. In one passage for Population II the median error rate was approximately 80 per cent but in Chile, India and Iran the rate was 47, 63 and 61 per cent respectively. For Population IV for one passage the median error rate was 25 per cent but for the above-mentioned developing countries was 44, 66 and 68 per cent respectively (see pages 133-139 in the Reading Comprehension Report). 1/

1/ Thorndike, Robert L., Reading Comprehension education in fifteen countries : An empirical study, International Studies in Evaluation, III, Almqvist & Wiksell, Stockholm, 1973.

Table 2. Percent failing - Population I

Country	Item 1	Item 2	Item 3	Item 4	Item 5
Belgium (Flemish)	5	44	53	40	43
Belgium (French)	12	27	52	27	30
Chile	16	48	53	41	60
England	10	37	37	27	37
Finland	16	36	26	22	31
Hungary	8	40	32	39	48
India	16	66	58	54	61
Iran	56	75	65	56	66
Israel	12	49	42	46	56
Italy	8	32	80	23	32
Netherlands	8	28	48	41	38
Scotland	11	38	35	30	38
Sweden	10	35	39	27	36
U.S.A.	13	35	46	38	33

A further indication of the base level of competence in the developing countries is provided by error rate on the first nine items of the test of reading speed. These are items such as:

"Peter has a little dog. The dog is black with a white spot on his back and one white leg. The colour of Peter's dog is mostly (i) black, (ii) brown, or (iii) grey."

The percentage of wrong answers on items such as this is as follows:

Country	10-year-olds	14-year-olds
Belgium (Flemish)	8	3
Belgium (French)	9	2
Chile	26	16
England	9	4
Finland	11	8
Hungary	20	8
India	36	33
Iran	52	20
Israel	17	9
Italy	11	9

Netherlands	8	4
New Zealand	(not taken)	4
Scotland	10	3
Sweden	7	2
U.S.A.	11	4

In the European countries, a typical error rate on these items is about 10 per cent for 10-year-olds and 4 per cent for 14-year-olds. With these values, one must contrast percentages of 26, 36 and 52 for 10-year-olds of the developing countries (Chile, India, Iran) and 16 per cent, 33 per cent and 20 per cent for the 14-year-olds. Admittedly, this material was given as a speed test. But it was also given as a reading test. If a substantial proportion of the students in a school system have real difficulty in reading these materials, one must question whether any more than a minimal level of literacy has been achieved in that school system.

Literature

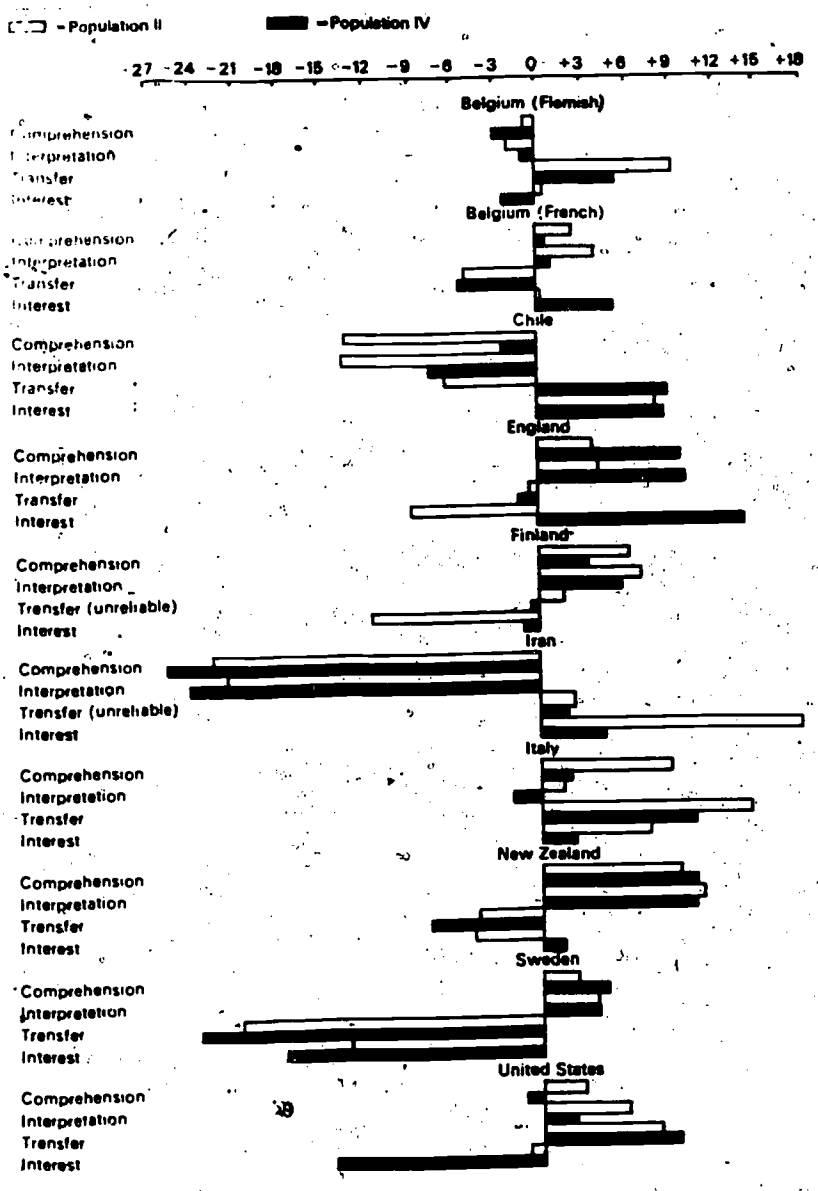
Literature was only tested by Populations II and IV in ten countries. Figure 1 presents a country profile of four scores:

- Comprehension is a reading comprehension score of literary prose.
- Interpretation is an interpretation (primarily inference) score of literary prose.
- Transfer is a measure of the extent to which a student brings his literary experiences to bear on the rest of his life and vice-versa.
- Interest is a measure of the amount of interest a student expressed in reading literature.

It can be seen that the achievement scores (that is, Comprehension and Interpretation) are lower in Chile and Iran than in the participating developed countries. The Transfer and Interest scores are presented as an illustration of non-cognitive outcomes of an educational system. However, it should be pointed out that probably the most important measure in Literature was the response-preference measure, where each student responded to the measure three times - once in a questionnaire about literary works generally and twice as response to two literary extracts.

The response-preference measure sought to determine the characteristic way by which a group of students approached a particular literary work or literature in general. For example, some students might be more concerned with the historical aspects of the work; other students might be more concerned with its aesthetic form; still others with its theme or meaning; still others with its affective value. Students in both Populations II and IV were asked to indicate the questions they thought most important following a reading of each of two short stories and with reference to literature in general.

Figure 1. Cross-national score profiles : standard scores



A multivariate analysis of variance indicated that students' response preferences were strongly influenced by the nature of the selection that they read, by their age (that is, by which population they were in) and by their country. This analysis as well as an examination of the response preference patterns for the Population II and Population IV students led the committee to conclude that a student's response is determined by the nature of what he reads and by the educational system. During the years of secondary school, students in all of the countries examined appeared to have learned to focus their critical attention to certain aspects of literary works. A comparison of the students' response preferences with those of the teachers of literature in the schools shows that the students learned to approximate the response patterns of their teachers. The study of response preference also indicates quite clear similarities and differences between countries as the students in those countries approach literary work. There emerge certain quite clear cultural patterns. Students in Chile, England and New Zealand are similar to each other in their interest in formal aspects of literary works. Students in Belgium, Finland and the United States of America are marked by their interest in the thematic aspects of the work. Students in Italy are primarily interested in the historical aspects of the work.

Science

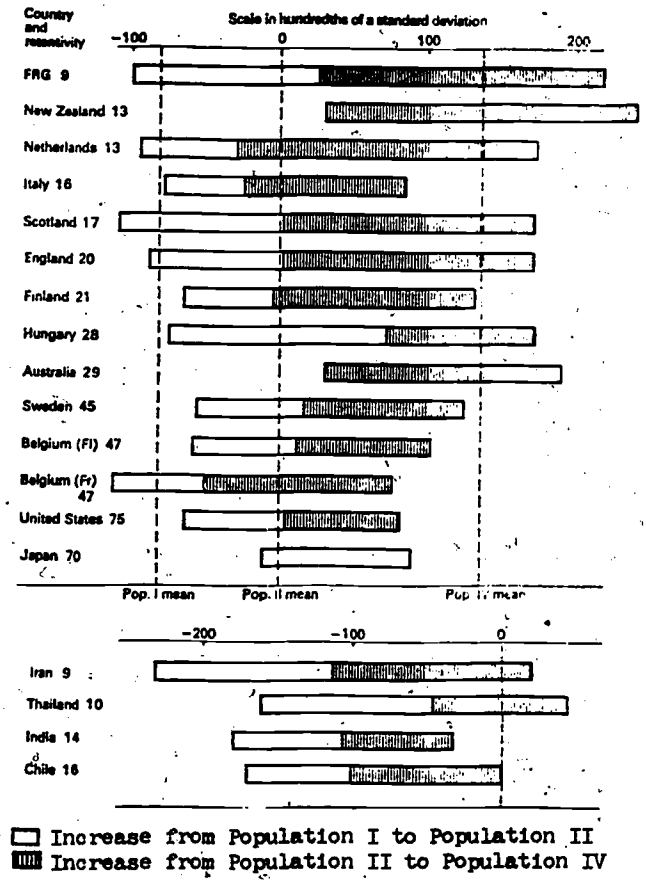
Because there were anchor items between the tests for the various populations it was possible to bring the population test scores on to a common scale and examine the increment from one population level to another in Science, as measured by the IEA tests. Figure 2 presents the increases. Again, the developing countries are presented separately - note the different position of the zero point. The 'retentivity' figure given next to each country's name is an estimate of the percentage of an age-group still in school in the pre-university year grade. It is also at Population IV level that the differences in interpretation of what this population was has operated to enhance differences between countries. Note that Australia and New Zealand did not test 10-year-olds and Japan did not test the final year grade group. France is missing from the table since it only tested one population, namely "la classe terminale" and hence no increment measure was available. The left-hand side of the white bar represents the scaled mean score at age 10, the right-hand side the scaled mean score at age 14 and the right-hand side of the crossed bar represents the scaled mean score of the pre-university year group.

In the Science report, the authors have broken down the total Science score into a series of sub-scores: earth science, physics, chemistry, biology, and practical science on the one hand, and on the other hand a classification of educational objectives such as functional knowledge, comprehension of a scientific principle, application of knowledge for problem-solving and analysis, synthesis and evaluation. 1/ The country-by-country profiles of such sub-scores help to identify the relative

1/ Bloom, B.S. (ed.), Taxonomy of educational objectives : the classification of educational goals, Handbook I, Cognitive Domain, David McKay Company Inc., 1956.

strengths and weaknesses of the students performance. Further to this, the item-by-item analyses of the Science items show not only the percentage of the population answering the item correctly (which can in a certain sense be construed as a criterion-referenced test) but also the types of wrong answers typically produced. Both the profiles and item analyses are extremely useful feedbacks to the curriculum developers as a check to what extent performance matches the global and detailed objectives set for Science education in a particular country.

Figure 2. Increase in level of performance in Science from the 10-year-old level to the terminal secondary school stage



Each participating research centre will be carrying out a set of special national analyses. Where different curricula exist for different school types or regions within a country, such profile and item analyses will be important as feedback to the curriculum developers. In a centralised system of education the feedback links are, of course, easier to forge or maintain than in a decentralised system but it is hoped that all research centres will create the necessary links, however complex this may be.

Experience suggests that boys show a greater "interest in Science" than girls and perform better on Science tests. These impressions receive powerful support from the IEA study. These differences exist across all countries and the differences increase as the students grow older. Thus, at ten years of age, boys performed on average about one-quarter of a standard deviation higher than girls, half a standard deviation at 14 years of age, and three-quarters of a standard deviation higher in the terminal year of secondary education. Boys are more strongly attracted to the Physical Sciences and girls to the Biological Sciences. The score differences in Biology were less than in the other branches. This is clearly a problem which deserves earnest attention in the near future in all countries.

French as a Foreign Language 1/

Table 3 presents the means, standard deviations and numbers of students for the four skills in French for each participating country at both Populations II and IV levels. When examining these scores, it

Table 3. Means, standard deviations and numbers of students for scores in four major skills of French

	Reading			Listening			Total Writing			Speaking Fluency		
	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.
Population II												
England	9.3	9.2	2 076	9.6	9.7	455	38.4	27.0	437	53.6	32.5	102
Netherlands	12.8	7.8	1 545	-	-	-	-	-	-	-	-	-
New Zealand	12.2	8.3	1 794	4.5	6.3	1 790	45.6	27.3	909	58.8	30.3	185
Rumania	26.8	8.7	2 271	24.1	12.5	117	60.1	31.8	1 825	97.2	25.4	64
Scotland	13.1	9.8	835	8.8	8.4	834	49.2	31.4	826	64.4	32.5	193
U.S.A.	7.6	7.5	4 177	6.1	7.2	1 927	46.6	27.7	2 329	59.2	27.9	174
Population IV												
Chile	6.6	5.7	1 440	3.2	5.1	173	28.0	21.9	149	28.3	25.1	132
England	32.1	5.2	702	27.9	7.3	180	86.4	28.5	181	111.4	34.1	83
Netherlands	26.3	6.1	1 753	-	-	-	-	-	-	-	-	-
New Zealand	27.4	5.4	363	12.8	8.3	361	73.2	26.0	363	93.1	29.3	150
Rumania	28.6	8.7	2 247	26.2	11.9	77	62.4	33.0	2 073	137.6	39.4	64
Scotland	25.2	7.0	972	17.3	8.9	972	79.3	30.1	966	96.6	32.3	237
Sweden	19.6	9.4	1 755	19.3	10.2	1 742	67.4	28.2	1 670	107.9	41.8	186
U.S.A.	17.5	9.5	3 069	13.8	11.5	1 370	57.6	27.3	1 776	91.8	37.1	178

1/ Carroll, John B., International Studies in Evaluation. V : French as a Foreign Language in Seven Countries. An Empirical Study. Almqvist and Wiksell, Stockholm (in press).

is important to bear in mind what Carroll 1/ has written elsewhere. The national mean scores were in general highly correlated with the average number of years that French had been studied and as a corollary to this average scores related negatively to the average grade in which the study of French began.

English as a Foreign Language2/

Table 4 presents the means, standard deviations and numbers of students for the participating countries in the four major skills of Reading Comprehension, Listening Comprehension, Writing and Speaking Fluency. Blanks denote non-participation. The large standard deviations for Population II Reading are noteworthy.

Table 4. Means, standard deviations and numbers of students for scores in four major skills of English

	Reading			Listening			Total Writing			Speaking Fluency		
	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.
Population II												
Belgium (French)	22.8	15.2	687	15.4	8.4	671	12.5	6.8	697	20.9	7.4	162
Fed. Rep. of Germany	29.9	16.2	1 074	-	-	-	-	-	-	-	-	-
Finland	16.7	17.8	2 100	12.1	9.9	2 118	9.7	8.4	2 087	19.8	9.7	200
Israel	26.3	18.7	1 065	-	-	-	-	-	-	-	-	-
Italy	18.6	17.3	791	9.1	6.0	694	11.3	7.0	731	19.8	7.7	71
Netherlands	37.4	16.4	2 090	-	-	-	-	-	-	-	-	-
Thailand	18.4	13.4	1 951	-	-	-	-	-	-	-	-	-
Population IV												
Belgium (French)	24.2	10.8	1 440	16.9	7.0	1 394	20.5	6.0	1 391	24.2	6.9	192
Chile	10.6	9.8	2 118	7.1	4.5	231	7.2	4.8	183	14.4	6.5	94
Fed. Rep. of Germany	41.3	6.1	1 374	-	-	-	-	-	-	-	-	-
Finland	35.6	8.3	2 310	23.2	6.1	2 296	26.6	3.2	2 238	29.9	8.7	279
Hungary	20.1	11.2	1 063	16.3	7.3	1 063	18.2	7.2	1 063	30.3	12.3	119
Israel	27.4	12.9	604	-	-	-	-	-	-	-	-	-
Italy	20.2	12.6	324	14.7	7.6	314	18.9	6.0	322	27.0	7.2	28
Netherlands	42.9	6.5	1 561	-	-	-	-	-	-	-	-	-
Sweden	39.5	9.0	1 626	26.2	5.4	1 602	25.3	4.6	1 544	34.0	8.8	197
Thailand	19.7	9.1	937	-	-	-	-	-	-	-	-	-

1/ Carroll, John B., Factors accounting for between-student, between-school and between-nation differences in performance in French as a Foreign Language. A presentation at AERA Convention, New Orleans, Louisiana, February, 1973.
2/ Lewis, E. Glyn, International Studies in Evaluation, IV : English as a Foreign Language in Ten Countries. An Empirical Study. Almqvist and Wiksell, Stockholm, (in press).

Civic Education

As can be seen from Appendix I, there were many outcomes of Civic Education measured - attitude scores in addition to the Cognitive Test Total and subscores. In Table 5, means for Civics Cognitive Test Total are compared for different countries within population. (Comparisons between age population groups are not justified because the Cognitive Tests were composed of different items for each population.) The phenomenon noted in the other subjects, of considerably lower scores for less developed nations, holds true in Civics for Iran. The other countries' mean scores are fairly closely clustered; it is interesting to note however that at both the Population II and Population IV levels, students in Ireland achieve relatively low mean scores. The anti-authoritarianism scale, which is derived from a series of attitude items concerning democratic values, administered in the same form at each age level can be compared across population. Anti-authoritarianism is a good representation of the first major attitudinal factor derived in factor analysis on the 14-year-old (Population II) students and was used as a criterion in the regression analysis (to be briefly summarized later in this paper and completely presented in the Civic volume) 1/. Although the between-country and between-population differences are interesting for speculation, of more interest are patterns of attitude scale difference which are too complex to be examined here. For example, in some countries high anti-authoritarianism is combined with a low sense of citizen efficacy; in other countries both types of attitudes are low.

Other between-country differences

It is not possible to carry out a multivariate analysis between countries (with so few degrees of freedom) to identify the factors associated with differences between national mean scores. However, certain differences in certain factors are striking and provide interesting leads. Four examples are given below:

(i) Opportunity to learn in Science

In each school in which Science was tested the Science teachers in that school were asked collectively to rate each item in the Science tests as to the appropriate percentage of students in the target population who had had "the opportunity to learn" the substance tested in the item. The scale was:

- all students	4
- more than 75 per cent of the students	3
- between 25 and 75 per cent of the students	2
- less than 25 per cent of the students	1
- none of the students	0

1/ Farnen, R.F., Marklund, S., Oppenheim, A.N., and Torney, Judith V., International studies in evaluation, VI : Civic Education in ten countries - an empirical study, Almqvist and Wiksell, Stockholm (in press).

Table 5. Means, standard deviations and numbers of students for scores in three outcomes in Civic Education

	Total cognitive score			Anti-authoritarian (attitude)		
	$\bar{X}$	σ	No.	$\bar{X}$	σ	No.
Population I						
Fed. Rep. of Germany	13.9	8.5	1 070	3.4	0.6	1 051
Israel	-	-	-	3.3	0.6	402
Italy	18.6	11.1	2 390	3.5	0.6	2 325
Netherlands	15.6	7.6	1 746	3.4	0.6	1 730
Population II						
Fed. Rep. of Germany	26.0	8.7	1 313	4.0	0.5	1 275
Finland	24.5	9.8	2 370	3.9	0.6	2 356
Iran	9.7	5.9	2 204	3.2	0.5	2 032
Ireland	20.8	10.4	834	3.9	0.6	817
Israel	25.6	9.6	1 039	3.7	0.6	952
Italy	22.9	9.4	930	3.9	0.6	918
Netherlands	27.3	7.9	1 685	4.1	0.5	1 645
New Zealand	24.3	9.6	1 983	4.1	0.5	1 969
U.S.A.	24.7	9.9	3 186	4.0	0.5	3 119
Population IV						
Fed. Rep. of Germany	28.2	5.8	1 163	4.6	0.4	1 176
Finland	26.1	6.6	2 315	4.3	0.4	2 282
Iran	6.8	4.9	2 159	3.5	0.5	2 028
Ireland	16.9	8.4	786	4.2	0.4	782
Netherlands	25.5	6.5	1 296	4.3	0.5	1 203
New Zealand	28.4	7.1	1 665	4.3	0.4	1 668
Sweden	27.0	7.8	1 723	4.5	0.5	1 636
U.S.A.	21.4	9.7	3 016	4.1	0.5	2 928

Despite the crudity of the measure it shows a considerable difference between countries when summed to a national level. Figure 3 presents graphs of opportunity to learn against Science total test scores for Populations I, II and IV. The relationships are striking. However, one cannot help wondering if, for example, in Population II, the tests are more appropriate for Japan than for the Netherlands and if the Netherlands' students are learning other things in Science not measured by the Science tests. This is for the curriculum developers to decide, but it should be pointed out that the test construction was a lengthy and detailed exercise involving Science educators in all countries. All countries expressed themselves as agreeing that the tests were a reasonable general measure of Science as taught at the particular level. The lack of relationship for Population I may be attributed to the unstructured curriculum in Science in most countries at that stage of schooling.

(ii) Social bias

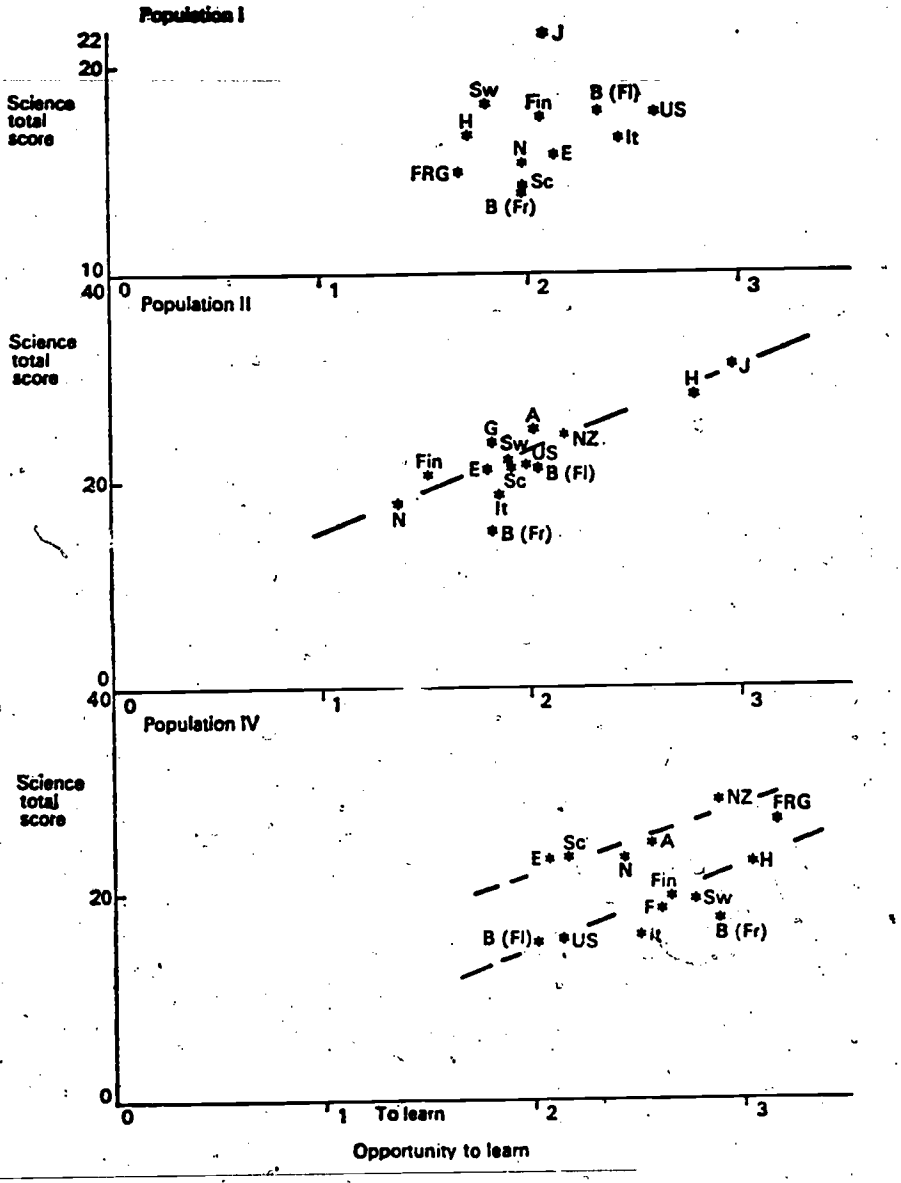
There has been much discussion recently concerning the equality of participation, in the whole school system, of students from all social classes. Each research centre utilised its own social class set of categories. (This differed from the IEA mathematics study, 1962-1967, where one international set of categories was used.) Therefore, although the categories are not exactly the same from country to country and therefore it is difficult to make comparisons between countries, it is possible within countries. Table 6 presents the percentage of students for each population in selected centres occupational categories for seven countries.

The progression from population to population is of interest and the difference between England, the Federal Republic of Germany and the Netherlands on one hand, and Finland, Hungary, Sweden and the United States of America on the other hand, is striking. It should be recalled that the Federal Republic of Germany only tested students in the Oberprimaria in the Gymnasium.

Table 6. Percentage of students for each population in selected categories of father's occupation (Science, Reading Comprehension and Literature)

	Population I		Population II		Population IV	
	Professional & managerial	Unskilled & semi-skilled	Professional & managerial	Unskilled & semi-skilled	Professional & managerial	Unskilled & semi-skilled
England	16	21	14	14	38	5
Fed. Rep. of Germany	13	7	14	8	49	1
Finland	9	35	10	34	20	15
Hungary	15	43	20	36	38	18
Netherlands	26	12	20	12	55	5
Sweden	23	31	26	27	35	15
U.S.A.	24	18	31	16	34	14

Figure 3. Graphs of opportunity to learn against Science Total Test Scores for Populations I, II and IV

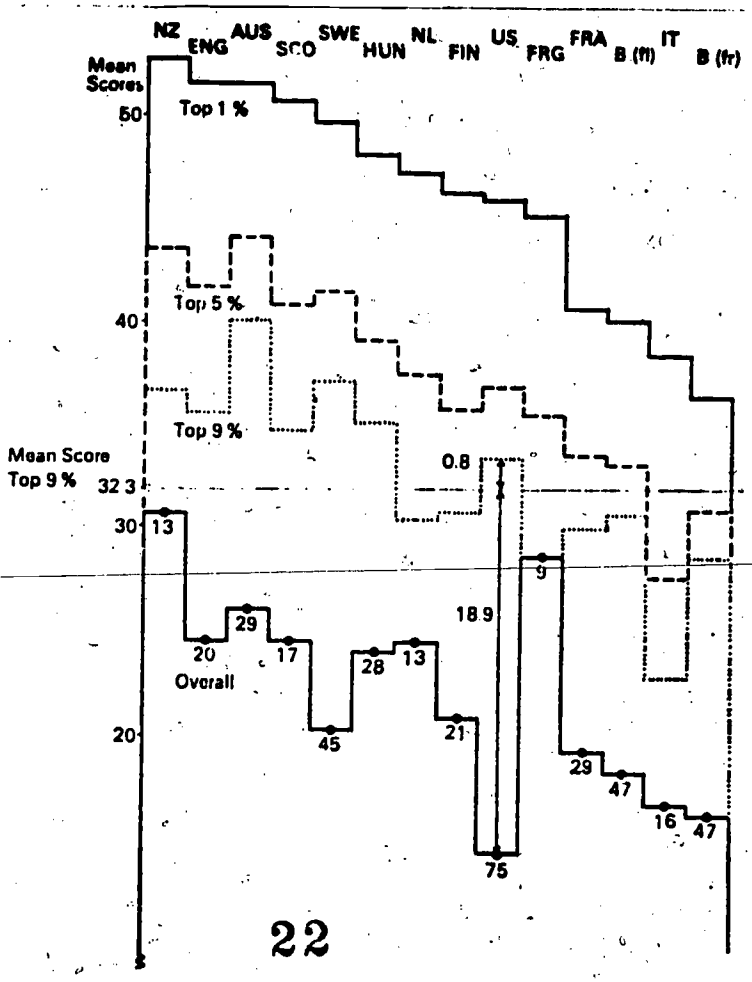


(iii) Does more mean worse?

It is often argued that "more means worse" in the sense that the higher the proportion of an age-group allowed into the final year of schooling, not only will the average achievement be lower but also the achievement of the élite students will be lower.

Taking Science as an example, Figure 4 presents the average mean score of those in full-time schooling in the pre-university grade - the bottom line (the vertical axis gives the score; the figure under the bottom line for each country represents the estimated percentage of an age-group in school; hence Sweden with 45 per cent of an age-group in school has a higher average for that group than France with only 29 per cent of an age-group in school); the average of the top 9 per cent of an age-group (since the Federal Republic of Germany only has 9 per cent of an age-group in the Gymnasium) - the line second from the bottom; the average of the top 5 per cent - the second line from the top; and the top 1 per cent - the top line. The comparisons between countries are fascinating but it is clear that the differences in score shown by the top 1 per cent, or 5 per cent or 9 per cent are not related to the differences in percentages retained in school; "more does not mean worse".

Figure 4. Science mean scores of top 1 per cent, top 5 per cent, top 9 per cent of an age-group and overall group



(iv) Single variables

For all questionnaire items (student, teacher and school questionnaires) means, standard deviations and frequency distributions were produced for every population in every country. Such extensive information is typically not collected by Statistics Divisions of Ministries of Education. For each variable there is a known standard error of sampling. The data collected on the hundreds of variables used in the IEA survey constitutes a mine of information not only for national Ministries of Education but also for international agencies such as Unesco and OECD. Furthermore, a comparison of the national means of a particular variable is of interest. As an example, Table 7 presents the means of the student-teacher ratios for each of the populations tested.

Table 7. Student-teacher ratios by country and by population (Science, Reading Comprehension and Literature

	Student-teacher ratio 1/		
	Population I	Population II	Population IV
Australia	-	19	20
Belgium (Flemish)	22	15	14
Belgium (French)	18	12	14
Chile	40	39	41
England	28	18	17
Fed. Rep. of Germany	33	27	20
Finland	19	20	22
France	-	-	16
Hungary	20	20	17
India	36	27	26
Iran	46	74	72
Israel	20	15	14
Italy	20	16	17
Japan	27	22	-
Netherlands	31	18	17
New Zealand	-	22	22
Scotland	27	17	17
Sweden	17	14	15
Thailand	24	24	24
U.S.A.	26	20	20

1/ Data reported by schools. This ratio comprises the total number of students in a school divided by the total number of full-time teachers (two half-time equal one full-time, etc.) in that school. It does not include teacher aides or technical assistants.

2. Within countries

Within countries the student score in the various subject-matters was used as the school outcome (dependent variable) and the input and process factors as the independent variables in a regression analysis whereby an attempt was made to identify the importance of groups of variables in accounting for variance among students in their performance. Certain groups of variables were formed. The details of how these groups were formed, both through the sieving and compositing of variables, is reported in detail in the various reports and particularly in the technical volume. 1/ For each subject area there were, as has been mentioned above, some 500 independent variables describing the students home background, their previous schooling and current schooling, the teachers and their teaching, the school principal, the organization of the school, its size and budget, etc.

Clustering of variables

The major groups (or blocks) of variables used were:

1. Home background. A weighted composite of father's education, mother's education, father's occupation, number of books in the home, use of dictionary in the home and size of family. (There were minor deviations in this composite between subject areas, and in one country, Sweden, the first two variables had to be omitted because of ambiguity in the Swedish translation of the questions).

Sex, age. (Age was entered in Block 3 for French and English.)

2. Type of school and/or programme in which the student was then enrolled. (The elements of each variable were criterion scaled against Reading, Word Knowledge or Science in order to form the variables type of school and/or type of programme.) This variable was intended as a surrogate (perhaps a weak one) for previous schooling.

3. School and teacher variables (for Science, Reading Comprehension, Literature and Civics).

Time variables, e.g. years studied subject, current grade, grade beginning study of subject and age (for French and English).

4. Kindred variables, e.g. expected occupation, education, attitudes, etc., (for Science, Reading Comprehension, Literature and Civics).

School and teacher variables (for French and English). (In French and English the kindred variables formed a fifth block.)

Typically, two further groups of variables were used but not consistently from subject to subject. The Word Knowledge scores (on an antonym-synonym test) were used and the Reading scores where Reading was not the criterion.

Within countries two major types of analyses were undertaken - a between-student analysis and a between-school analysis. The between-school analysis was undertaken for Science, Reading Comprehension and Literature only. In each case the sieving of independent variables (particularly

1/ Peaker, G.F., An empirical study of education in twenty-one countries : a technical report, 'International Studies in Evaluation, VIII', Almqvist and Wiksell, Stockholm (in press).

the instructional, school organizational and kindred variables) was based on the partial correlation, after home background and/or type of school, type of programme had been partialled out, exceeding twice the standard error of sampling and the importance attached to a particular variable by the subject committee. In this way, the effect of school and teacher variables could be judged after the calibre of input of students to the school had been taken account of. Cases of detected high multicollinearity were 'solved' by compositing the variables in question. Where the school and teacher variables were based on fewer than 50 observations (i.e. 50 schools) there is a problem of stability of estimates. The populations in this category are the two Belgiums (Populations II and IV), Iran (all Populations) and the Netherlands and Hungary (Population IV). It should also be pointed out that, where a variable had 20 per cent or more missing data, it was dropped from the analysis. This was the case for all data relating to the school budget. However, there are countries where these data are complete and further analyses should be run whereby the effect of differential school budget aspects are examined for their effectiveness after at least home background has been partialled out. Some variables (e.g. class size, laboratory class size) were omitted because they were curvilinear. The separate analyses of these were never undertaken because of temporal and financial constraints, and clearly such analyses have high priority.

Table 8 presents as an example the surviving variables by block for New Zealand Science Population II between-school analysis. There are forty single variables which with compositing form thirty variables. This was a typical number of variables for the final regression analyses.

Incremental variance

Appendixes III to VIII present the incremental variance for various blocks of variables, as well as the total variance accounted for in each subject area in each population for the between-student analyses. Table 9 presents a summary of the block variances and total variances account.

The percentage of total variance accounted for varies from a low of 11 per cent in Iran for Population IV Science to a high of 79 per cent in Germany for Population II English. The average percent total variance accounted for was 38.9 per cent which indicates that there is a great deal of unaccounted variance and which implies work for educators to identify and measure other factors not yet included in the IEA analyses. The total variance accounted for is notably lower in developing countries.

The home background variables vary from 1 per cent to 30 per cent, with an average of 11.5 per cent, which raises the question of to what extent are the appropriate home variables being tapped in some countries.

Learning conditions vary from a low of 1 per cent to a high of 52 per cent. The average is 10 per cent across all populations in all subjects in all countries. The percentages are clearly lower for Reading Comprehension, Literature and Civic Education than for Science, French and English.

Table 8. Variables surviving to final between-school regression analyses in New Zealand Science Population II

Block 1	Father's occupation	Home background
	Father's education	Home background
	Mother's education	Home background
	Use of dictionary	Home background
	Books in home	Home background
	Family size	Home background
	Age	
	Sex	
Block 2	Type of school	
	Type of programme	
Block 3	Percent male teachers in school	
	No. of laboratory assistants	
	Sex of teacher	
	'Opportunity to learn'	
	School environment	
	Total homework per week	Homework composite
	Total science homework per week	Homework composite
	Taking science	Study of science composite
	Total years study of science	Study of science composite
	Total hours study of science	Study of science composite
	Principal's teaching experience	
	Total enrolment	
	Decision-making - syllabus	
	Admission criteria - residence	
	- exams	
	Economy of school region	
	Methods - audio-visual	
	- field trips	
	Science expeditions	
	In-service biology course	
	Total in-service training	
	Grade in school	
	Science teacher training	
Block 4	Interest in science	Attitude composite
	Science in work	Attitude composite
	Importance of maths	Attitude composite
	Reading science/technical	
	Reading science fiction	
	Reading science articles	
	Watching TV - science-related programmes	

Table 9. Averages and ranges of all countries of contributions by selected blocks to total variance in the between-student analyses for each subject

	Population I		Population II		Population IV	
	Average	Range	Average	Range	Average	Range
Home background (Block 1)						
Science	11	1 - 27	16	3 - 29	13	3 - 25
Reading Comprehension	14	1 - 25	16	1 - 27	8	1 - 18
Literature	-		15	5 - 25	9	4 - 17
Civics (cog.)	-		16	9 - 20	10	5 - 18
French (reading)	-		15	3 - 15	4	1 - 16
English (reading)	-		14	1 - 26	7	1 - 16
Previous schooling (Block 2)						
Science	1	0 - 14	6	0 - 17	6	0 - 26
Reading Comprehension	1	0 - 9	10	0 - 21	7	0 - 19
Literature	-		8	0 - 14	3	0 - 7
Civics			11	5 - 20	5	0 - 15
French			22	1 - 22	3	0 - 9
English			20	0 - 52	9	0 - 31
Learning conditions (Block 3 for Sc., R.C., Lit. and Civics, Blocks 3 and 4 for Fr. and Engl.)						
Science	8	1 - 21	9	4 - 23	15	4 - 41
Reading Comprehension	6	2 - 18	6	3 - 10	5	2 - 15
Literature	-		7	3 - 12	4	1 - 6
Civics			11	3 - 18	5	1 - 9
French 1/			17	12 - 24	19	10 - 52
English			16	8 - 27	17	4 - 29
Total variance 2/						
Science	27	14 - 36	36	17 - 55	39	11 - 63
Reading Comprehension	28	18 - 39	39	20 - 51	25	13 - 47
Literature			55	36 - 66	37	28 - 53
Civics			59	55 - 62	39	28 - 57
French 1/			45	28 - 65	43	27 - 61
English			60	47 - 79	44	22 - 75

1/ This is with Reading as the criterion. For Listening the percentages accounted for are greater.

2/ The total variance is always unity or in this case unity x 100. Both in this table and in Appendixes III to VIII, it will be noted that the sum of the percentages of the three blocks for any one subject do not total to the percentage in the 'total variance' block. This is because the 'total variance' in each subject is the sum of the first three blocks plus a kindred variables block and typically one or two other blocks which vary from subject to subject but which usually include a Word Knowledge score which might be regarded as a partial surrogate for 'intelligence' and 'previous experiences' not measured in Blocks 1 and 2.

Furthermore, the percentage in these last three subjects rises with the school population being examined. Let us be clear on what these percentages are. In a cross-sectional survey of this kind we are saying that after the home background of children has been taken account of, the differences between schools (as they now exist) in the learning conditions they provide are associated to a considerable degree with differences in performance between students on the criteria under consideration. In some cases, the learning conditions are two-thirds of the total variance accounted for.

But then why should these differences between learning conditions be more important for Science, French and English than for Reading Comprehension, Literature and Civics. Is it that the first three are more school-oriented in that specialized knowledge is being learned where there exists little knowledge in the home and therefore little backing? Or is it that Reading Comprehension is not systematically taught in schools once the decoding and encoding of the mechanics of reading have been accomplished? As a corollary to this, would we expect differences in learning conditions to be greater amd more associated with student performance if it were taught systematically. Is the same true of the sort of civics and literature tested by the IEA tests? Several alternative explanations can be conceived but what do the educators in these subject areas think?

Incremental variance between schools

For the details of the between-school analyses, readers are referred to the main publications 1/, (between-school analyses were undertaken for Science, Reading Comprehension and Literature only) but an example is given here for Scotland and Sweden Population II in Science.

	Scotland	Sweden
1. Home background	80	8
- age	1	-
- sex	0	5
2. Type of programme	2	28
3. School	3	16
4. Kindred	2	6
Total	89 x	63

x Figures making up this table have been rounded.

When the home background of students is aggregated to the school level (neighbourhood effect?) the difference in the percentage of variance accounted for is dramatic between Scotland and Sweden. In Scotland the society is clearly stratified in its housing and district school provision according to home background whereas the difference between neighbourhoods in Sweden is minimal or another way of

1/ International Studies in Evaluation, Vol. I: Science education in nineteen countries, an empirical study by L.C. Comber and John P. Keeves; Vol. II: Literature education in ten countries, an empirical study by Alan C. Purves; Vol. III: Reading comprehension education in fifteen countries by Robert L. Thorndike, Almqvist and Wiksell, Stockholm, 1973.

saying this is that school neighbourhoods are more homogeneous in Scotland. With the home background accounting for so much variance in Scotland, clearly there is little variance left whereby differences between school processes (or learning conditions) can account for further differences. On the contrary in Sweden, with its socially relatively homogeneous society, the differences between schools as they are run accounts for two-thirds of the total variation. It should however be recalled that the two variables, father's/mother's education had to be omitted from the home background composite, but it is unlikely that this would have raised the 8 per cent to more than 10 per cent.

Learning conditions variables

In those cases where differences in learning conditions between schools and students exist, after the calibre of input of students to those conditions has been taken account of, it is important to identify the factors important in the learning conditions. To generalize, the following factors stand out as important:

1. Time, i.e. years of study, number of hours instruction and homework per week prove to be important. Unfortunately, however, the IEA study did not examine the trade-off. If, for example, a student who spends more time on Science performs better than one who spends less time, how does the student who spends less time perform in some other aspect of school behaviour on which he has spent more time than the other student. The study of multiple outcomes for students and schools is clearly important in future research of this kind.

2. The 'opportunity to learn' (or in more general terms the curriculum) is important in that the more a student is allowed to learn the subject, the more he will learn. This is in general a powerful variable and the detailed study of this will be extremely useful to curriculum planners.

3. The amount of teacher tertiary education plus pre-service teacher training is important for student performance, particularly in the higher grades in school.

4. In foreign languages, the extent to which the students and teacher speak in the foreign language is important. This does not imply, however, that total use of the direct method is a panacea for high standards of performance in all aspects of the foreign language.

In French the major variables emerging from the multivariate analyses are always: time, aptitude, interest, and quality of instruction - in that order of importance. It should also be added that there was considerable variation between countries in 'teacher competence' as indicated by such variables as rated competence in French and amount of time spent in a French-speaking country.

5. Measures of classroom climate and the use of practices such as patriotic rituals in the classroom were associated strongly with variation in civic attitudes and to a lesser extent with the civic cognitive score.

For more subject-areas at each population level detailed comments have been made about the learning conditions variables and, as an example, the authors of the Science report comment on each of the following variables at Population IV level for the between-student analysis:

Population IV. Grade, science study and homework, total years study of science, students' planning of investigation, total enrolment of school, science teachers' involvement in science curriculum reform, teachers' post-secondary schooling, teacher training in Biology, teacher subject association membership, teacher's preparation of lessons.

II. FINAL REMARKS

The worthwhileness of the results will depend to a great extent on the way in which policy-makers, at whatever level, understand both the strengths and weaknesses of the results. In many cases, this will depend to a large extent on the systematic links already forged or in the process of being forged between the IEA researchers at the national level and policy-makers in each nation. The univariates descriptive statistics on so many input, process and output variables for various levels in each school system are a mine of information for each national system of education. Some of the between-nation differences, as can be seen from the selected examples given earlier in this paper, have very pointed messages for some national systems. More will surely be provided in the forthcoming Passow-Noah-Eckstein report. 1/

The within-country analyses, despite the problems of correlations between the variables in the final regression analysis and exogeneous variables, have identified aspects of learning conditions (as they now exist) which are highly associated with differences in student achievement and which suggest further experimentation and research.

In a cross-sectional study of this kind, cause and effect cannot be proven, only inferred, and there is a strong case for submitting some of the variables shown to be important to a strict experimentation.

IEA is in the process of establishing a data bank of all these data. Certain parts of the IEA data were never analysed (e.g. the data concerning students specializing in Science at the pre-university level, budget data, etc.) and they must surely be analysed. The within-country analyses were undertaken on total national samples. It is important to undertake regression analyses on specific sub-groups (whether within school type, social class or minority sub-groups, etc.) to discover the deviations from the overall findings for such sub-groups. Although many variables did not reach 'significance' and were therefore dropped from the analyses, certain of these are of importance to different disciplines of education. Their inter-relationships with other variables, possible compositing and differential 'effects' on educational output should be explored. Many of the variables which emerged as important can be coded. Will economists undertake cost-effectiveness analyses using these data? One could proceed with such a list for several pages.

For international organizations one of the problems emerging from the IEA study concerns the types of variables on which it would be useful (and why?) to collect systematically standardized data (and how?).

1/ Passow, A. Harry, Noah, Harold J., and Eckstein, Max, The national case study : an empirical comparative study of twenty-one educational systems, 'International Studies in Evaluation, VII', Almqvist and Wiksell, Stockholm (in press).

To return to the first result presented in this paper, what sorts of pilot work can be undertaken to determine how, at low cost, can one collect reliable data from national samples of children in developing countries where the illiteracy rate is high? What type of interdisciplinary work should be undertaken in developing countries to help identify the types of factors on which data should be collected to help 'explain' differences between students and between schools.

Although evaluators do conceptualize in their roles as researchers, it is clear that in all countries the state of the art of educational theory (or in some cases perhaps the communication of it to the educational researchers and evaluators) is poor, as witnessed by the low total variances accounted for. What types of work can be undertaken to improve this state of affairs?

Finally, it is through evaluation projects of this kind that more 'hard' information can be collected and used to improve the educational provisions for children in the coming years. The knowledge explosion in measurement and evaluation techniques has been rapid. The co-operative nature of IEA's work has helped researchers in some countries to advance their competency by twenty years in a five-year period. More such enterprises of this kind are needed. Research centres and grant-giving agencies should not be deterred by the absence, in many instances, of black and white results. Progress is always slow but decisions concerning the future content and strategies of education must surely be ameliorated by the presence of more research findings.

Appendixes

Appendix I

MAJOR OUTCOME VARIABLES - COGNITIVE AND AFFECTIVE

1. Science

Cognitive

Total Science Score (different tests Populations I, II and IV, but with anchor items).

Sub-scores:
- Earth Science (Population I only)
- Physics
- Chemistry
- Biology
- Practical Science
- Functional Knowledge
- Comprehension
- Application
- Analysis, Synthesis, Evaluation

Specialist Science tests in Biology, Chemistry, Physics (Population IV)

Practical tests (Populations II and IV)

Test of Understanding Science (Populations II and IV)

Attitudinal and Descriptive

Interest in Science

Attitude towards School Science

Attitude towards Science in world

Description of Science teaching (textbook/experimental)

Description of Science teaching (laboratory work - structured/unstructured)

2. Reading Comprehension

Total Reading Comprehension Score (different tests, Populations I, II and IV, but with anchor items).

Sub-scores:
- ability to determine the meaning of a word or phrase in context (Population I only)
- ability to follow the organization of a passage and to identify antecedents and references in it (Populations II and IV only)

- ability to answer questions that are specifically answered in the passage
- ability to draw inferences from a passage about its content
- ability to determine the writer's purpose, intent and point of view.

Word knowledge test (antonym/synonym test - different for each population but with anchor items).

Reading speed test (Populations I and II only - score = item reached)

Sub-score - error rate on first nine items.

3. Literature

Cognitive

Total Literature Score

Sub-scores: Comprehension
Interpretation

Attitudinal

Response preference

Transfer

Interest in Literature

Degree of liking passage

4. English as a foreign language

Cognitive

Reading total (different for Populations II and IV, but with anchor items)

Listening total

Sub-scores: Listening sections I, II, III and IV

Writing total

Sub-scores: Writing Quantity
Writing Quality
Writing Variety

Speaking total

Sub-scores: Structural Control
Oral Reading
Various Fluency Scores

Attitudinal and Descriptive

Interest in English

Perceived Utility of English

English activities out of school

5. French as a foreign language

Cognitive

Reading total (different for different populations but with anchor items)

Listening total (different for different populations but with anchor items)

Writing total

Sub-scores:
- Fill-in writing test
- Writing Quantity
- Writing Quality

Speaking total

Sub-scores:
- Pronunciation
- Structural Control
- Oral Reading
- Various fluency scores

Attitudinal and Descriptive

Interest in French

Perceived Utility of French

French activities out of school

6. Civic Education

Cognitive

Cognitive total (different for different populations but with anchor items)

Sub-scores:
- Citizenship
- Institutions
- Processes
- International
- Economic Processes
- Social Processes
- Simple Behaviours
- Complex Behaviours
- Abstract Behaviours

Attitudinal and Descriptive

Democratic system difference scores:
- Business
- Police
- Welfare
- Unions

Total "don't know" Score

Spread of How Society Works

Political Conflict Perception - How Society Works

Town Council Evaluation

Town Council Responsiveness

National Government Evaluation
National Government Responsiveness
Anti-authoritarianism
Tolerance/Civil Liberties
General Tolerance
Efficacy
Women's Rights
Value of Criticism
Attitude "don't know"
Influence "don't know"
Pressure Group Influence on Laws
Media Influence on Laws
Equality Attitudes
Equality "don't know"
Avoidance of Conflict
Active Good Citizenship
Disengaged Good Citizenship
Non-political Good Citizenship
Personalised Egocentric Attitudes
Personalised Sociocentric Attitudes
Abstract Reasoning
Interest in Civic TV
Discussion with Friends
Discussion with Parents
Discussion with Teachers
Total Discussion
Unawareness of Political Opinions
Agreement with Family
Agreement with Friends
Agreement with Teachers

Appendixes

Appendix II

Student, teacher and school variables in French study

Student	Teacher	School
Age	French in classroom	Class size
Grade	Teach other subjects in French	Principal's degree
Father's occupation	Hours instruction in French	Years as principal
Expected occupation	Sex	Years as principal of this school
Father's language	Age	Years teaching experience
Mother's language	Specialist teacher	Type of community
First language	Full-time education	Museum available
Second language	Post-secondary education	Zoo available
Sex	Teaching experience	Public library available
Father's education	Time teaching at this school	Concert hall available
Mother's education	Hours preparation of lessons	Opera available
Expected education	Hours marking papers	Foreign language societies available
Hours homework	Membership of teachers association	Amenities score
Place of homework	Teaching journals	Total enrolment, boys
Fixed time for homework	Subject journals	Total enrolment, girls
Help with homework	Attended conferences	Total enrolment
Parents correct speech	Assessment, standardized tests	Lowest grade in school
Parents check spelling	Assessment, essay tests	Highest grade in school
Dictionary in home	Assessment, objective tests	Mean grade of school
Parents encourage reading	Assessment, homework	Grade range of school
Parents' interest in school	Assessment, projects and papers	Enrolment of Population I
Encourage culture	Assessment, variety and extent	Enrolment of Population II
Daily newspapers	Importance of student needs	Enrolment of Population IV
Books in home	Importance of curriculum	Co-education, Population I
Hours pleasure reading	Importance of textbooks	Co-education, Population II
TV and radio	Importance of examinations	Co-education, Population IV
Size of family	Importance of next grade	Day or boarding school
Position in family	Methods textbooks	Beginning French grade
Like/dislike of different subjects	Methods drill materials	Beginning English grade
	Methods individual materials	Beginning Social Studies grade
Interest in French	Methods small group work	Decision-making, textbooks
French activities out of school	Methods individual tutoring	Decision-making, rules

Student

Mother studied French
Parents interest
Help with French homework
Utility of French
Aspiration in French skills
Grade beginning French
Entry knowledge
Years studied French
Perception spoken French
Perception listening French
Perception reading French
Perception writing French
Opportunity to speak French
Time in French country
Ease/difficulty French
Perceived industry in French
French books in home
Frequency of exposure to French books

Teacher

Methods audio and visual
Methods field trips
Methods lectures
Methods questioning
Methods discussion
Methods variety and extent of approaches
Within class grouping
Part of time employed
Teacher-training institution
Foreign language teaching load
French in other subjects
Other foreign language
Grade range of French teaching
Number of groups taught
Total time teaching French
Mother-tongue
Age beginning French
Perceived listening skill, speaking skill, reading skill, writing skill, and pronunciation skill
Residence French country
Tertiary French in years
Method training
Years teaching foreign language
Foreign language association member
Use of mother tongue, beginning
Use of mother tongue, intermediate
Use of mother tongue, advanced
Emphasis choices: listening comprehension
Emphasis choices: speaking fluency
Emphasis choices: correct pronunciation

School

Decision-making, choosing teachers
Decision-making, conditions
Decision-making, student selection
Decision-making, expenditure
Decision-making, tuition fees
Decision-making, administration
Decision-making, syllabus
Decision-making, methods
Decision-making, syllabus and methods
Inspection
Report to authorities
Advice on school problems
Advise teachers
Assess teachers
Role of inspection
School operating costs
Teacher salaries
Non-teaching staff salaries
Maintenance and repair
Books and stationery
Purchase of equipment
Other (loan charges)
Total budget
Total budget per pupil
Teacher salaries per pupil
Foreign language percentage budget
Number full-time staff
Pupil/teacher ratio
Teacher salaries per staff
Number of teachers male
Number of teachers in French
Number of teachers in English
Number of teachers in Social Studies
School librarian

Student

Teacher

Emphasis choices: reading comprehension

Emphasis choices: ability to write

Order of spoken and written French

Grammar teaching

Speaking emphasis

Pronunciation methods

Teaching aids, blackboard

Teaching aids, pictures

Teaching aids, "props"

Teaching aids, film strips

Teaching aids, sound movies

Teaching aids, phonograph records

Teaching aids, tape recorder

Teaching aids, language laboratory

Teaching aids, variety and extent

School

French foreign language assistant

English foreign language assistant

Foreign language laboratory technician

Total number of language auxiliaries

Admission criteria, residence

Admission criteria, performance

Admission criteria, interview

Admission criteria, examination

Admission criteria, graduation

Admission criteria, membership

Streaming practices

School programme, variety

Type of course

Language of instruction

Student decision-making

Hours per week schooling, Pop. I, II and IV

Weeks per year schooling

Hours per year schooling, Pop. I, II and IV

When spoken French introduced

When reading and writing French introduced

Class conducted in French

Selection of French students

Percentage study French, Pop. I, II and IV

Number of periods French, beginning, intermediate, advanced

Length of periods French, beginning, intermediate, advanced

Level present, beginning

Total time French, beginning, intermediate, advanced

Level present, advanced

Native teachers French

Appendix III

Contribution by blocks to total variance from between-student regression analyses - Science

Increment / Country	Population I				Population II				Population IV			
	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total
Australia	-	-	-	-	16	7	11	39	13	3	20	44
Belgium (Flemish)	4	2	8	21	8	3	12	26	-	-	-	-
Belgium (French)	12	0	21	36	-	-	-	-	-	-	-	-
Chile	4	0	9	26	13	4	6	25	19	3	8	32
England	21	0	3	32	23	17	7	52	13	2	41	61
Fed. Rep. of Germany	8	2	10	24	18	2	14	34	13	12	8	39
Finland	14	0	4	26	22	6	10	44	25	19	7	56
France	-	-	-	-	-	-	-	-	16	26	5	49
Hungary	8	0	7	20	14	3	5	31	11	13	10	41
India	1	0	20	29	3	10	8	24	4	4	17	26
Iran	6	14	6	32	5	1	9	17	3	0	4	11
Italy	4	0	4	14	10	4	6	24	10	2	16	30
Japan	17	0	1	22	23	0	4	40	-	-	-	-
Netherlands	16	1	7	29	19	15	10	49	21	5	31	63
New Zealand	-	-	-	-	17	12	8	45	13	1	31	54
Scotland	22	1	5	36	29	11	9	55	19	1	34	63
Sweden	16	0	5	24	18	0	7	36	18	8	20	52
Thailand	-	-	-	-	10	3	23	37	-	-	-	-
U.S.A.	18	1	9	34	22	2	7	36	18	9	8	39
Mean	11.4	1.4	7.9	27.0	15.8	5.9	9.2	36.1	12.7	6.4	15.3	38.8
Highest	22	14	21	36	29	17	23	55	25	26	41	63
Lowest	1	0	1	14	3	0	4	17	3	0	4	11

Appendix IV

Contribution by blocks to total variance from between-student regression analyses - Reading Comprehension

Increment / Country	Population I				Population II				Population IV			
	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total
Belgium (Flemish)	1.7	0.0	11.6	22.5	7.5	10.7	9.3	33.7	12.7	18.5	3.4	42.3
Belgium (French)	16.7	1.1	18.4	38.8	11.2	14.5	10.3	43.2	7.2	8.4	9.0	29.7
Chile	1.4	1.2	8.3	22.3	20.1	5.9	6.6	38.6	14.9	4.9	5.1	28.9
England	22.1	0.4	1.5	31.4	27.3	13.7	2.6	50.8	2.4	1.8	4.6	16.4
Finland	17.7	0.0	2.8	27.8	20.3	13.4	4.3	46.8	11.5	2.3	2.3	22.0
Hungary	18.7	0.3	3.8	28.4	18.6	3.9	4.0	36.2	6.9	8.5	4.8	27.1
India	1.6	0.3	14.9	30.8	1.4	3.1	9.7	20.0	4.0	2.2	4.7	12.9
Iran	8.7	9.1	7.2	30.3	6.1	3.2	6.6	20.0	5.5	0.0	7.6	14.5
Israel	25.4	1.9	3.6	36.6	25.2	13.9	4.3	48.3	8.6	19.4	15.2	46.9
Italy	9.6	0.2	4.4	18.0	10.4	12.4	3.2	33.6	9.6	7.8	5.0	25.3
Netherlands	11.1	1.6	4.0	26.4	12.5	20.7	3.9	45.8	1.0	7.4	4.8	16.3
New Zealand	-	-	-	-	13.5	17.6	8.2	46.8	9.0	1.6	1.9	19.2
Scotland	23.7	0.3	2.3	33.8	26.1	12.3	3.9	50.8	4.0	1.4	3.5	20.2
Sweden	11.4	0.3	2.9	18.3	16.1	0.0	2.9	33.7	4.2	14.0	1.8	24.9
U.S.A.	19.8	2.0	3.8	31.8	22.1	3.4	6.5	42.6	17.5	8.0	2.9	33.6
Mean	13.5	1.3	6.4	28.4	15.9	9.9	5.8	39.4	8.0	7.1	5.1	25.3
Highest	25.4	9.1	18.4	38.8	27.3	20.7	10.3	50.8	17.5	19.4	15.2	46.9
Lowest	1.4	0.0	1.5	18.0	1.4	0.0	2.6	20.0	1.0	0.0	1.8	12.9

Appendix V

Contribution by blocks to total variance from between-student regression analyses - Literature

Increment / Country	Population II				Population IV Cognitive				Population IV Interest			
	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total
Belgium (Flemish)	12.4	13.7	6.9	58.0	13.2	6.9	2.3	44.5	11.8	3.9	6.6	38.6
Belgium (French)	18.8	6.3	9.0	55.3	4.1	3.2	3.3	33.6	2.2	3.8	5.1	40.8
Chile	14.4	6.2	8.9	52.7	10.5	3.5	5.7	37.7	1.5	0.9	4.5	30.2
England	25.2	11.9	3.9	66.1	7.3	0.8	0.7	33.7	4.4	8.5	5.7	49.7
Finland	18.1	9.7	5.5	55.8	5.7	0.3	2.7	28.1	2.7	6.3	3.0	34.4
Iran	4.5	2.4	12.1	36.4	8.3	0.0	5.7	28.8	4.4	3.0	7.9	23.3
Italy	10.6	8.7	3.0	43.8	10.7	3.3	5.4	40.0	3.0	4.5	5.0	29.8
New Zealand	15.2	13.9	8.7	63.4	3.7	1.7	2.0	35.1	2.8	7.5	4.2	40.4
Sweden	15.2	0.0	6.1	57.7	5.1	4.4	1.8	37.7	2.1	5.0	4.9	36.2
U.S.A.	18.5	2.4	7.7	61.5	16.5	3.4	5.0	53.0	5.7	1.2	5.9	35.6
Mean	15.3	7.5	7.2	55.1	9.0	2.8	3.5	37.2	4.1	4.5	5.3	35.9
Highest	25.2	13.9	12.1	66.1	16.5	6.9	5.7	53.0	11.8	8.5	7.9	49.7
Lowest	4.5	0.0	3.0	36.4	4.1	0.0	0.7	28.1	1.5	0.9	3.0	23.3

Appendix VI

Contribution by blocks to total variance from between-student regression analyses - Civics Cognitive

Increment / Country	Population II				Population IV			
	Block 1	Block 2	Block 3	Total	Block 1	Block 2	Block 3	Total
Fed. Rep. of Germany	17.6	6.5	17.7	57.2	4.7	0.6	6.8	27.8
Finland	19.0	11.2	9.1	60.6	13.8	2.9	0.9	37.2
Ireland	18.6	4.7	17.0	62.2	9.0	0.8	9.0	43.5
Italy	8.5	10.5	12.2	56.3	-	-	-	-
Netherlands	15.3	20.2	7.5	55.2	8.2	7.9	5.2	31.2
New Zealand	14.4	14.7	12.1	60.0	8.0	0.2	3.3	33.7
Sweden	-	-	-	-	7.5	15.0	6.0	43.2
U.S.A.	20.4	7.0	2.8	61.8	18.3	9.0	4.1	56.7
Mean	16.3	10.7	11.2	59.0	9.9	5.2	5.0	39.0
Highest	20.4	20.2	17.7	62.2	18.3	15.0	9.0	56.7
Lowest	8.5	4.7	2.8	55.2	4.7	0.2	0.9	27.8

Appendix VII

Contribution by blocks to total variance from between-student regression analyses - French Reading

Country \ Increment	Population II					Population IV				
	Block 1	Block 2	Block 3	Block 4	Total	Block 1	Block 2	Block 3	Block 4	Total
Chile	-	-	-	-	-	1.4	3.9	1.9	7.2	27.1
England	12.0	22.1	10.8	3.6	58.5	2.6	4.9	7.1	6.0	44.0
Netherlands	3.7	12.3	14.0	7.4	42.4	0.8	0.4	4.9	5.0	26.8
New Zealand	4.5	1.5	9.2	5.5	43.8	2.0	0.0	2.5	14.3	37.3
Romania	10.6	2.8	0.0	14.8	33.0	15.8	8.6	3.8	14.1	47.4
Scotland	14.6	12.3	13.7	9.8	64.8	2.7	0.4	5.1	16.6	52.8
Sweden	-	-	-	-	-	2.5	0.0	29.0	22.5	61.0
U.S.A.	2.6	0.8	4.1	7.9	27.8	3.3	2.4	7.1	13.5	47.1
Mean	8.0	8.6	8.6	8.2	45.1	4.0	2.6	7.7	11.7	43.0
Highest	14.6	22.1	14.0	14.8	64.8	15.8	8.6	29.0	22.5	61.0
Lowest	2.6	0.8	0.0	3.6	27.8	0.8	0.0	2.5	5.0	26.8

Appendix VIII

Contribution by blocks to total variance from between-student regression analyses - English Reading

Country \ Increment	Population II					Population IV				
	Block 1	Block 2	Block 3	Block 4	Total	Block 1	Block 2	Block 3	Block 4	Total
Belgium (French)	14.2	21.0	6.5	9.8	57.0	6.0	3.4	10.7	15.6	41.9
Chile	-	-	-	-	-	8.2	8.5	4.7	15.2	42.0
Fed. Rep. of Germany	15.1	30.7	14.0	13.0	78.5	1.0	0.5	7.5	8.9	34.0
Finland	25.7	40.4	5.8	2.2	78.2	4.3	0.2	19.3	2.3	43.1
Hungary	-	-	-	-	-	2.9	12.7	2.1	17.5	41.6
Israel	22.4	4.1	5.5	4.4	53.5	9.7	12.1	7.7	5.2	42.3
Italy	1.2	7.5	5.4	13.2	36.1	16.4	30.9	5.8	15.2	75.0
Netherlands	10.6	51.7	8.4	2.1	77.5	0.6	1.2	2.8	5.8	22.4
Sweden	12.5	0.0	8.6	4.2	54.4	3.1	19.5	1.5	2.3	47.0
Thailand	8.2	1.4	14.5	9.9	47.1	12.8	5.3	5.0	17.1	48.3
Mean	13.7	19.6	8.6	7.4	60.3	6.5	9.4	6.7	10.5	43.8
Highest	25.7	51.7	14.5	13.2	78.5	16.4	30.9	19.3	17.5	75.0
Lowest	1.2	0.0	5.4	2.1	36.1	0.6	0.2	1.5	2.3	22.4

OCCASIONAL PAPER No. 30: discusses the major results of a large-scale survey of educational achievement in twenty-two countries by the International Association for the Evaluation of Educational Achievement (IEA). Testing was undertaken in six different subject areas at two or three different levels in each school system. Particular attention is given to between-country differences in achievement and factors accounting for differences between schools and between students within countries.

T. NEVILLE POSTLETHWAITE Ph. D. (Stockholm) was Executive Director of the International Association for the Evaluation of Educational Achievement (IEA) from 1962-1972. He is currently an Associate Staff Member at the International Institute for Educational Planning in Paris, France.