SECTION 2

BRIEF HISTORY OF IEA

IN THE LATE 1950S RESEARCHERS FROM LEADING EDUCATIONAL RESEARCH INSTITUTIONS IN ABOUT A DOZEN COUNTRIES MET UNDER THE AUSPICES OF THE UNESCO INSTITUTE FOR EDUCATION IN HAMBURG. THEY FELT A STRONG NEED FOR AN EMPIRICALLY ORIENTATED, COMPARATIVE RESEARCH PROGRAM IN EDUCATION THAT SHOULD INVESTIGATE PROBLEMS COMMON TO MANY NATIONAL SYSTEMS OF EDUCATION. FOR A LONG TIME MANY IDEAS HAD BEEN ADVANCED ABOUT RELATIVE FAILINGS AND VIRTUES OF THE VARIOUS NATIONAL SYSTEMS. THE GROUP FELT THAT IT SHOULD BE POSSIBLE TO CONDUCT QUALITATIVE EVALUATIONS OF THESE SYSTEMS BY MEANS OF MODERN SURVEY TECHNIQUES. THE WORLD COULD BE CONCEIVED AS A HUGE EDUCATIONAL LABORATORY WHERE DIFFERENT NATIONAL PRACTICES LENT THEMSELVES TO COMPARISON THAT COULD YIELD NEW INSIGHTS INTO DETERMINANTS OF EDUCATIONAL OUTCOMES.

THE WORLD AS AN EDUCATIONAL LABORATORY

by Alan C. Purves, The University at Albany, U.S.A.

One kind of educational research that has grown over the years has been that of comparative and international studies. Although historians and antropologists of education as well as educational planners had been looking at other systems of education for a long time it was not until the 1960s that the idea of comparative studies other than the descriptive came into prominence. One of the major catalysts in that change was the formation of the organization that came to be known as the International Association for the Evaluation of Educational Achievement or IEA. A recent issue of the Comparative Education Review gives something of the history and flavor of the organization (Husén, 1987; Postlethwaite, 1987; Purves, 1987) which I shall summarize briefly. I shall concentrate on the work of IEA in this paper, in part because it represents the sort of research I best know, and in part because I think it is illustrative of the value of comparative educational research as opposed to national or subnational research on the one hand and comparative statistics on the other.

In the late 1950s IEA started as an organization of researchers from around the world who found that they were concerned with a number of issues that could not be studied well within the confines of one schoolsystem. The reason for this is that most systems are more or less uniform with respect to such matters as class size, age of school starting, length of the school year, comprehensive secondary schooling and the like. If one wanted to study the effects of these variables, one needed to design costly and politically risky experiments. There was however "natural variation" if one were to go beyond the borders of a single geographic unit. The idea of comparative empirical studies of achievement and its antecedents and consequents was thus born.

One of the first problems the group faced was that of creating comparable measures, tests that could be used across languages and cultures. The initial experiments showed that this task was feasible, so a full-scale study of mathematics was launched in the early 1960s and was followed by the "six-subject survey" (reading, literature, science, civic education, and English and French as foreign languages) in the 1970s. During the course of the past decade, IEA has conducted a second study in mathematics and science, a study of written composition, and one of classroom environments. IEA is currently conducting studies of pre-primary education and computers in education. It is also launching a study of reading literacy, the first in a series of projected cyclic studies of learning in the basic school subjects, and is contemplating a study of social values and moral reasoning.

In general IEA's methodology has been one of survey research, with an emphasis on careful test construction combined with sets of questionnaires for students, teachers and school heads as well as "national" curriculum questionnaires and supplementary histories and interviews. Over the course of its history, IEA has used various approaches to the analysis of the data and has been among the pioneers of various sorts of casual modelling and analysis. Although the studies are surveys, they are surveys that take into account and, indeed, are predicted upon the differences that exists among the schools and systems surveyed. The studies are what Mike Cole refers to as "comparative" studies, those which take research out of the laboratory and into the real world, or that see the world as a "naturally" existing laboratory. The IEA

studies enable researchers and policy-makers to view alternative strategies and structures in education.

In the last decade of this century the sort of surveys that IEA pioneered have become commonplace within educational systems as national assessments and various international organizations contemplate the collection of data on student achievement, particularly in the basic skills, as well as on other educational indicators. In many cases, they do so with a view of the world and the nation not as a laboratory but as a competition. Such an approach appeals to educational policy-makers, who are able to tie budgets to relative succes or failure in educational horserace. They may ask why it is important to do cross-national comparative research, as an addition to the provision of cross-national indicators for monitoring systems.

I would suggest that there are a number of answers to that question. All of the answers point to the importance of seeing the world as a laboratory containing natural variation. All of the answers would see that the focus of research must be on the educational system as a system involving schools, classes, teachers, students and communities. Comparatively speaking each of these systems has its unique features; at the same time all have common threads. In particular, all systems must cope with a number of issues, some of which have been raised by prior IEA research, and some have yet to be studied. Among the first group of issues I would set the following: exploration of the curriculum and particularly opportunity to learn; exploration of the effects of tracking and streaming; and exploration of the relationship of achievement in a positivist sense to styles and patterns of thinking about the subject. Among the second group I would argue that the following issues could benefit from cross-national study: the educational fortunes of ethnic and linguistic minorities; the tracking of efforts to educate the semi-literate underclass; and the changing educational patterns in third world countries. These are all fundamental educational problems which occur in many countries around the world.

Opportunity to Learn

Over the course of their history, the IEA studies have been remarkably consistent in identifying as one of the main variables that lay behind differences between systems of education in student achievement, "opportunity to learn". This phrase has been used to describe the actual class instruction in a subject, which may differ from what is in the official syllabus. Opportunity to Learn (OTL) has been a remarkably good predictor of the relative achievement of large groups, although not as good a predictor of differences between students. As far as I know, it is a feature unique to IEA measures, and one that has been constantly used since the first mathematics study in the early 1960s. The index is based on an indication by the students tested or their teachers as to whether the material or process or concept measured by a specific item has been presented to the students and how recently it has been presented. The various IEA studies have experimented with different ways of determining OTL. In some cases it has been by a single question; in some by a set of questions that seek to disentangle curricular history; and in some by a series of detailed questions on methods of teaching. There have also been interviews of teachers after the testing as well as interviews and questionnaires given to the students themselves. In two studies the same measures have been given to both student and teacher with a request to estimate the order of importance of various items.

Whatever the method, the result has been to show that at the operational level of schooling there is variation between systems of education as to the opportunity students have had to master a particular concept, learn a particular procedure, or adopt a particular cognitive style, each of which is seen as an important aspect of schooling. The OTL indices have often explained systematic differences in student performance in terms of choice, not chance. If one looks across systems of education, one finds that in mathematics and science , different sub-topics are covered; in reading, literature, and writing different genres and different strategies are stressed; writing, speaking, and listening; in civic education different priorities are set among the various civic and familial values.

Across this variation, the IEA studies have been remarkable in the degree to which they have identified common concepts, common skills, and common values. These might be thought of as the International Basic Skills. The common skills are modified by the tradition and history of a culture and its school system. Even mathematics, seemingly most immune from cultural influence, has been shown in the recently completed IEA study to have cultural overtones that distinguish the learning of algebra in Japan from that in Swaziland or Costa Rica. In reading comprehension, science, and written composition studies, the same picture emerges. Although there may be variation as to the types of texts read or written and the particular linguistic and orthographic systems, such aspects of reading comprehension as analysis and interpretation or of writing as structural coherence and appropriateness of style have universal dimensions. The IEA studies reflect not cultural or educational imperialism but the subtle interplay between national and international definitions of performance in the basic skills.

Opportunity to Learn has appeared in a variety of subjects that IEA has studied, and as a phenomenon has

had the potential of challenging the idea of growth or development as an explanation of why students within a system do not do well on a particular task. Within a system the argument is often that "they are not ready for it", or are not mature enough to do it. When it is shown that in other systems of education students do perfectly well on the same task, this argument loses its strength. This use of Opportunity to Learn has not been made fully enough, and the very concept has not been studied fully as it might. One of the research questions that might arise asks in what other aspects of the curriculum might we see the phenomenon working. It seems that cross-national studies call into question the assumptions about "growth" or "development" that operate within a given culture and show that many of the so-called "universals" of human learning can be reinterpreted as functions of a society's particular curriculum. Such questions need to be sharpened and studied more carefully in future research. They can only be adressed in an international comparative framework.

The effects of tracking and streaming

The previous IEA studies have shown that students who are in the "lower" or less stable tracks or streams of the educational system tend to fall behind. Although this might not seem to be a particularly significant finding, some of the early studies showed that students in non-streamed systems performed better than their counterparts in streamed systems. The reasons for this finding are not clear. They may have something to do with the mix of students and their impact on the teacher, or they may result from the curriculum itself (see Entwistle, Noel, Teaching for learning: retrospect and prospect, ibid., pp. 175-207). More recently, certain analyses of IEA studies have suggested that that the curriculum may be the better answer. In the recent mathematics study and in the Written Composition study, there is evidence in several countries that students in the "lower" tracks are presented with a curriculum that almost garantuees that the students will not achieve at more than a minimum level. These curricula focus on the repetition and drill of very simple material, material that keeps the students attending to the surface features of the subject at the expense of the more important concepts. Students in lower mathematics groups repeat the same sort of arithmetical excersises and so never really move on to the next unit and to more salient mathematical concepts. Similarly, students in mother-tongue keep practicing work in writing grammatical and error-free sentences and so never attend to the discourse-level issues of composition.

These findings parallel similar findings in the earlier IEA studies, particularly in reading and literature, where lower track students practiced decoding skills and never got to comprehension and interpretation. In some school systems, these same students were not given materials to read, but films and other mass media and so were denied access to the culture that they were blamed for not having. The analysis of the science data in several systems of education suggests the same phenomenon occurs in that subject as well; there appear to be major differences in the curriculum depending upon the program in which the students are enrolled.

The question that needs to be examined through the natural laboratory approach is what actual curricular variables operate in those systems that have grouping. Another thrust would be to study those systems in which there is no ability grouping or in which the curricular materials and offerings remain constant across ability groups. What happens when students who are supposedly less able are indeed offered the material that is given to the brighter or more fortunate? There is enough anecdotal evidence across systems of education that the lower group of students can learn if only they are given something of importance to learn. In sampling their school systems, the participating members should seek to indentify and single out the schools which appear to be anomalous. Such has been the planned as an option in the IEA Reading Literacy Study so that one can single out those schools where students perform higher than would be expected given the social class of the students and the type of school in which they are enrolled. That having been done, the research team can visit those schools and explore in depth the possible explanations of their success.

Achievement and styles of conceiving the subject

In some of the earlier IEA studies, attention was paid to aspects of achievement besides a cognitive test score. In the first Science Study there was a Test of the Understanding of Science which tapped something of the ways in which the students conceived of the subject science. In the Literature Study there was a response preference measure, which determined the characteristic approach to texts that students had acquired as an aspect of their instruction in literature. In the recent Written Composition Study there was a measure which asked the students to define achievement in writing in their school. There was also an opportunity to examine differences in approach to the composition assignment among students from the various systems to see if there was such a phenomenon as a national style. The Second Mathematics Study was able to determine characteristic modes of conceiving the subject among the teachers in the participating

systems. The Reading Literacy Study is going to include a measure of reading practices, which can be seen as an index of the uses of reading as learned in the schools and supported by the community.

Different as these studies are, they all suggest that learning in a school subject is not simply the matter of the acquisition of particular skills or pieces of knowledge. Learning a school subject in any system of education bears with it the acquisition of sets of habits and preferences concerning the subject of study. These sets are complex and perhaps culturally determined, but they are more than side-effects of education, they may - in some instances - be the central outcome of learning.

This finding suggests that it is important for researchers to see the extent to which school systems tend to establish communities of learning that are regional, national, or trans-national. Is there a common conception of the subject that enables easy communication across regional or national groups or are there some barriers to communication? To what extent do cultural differences occur in the conception of a particular school subject that may enhance or inhibit learning? At a more particular level the issue may be framed as asking whether certain groups of students are limited in their achievement because they have a conception of the subject that is at variance with that which prevails. In composition, for example, certain groups of students appear to see achievement only in terms of handwriting, neatness, and spelling. These are the students who appear to fail to produce meaningful discourse when they write. Is the problem that they do not have a larger view of the subject or is it that the instruction they have undergone fastens on the surface features rather than such aspects of discourse as content structure, or style? It would appear that the problem exists in subjects like mathematics and science as well, where conceptions about the subject field may both be culturally determined and potentially inhibiting of learning for some groups.

Comparative studies at their best seek to uncover the complexity of causes that may underlie the differences in learning of a particular subject by groups around the world. The IEA studies have come to show that learning in any subject is multivalent. Achievement is not a simple matter of a unidimensional score, but it involves the learning of procedures and approaches with respect to the subject. We may say that these and other more detailed studies have given empirical prove to the activity theory of Vygotsky and his followers. They have shown that school subjects must be seen as complex activities made up of their constituent acts and operations. Further they are culturally embedded particularly in the procedures and approaches, which constitute an important part of what it means being a member of the educated group in a society. Across the range of societies in the world these procedures and approaches are not uniform so that the activity of literacy or mathematics, or science, or even education is not monolithic. Educational systems and the subject communities that are contained in them are culturally situated, and a simple test of achievement runs the risk of neglecting the subtle and powerful dimensions of school learning. It is only through the comparative approach in a natural laboratory that these dimensions can be explored and brought to light.

The educational fortunes of ethnic and linguistic minorities

In most systems of education throughout the world there are ethnic and linguistic minorities. Some of these are indigenous populations as in large nations like the USSR or Indonesia, smaller ones like Finland or Yugoslavia, or new nations that emerged out of artificially carved out colonies such as many of the sub-Saharan African nations. Other minorities are immigrant or "guest" populations as in many of the European nations and others that have taken in large numbers of refugees. Whatever the particular history of these groups, their excistence has long constituted an issue and a concern for educational systems. Among the specific concerns are those of the language of instruction, the segregation or integration of these groups in schools or classrooms, the teaching of the "home" culture and its literature and history, and the importance of adult instruction as well as instruction of the children. IEA's Written Composition Study and Reading Literacy Study in particular have extensive background information on the language practices in the students'homes and schools. Such information can clearly be used to compare and contrast systems and to shed light on these concerns.

The approach to these concerns taken by different systems of education has varied; there is even a variety of approaches taken within a system depending on the size and nature of the minority group. To some extent the approaches taken result from political determination, but there are clear educational consequences of these political decisions. In those nations where a large segment of the population only spends a limited number of years in school the political decision to educate the children in two or three languages may result in a populace that is illiterate in several languages. It might prove more effective to concentrate on one language. Comparative research can shed some light on that issue. It may do so by examining systems that have different approaches to the bilingualism or multilingualism issue but all have a relatively short effective life in school for a large percentage of the population.

Other systems that have immigrant or minority populations can also study the effect of bilingual or monolingual programs and integration and segregation of the language groups by comparative study. These studies can look at performance both in the basic skills of literacy, numeracy, and scientific learning but also the cultural component of education and the extent to which the ways of thinking of the host culture clash with those of the parent culture of the students. The various systems of education around the globe have clearly distinct practices with respect to these issues. Their effectiveness and their effect clearly warrant study (see Siguan, Miguel, Multilingual or multicultural education, what for? Confronting ends and means, ibid.,pp. 121-140).

At the same time one must realize that studies of these concerns may elucidate issues and problems that have other than educational ramifications. The treatment of subgroups is often a political question rather than an educational one, and although the researcher may show the effectiveness of a particular policy or practice, its implementation in another context may prove impossible. Clearly the major political question is the desirability of integration as opposed to segregation. To take the example of the United States and Canada: the former has a long history of assimilation and integration which forms the core of its educational policy; the latter has an equally long history of seperate identities and the protection of ethnic and linguistic identities (see Cummins, Jim, Multilingual/multicultural education: Evaluation of underlying theoretical constructs and consequences for curriculum development, ibid., pp. 141-174). Officially the school systems view linguistic and ethnic groups so differently that there is virtually no common ground. This difference does not preclude the importance of cross-national research involving these two systems. It may be that one policy has a different effect from the other; whether one system can adopt the other's policy may prove impossible. The usefulness of the research is not to produce an educational solution, but to raise the awareness of the decision-makers to the fact that these are political rather than educational issues.

It is also clear that for other systems of education, the political decision has yet to be made and that research into the effects and consequences of decisions in other systems concerning the education of linguistic and ethnic minorities can help the decision-makers to make more thoughtful decisions and also help to educate the administrators and teachers to implement them succesfully. Despite the probability that decisions on this issue are based on a variaty of forces, comparative research using the natural laboratory of the world's educational systems can inform those who are in charge of implementing the decisions that have been made at the political level.

Tracking efforts to educate the semi-literate underclass

In many countries there is emerging a group of people who, while not functionally illiterate, do not succeed in gaining or retaining employement in institutions which require a fair amount of skill (Rosow & Zager, 1988). With the proper training, these people can do the work, but too often the schools and the industries fail to provide them with the training and they take lower-level jobs. These are the people who tend to pull down a nation's mean scores on a measure like those produced by IEA. Some of these people choose not to continue school beyond the compulsory level, either because they see no economic advantage to it, or because they have been deemed or doomed as failures by the educational system.

Analysis of the six-subject survey data in New Zealand and the United States (Purves, 1978, 1980) as well as analysis of the bottom 25% in the more recent mathematics and science studies indicate that the students who are unsuccesful do not necessarily come from the poorest families or the most deprived backgrounds. They are the group that has set relatively low aspirations for itself. The question must be raised as to whether these students as a group have been kept down by "the system", as the tracking issue suggests, or whether they have, in fact, opted out of the conventional set of values. There is an argument by cultural critics that the latter may well be more the case than the former. They refer to this group as being "alliterate", rather than "illiterate", for its members see no advantage to schooling and education in a materialistic culture where these have lost their value. They become the dropouts either in fact or in principle. By law they may stay in school, but they do so as a disaffected subgroup. In many of the Western countries this appears to be the case with an increasingly large population. In some systems of education, there may well be instances where the schools or the employers or some combination of the two is succeeding with this segment of youth. As comparative surveys are conducted in various nations, the national staff should seek to include these exceptional programs in the sampling frame, and perhaps isolate them for analysis to see the extent to which they are really working. Such a proposal means including those programs which often are not included as "schools" in educational systems, because they fall outside the traditional primary or secondary system. Some of these programs are offered by other government agencies such as the military, some of them by industry, some by community action groups. Alternative educational programs must be seen as part of a nation's total educational system and should be included in comparative studies.

Changing educational patterns in the third world

A recent meeting of various groups concerned with education in the third words indicated that a number of international institutions, including The World Bank, UNICEF, UNESCO, and UNDP see some alarming shifts in education including a declining status and income for teachers, increasing drop-out rates, and declining performance of various groups, especially women. The concern is great enough for these groups to call for a new thrust in education and particularly the basic skills and a strong effort to reduce what Benjamin Bloom called "The Achievement Gap" between rich and poor countries and between rich and poor within countries.

This new thrust is one that defines the acquisition of basic skills as a human right equal to health and food and shelter. It will seek to impress upon governments and citizens that having the children attend school is not an end but a means; the end is learning. It will seek particularly to focus on education of those groups which appear to have been denied the right, which is to say women. At the same time such a new policy will force systems of education in the developing nations to make hard choices about where to place their priorities. One of the hardest choices, given a limited economy, is the relative emphasis to be placed on teaching the adult or the child population. Another choice will be on the extent to which education is to be seen as a part of the economy as important as the military or the transportation system.

If education is to be viewed as important in every nation of the world, where should scarce funds be allocated? It is here that international comparative research comes into play. One may say that the resource allocation can be divided along traditional lines: plant and equipment, materials and supplies, and personnel. Each of these is in short supply in many of the developping counties. Some have been ravaged by internal and external strife and have no schools and no way of getting children to school if there a school. Those that have schools have insufficient desks and chairs or suitable means of keeping the noise of other activities in the community or in the school out of the classroom. Of the second group of resources, many systems of education have not the paper for textbooks, much less consumable paper for teacher writing. Some schools may have one textbook for each one hundred students; and only one chalkboard as well. Libraries are non-existent. Along with other studies, IEA research has shown how important materials are to the learning of students, particularly in reading and writing, but also in science and mathematics. Other research has shown the effect of providing hand-held calculators that teach spelling and words as well as those that have arithmetic functions. In an industrialized nation these are easy to provide. They represent a major item in the budget of a poor nation.

But are buildings, desks, and instructional materials at all useful if there are poorly trained or untrained teachers in the classroom? In many systems of education of the world primary teachers have no training save that provided in the same or another primary school. If they are to teach something of mathematics or science, not to mention reading or writing, they must have knowledge of the subject that they are teaching. This need is most important at the secondary and upper primary level, but it is important at the lower primary level as well. The IEA study of classroom environment has shown this to be the case both in industrialized and in developing nations (Anderson et al., 1989). All of the pedagogical technique in the world cannot make up for ignorance of the subject on the part of the teacher.

Another factor which must be considered in considering the issue of basic education for all is that providing such education may serve to undermine the existing cultures of the society and move the society willy-nilly into a standardized society. In a recent article in The Courier (1988), Frederico Mayor wrote, "...culture should be regarded as a direct source of inspiration for development, and in return, development should assign to culture a central role as a social regulator. This imperative applies not only to developing countries, where economic extraversion and cultural alienation have clearly and sometimes dramatically widened the gap between the creative and productive processes. It is also increasingly vital for industrialized countries, where the headlong race for growth in material wealth is detrimental to the spiritual, ethical and aesthetic aspects of life, and creates much disharmony between man and the natural environment" (Mayor, 1988, p.5). The thrust of such a statement or educational research and monitoring is that while it may be important to see education internationally as a race that will lead to economic gain, such a single-minded approach may cause great harm. Education leads people away from their past and their family; it is the main cause for alienation as well as the main cause for aculturation. Programs of international cooperative research should take the occasion of the laboratory to look at the costs and benefits of educations both in terms of the achievement of students and their attitudes and values. In such a way it may be possible to determine which educational programs serve best to educate people for development without destroying their cultural heritage and cultural pride.

Building buildings, providing materials, training teachers, preserving culture in the face of standardization: all are important to the provision of a basic education for all. Which of these come first, if not in order of priority in order of emphasis within a budget? Is there anyway in

helping a nation to decide? How can comparative research contribute to these decisions? One way that has demonstrated its usefulness to planners in a variety of educational systems has been monitoring and research using comparable measures of student performance in various school subjects. One can take the concept of the educational laboratory with systems that have natural variation in their allocation of resources and compare outcomes in terms of student learning and retention. One can also use these sorts of data to undertake simulations and on the basis of these inform policy-makers of the effects of different allocations of resources.

Conclusion

The reasons for having international measures of performance that go beyond monitoring to providing a strong research base are several: to estimate student performance according to standards set by an impartial group; to allow for comparisons of similar systems of education particularly with respect to the education of targeted subgroups; to provide comparisons over time using stable measures of both achievement outcomes and the various background variables that might affect those outcomes; and to allow researchs and evaluators to enter into a dialogue with their collegues around the world.

To an individual system of education, particularly one that has had little experience in doing educational research nationally, the advantages of joining ongoing comparative research projects as opposed to monitoring projects are several:

1. They represent cooperative work in test-construction, created by international teams of experts and based on detailed surveys of the subject domain, so that a participant can learn and share at the same time and not to be the recipient of an imperialistic approach to goal-setting and testing;
2. They are designed to be used across cultures and languages both within and between educational systems so that a system can see where it fits in a larger picture and add national options that suit its own circumstances;
3. They provide an international standard of achievement according to which a particular educational system can ascertain the performance of its students, but balanced by detailed student, teacher, and school and system questionnaires to allow for complex analysis and testing of alternative models and for comparative study of the effects of particular approaches to instruction and schooling.

In sum, the approach of cooperative international research studies allows nations and systems of education to join in the comparative education world and to be part of the laboratory for research. It is a laboratory in which the end is the improvement of education and learning for all. It is a laboratory in which each of the systems of education is an equal researcher and in which all are asking questions that can help the others. It is a laboratory which recognizes the similarities and differences among cultures, nations, and their educational systems and seeks to avoid cultural imperialism in educational solutions. It does assume that there is a global definition of learning and schooling, a broad definition with room for variation according to the particular history and aspirations of each participating system of education. This is a view of education and learning that allows for diversity within a larger global unity.

References

Cole, M. & B. Means (1982) Comparative studies of how people think. Cambridge: Harvard University Press.

Husèn, T. (1987) Policy Impact of IEA research. Comparative Education Review, Vol. 31, pp. 29-48

Mayor, F. (1988) The world decade for cultural development. The Courier, November, pp. 4-6

Postlethwaite, T.N. (1987) Comparative educational achievement research: Can it be improved. Comparative Educational Review, Vol. 31, pp. 150-158

Purves, A.C. (1987) The evolution of IEA: A memoir. Comparative Education Review, Vol. 31, pp. 10-28

Rosow, J.M. & R. Zager (1988) Training; The competitive edge. San Francisco: Work in America Institute, Jossey-Bass.

PAST IEA RESEARCH

FEASIBILITY STUDY

From 1959-1962, IEA ran a smallscale feasibility study in 12 countries (systems): Belgium, England, Finland, Germany, Israel, Poland, Scotland, Sweden, Switzerland, the United States and Yugoslavia. A strategic target population in those countries was the children of age 13:0 to 13:11, since this was the last point where practically all of an age group were still in school in all of those countries. In most cases, children of schools or areas which were known to be close to the national mean and standard deviation were tested, and thus, there was no strict probability sampling.

In total, about 10.000 students representing eight languages were administered tests of reading comprehension, mathematics, science, geography, and nonverbal ability. The venture proved to be succesful. The monograph reporting the results of the study (Foshay, 1962) demonstrated the feasibility aims and procedures of IEA.

FIRST MATHEMATICS STUDY

At meetings held at Eltham Palace, England, and at the UNESCO Institute for Education, Hamburg, in 1959 it was decided to embark on a major cross-national study in mathematics, where several populations at the secondary education level would be sampled using random probability sampling techniques and where specific testing instruments would be specially constructed.

The primary reason for focussing on mathematics was that most countries involved in IEA were concerned with improving their scientific and educational education, at the basis of which lies the learning of mathematics.

The list of participating countries is shown in Table 3: List of Systems in Previous IEA Studies.

In 1964, the instruments were administered to representative samples of 13-year old students, and to pre-university mathematics and non-mathematics students. The data processing took place at the University of Chicago and a two-volume publication appeared in 1967 summarizing the work and results of the study (Husén, 1967). Many of the research institutions participating in the study also undertook special analyses of their national data and produced their own national reports.

SIX-SUBJECT STUDY

Encouraged by the succes of the mathematics study, IEA decided to see if some of the results of the mathematics study were generalizable to other subject areas. In November 1966 it recommenced a study of achievement in science, reading comprehension, literature, French as a foreign language, English as a foreign language, and civic education.

In this study three international populations were identified: Population I included all students in full-time schooling aged 10:0 10 10:11 years, Population II was all students in full-time schooling aged 14:0 to 14:11 years and Population IV encompassed all students in the terminal year of full-time secondary education programs that were either pre-university programs or programs of the same length. The interpretation of this definition of Population IV varied as well as the range of ages of students in the population.

There was also a Population III, which was designed for national datacollection and analysis only. The above population definitions held 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 should have studied it for at least two years.

Table 3 shows the countries which participated in the Six-subjects Study and the subject areas selected.

Although there was some variation from the subject to subject in the kinds of performance outcomes tested, both cognitive and affective outcomes were generally studied. In the field of science, for example, tests were developed to indicate knowledge of various fields (earth science, biology, chemistry, physics); to indicate general understanding of science; to measure practical (laboratory) skills; and to measure ability to use higher-level cognitive skills (application, analysis, and synthesis) in relation to scientific subject matter. On the affective side, there were measures of interest in and attitudes toward science. Information was also obtained that permitted some description of the nature of science teaching.

In the field of reading, the tests were to measure reading comprehension, word knowledge, and speed. In

the related field of literature, scores were obtained on the comprehension and interpretation of literature as well as on preferred responses to it.

Where English or French was thaught as a foreign language, information was obtained on reading, listening, writing, and speaking skills as well as on interest in learning the language, on its utility, and on activities out of school that involved the use of the language.

In the area of civic education, tests were developed to yield a total 'cognitive' score as well as subscores on citizenship, institutions, processes, and ability to use simple, complex, and abstract behaviours. In particular, there were tests to measure attitudinal outcomes as well as perceptions about 'how society works'.

Three to five years of developmental work were required to construct measures of performance in each subject at each level. The resulting tests produced accurate scores (that is, they were reliable), and national subject panels judged the tests to be appropriate for testing what was to be learned in school (that is, the tests were regarded as valid). Formats of the tests included multiple-choice, open-ended, and fill-in items. For foreign languages, taperecorders were used in tests of listening comprehension and speaking. A series of background questions was also given to the students, teachers, and schoolprincipals in questionnaire booklets.

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 at three different levels, which described education as it was at the time of testing and not as it might be. Probability samples of schools and students within schools were drawn for each level for each subject (or group of subjects) within each country. The manner in which the samples were drawn, together with the resulting standard of errors of sampling and design effects for selected variables, was reported in the various subjectmatter and technical publications.

The main international results were published in a series of nine volumes, dealing with science (Comber and Keeves, 1973), literature (Purves, 1973), reading comprehension (Thorndike, 1973), English as a foreign language (Lewis and Massad, 1975), French as a foreign language (Carroll, 1975), civic education (Torney, Oppenheim and Farnen, 1976), technical aspects of the studies (Peaker, 1975) and an overall summary (Walker, 1976).

SECOND INTERNATIONAL MATHEMATICS STUDY

The first major IEA study was concerned with mathematics achievement, interests and attitudes. The data from this survey were collected in 1964 and the main report was published in 1967. In the 1960s and 1970s many countries invested large resources in the development of new mathematics curricula, new instructional methods, new materials and in the training of teachers. It was, therefore, deemed appropriate that IEA should undertake a second study of mathematics.

It was supported mainly by the National Institute of Education (US), the National Science Foundation (US), the National Center for Educational Statistics (US), the Ford Foundation and the Spencer Foundation for all international costs.

Aims and Expected Outcomes

There were three main aims: (a) to describe the mathematics curriculum in each system and also to examine changes in the curriculum since 1960; (b) to measure achievement in mathematics in each system and examine the relative strength of different determinants; (c) to measure growth in achievements over a one year period and assess the reasons for differential growth of students/classrooms in the participating systems.

Curriculum

An analysis of national curricula was used to determine the relative importance attached to various content areas and behavioural levels of mathematics. It also indicated general educational goals of the countries as reflected in the mathematics curricula, as these are embodied in syllabuses, curriculum guides and widely used textbooks. This detailed examination of national curricula provided guidelines for test construction as well as allowing a full description of the 'intended' curriculum. Changes in the 'intended' curriculum over the period since the first IEA study were also examined. The 'intended' curriculum was compared with the 'implemented' curriculum; that is, the curriculum as implemented by the teacher in day-to-day instructional practice in the classroom. The measure of the 'implemented' curriculum was obtained from 'opportunity-to-learn' ratings provided by the classroom teacher.

Classroom

The way in which the curriculum is implemented in the classroom (that is to say, how the subject is taught) is likely to have the greatest effect on what students learn. Surprisingly little is known about what teachers do as they go about teaching mathematics. The study, therefore, sought to determine how much time was made available

to the classroom teacher for mathematics and how the teacher allocated this time to both instructional time (explaining, questioning, leading discussions, and so on) and to keep non-instructional time (taking attendance, keeping records and maintaining order, etc.).
Within the framework of specified mathematical topics, a survey was conducted of the kinds of instructional strategies employed in the classroom and of the materials employed as these strategies are used. Statistical analyses were undertaken in an attempt to identify those instructional variables which are most powerfull in accounting for differences instudent achievement and attitudes.

Student Outputs
In this component of the study, attention was focussed upon the students: how much mathematics they know, at what attitudes they held towards mathematics, towards mathematics instructional practices and towards their own abilities to succeed in mathematics. These student outputs were examined in the light of the curricula (both intended and realized) and of the classroom processes in which the learning has taken place.

For those countries which participated in the first mathematics study in 1964, there was the opportunity to compare student performance on selected topics and objectives.

CLASSROOM ENVIRONMENT STUDY

For 20 years IEA has conducted cross-national surveys of educational outcomes. While the IEA surveys have provided important information about educational achievement and its correlates in home background and school characteristics in different countries, they have had limited potential for affecting teaching and learning practices. To adress more directly the need for improving education, the IEA General Assembly in 1978 approved a descriptive study of classroom processes which have the potential for facilitating student learning. Ten centers participated in the study. The study was supported in part by the Ford Foundation, IDCR (Canada), Japanese Funding Sources and Unilever, Germany.

Aims
The aims of the classroom environment study, were to describe the nature of teaching practices in classroom within the participating systems, to examine the relationship among contextual factors, student behaviours, and teaching practices and, if possible, to identify teaching practices predictive of high student achievement and attitudes.

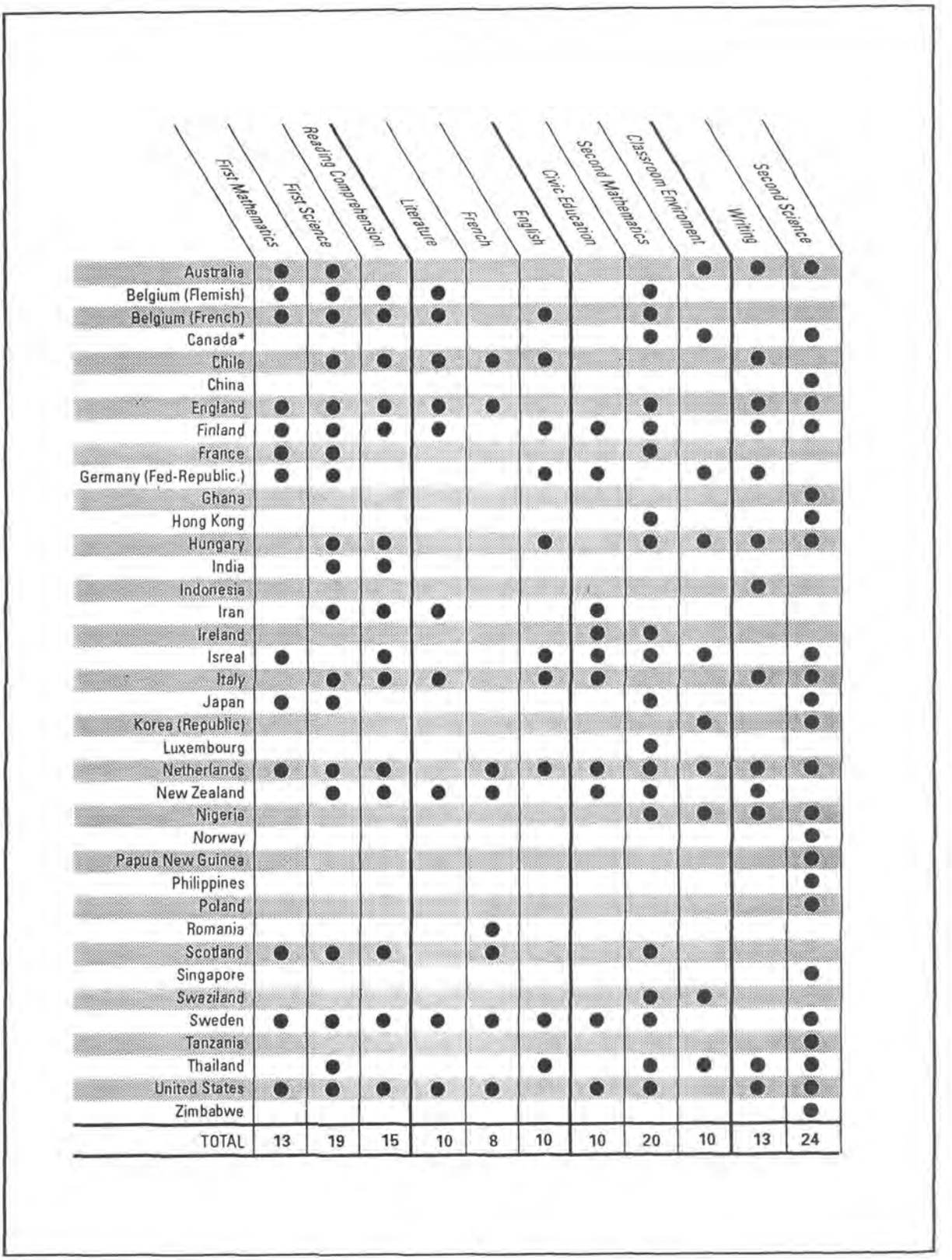
Research design
The conceptual model for the study and the constructs, variables and data sources are presented in An Organizing Framework for the IEA Classroom Environment Study: Teaching for Learning, which was written by D.W. Ryan in '81 (IEA-CE/ONT/TECH/ 100).

The major constructs in the model include context, instructional events and practices, student perceptions of instructional and managment events and practices, student learning processes and student outcomes. Variables subsumed under each construct were selected from literature reviews; some variables are latent and some are manifest. The basic design for correlational study is described in the original proposal written in 1980 by N.L. Gage and Adrian Fordham.

The study included the selection of a classroom sample drawn from the population selected as the focus of the study in each participating country.

Questionnaires and cognitive tests were administered to both teachers and students in the classroom sample at both the beginning and the end of a selected term period. During the school year a total of six to ten classroom observations were conducted. Teachers were interviewed with reference to the observed lesson.

Populations and Samples
The target population for the study (teachers who teach a particular subject-area for a particular grade or age level) was defined by each national centre. Most participating countries selected mathematics as the target subject area and grade 8 as the population level. Judgement samples were drawn.



	First Mathematics	First Science	Reading Comprehension	Literature	French	English	Civic Education	Second Mathematics	Classroom Environment	Writing	Second Science
Australia	●	●							●	●	●
Belgium (Flemish)	●	●	●	●				●			
Belgium (French)	●	●	●	●		●		●			
Canada*								●	●		●
Chile		●	●	●	●	●				●	
China											●
England	●	●	●	●	●			●		●	●
Finland	●	●	●	●		●	●	●		●	●
France	●	●						●			
Germany (Fed-Republic.)	●	●				●	●		●	●	
Ghana											●
Hong Kong								●			●
Hungary		●	●			●		●	●	●	●
India		●	●								
Indonesia										●	
Iran		●	●	●			●				
Ireland							●	●			
Isreal	●		●			●	●	●	●		●
Italy		●	●	●		●	●			●	●
Japan	●	●						●			●
Korea (Republic)									●		●
Luxembourg								●			
Netherlands	●	●	●		●	●	●	●	●	●	●
New Zealand		●	●	●	●		●	●		●	
Nigeria								●	●	●	●
Norway											●
Papua New Guinea											●
Philippines											●
Poland											●
Romania					●						
Scotland	●	●	●		●				●		
Singapore											●
Swaziland									●	●	
Sweden	●	●	●	●	●	●	●	●			●
Tanzania											●
Thailand		●				●		●	●	●	●
United States	●	●	●	●	●		●	●		●	●
Zimbabwe											●
TOTAL	13	19	15	10	8	10	10	20	10	13	24

TABLE 3:
List of Systems in Previous IEA studies

INTERNATIONAL STUDY OF ACHIEVEMENT IN WRITTEN COMPOSITION (WR)

In 1963 when Professor A.W. Foshay of Teachers College, Columbia University was exploring the possiblity of conducting an IEA study in the area of the humanities, the idea of doing a study on both literature and essay writing was seriously considered. Mainly for financial and practical reasons the study had to be limited to one area and literature was chosen, primarily because it seemed more feasible. The literature study was reported in 1973 as part of the Six-subject Study.

In December 1980, initial support for international costs was obtained through a grant from the Spencer Foundation. Data collection took place mainly during 1984 and 1985, with data recording and preliminary analyses beginning in 1985 and 1986. In addition, initial planning is in progress for the permanent storage on microfiche of a large representative sample of student scripts at the Foundation for Educational Research (SVO, the Netherlands) and with their financial support.

Aims and Expected Outcomes

The aims of the study were:

- To contribute to the conceptualization of the domain of writing and particularly the domain of school-based written composition;
- To describe the recent development and the current state of instruction in written composition across the world;
- To identify factors which explain differences and patterns in the performance of written composition and other outcomes, with particular attention to cultural back ground, curriculum, and teaching practices;
- To make a contribution toward solving problems related to the assessment of essay-type answers, particularly when more than one language is involved; and
- To make a contribution towards establishing a more active dialogue between educational sciences and the humanities.

Research Design

The definitions of the populations were:

Population A.
Students at or near the end of primary education and the self-contained classroom.

Population B.
Students at or near the end of comprehensive education; that is students who were in the last year of the shortest secondary program and those in longer programs who had completed the same number of years of schooling whether or not they had finished their program.

Population C.
Students at or near the end of academic secondary school.

Curriculum Analysis

One major aspect of the study was a comparative analysis of mother tongue curricula in the participating countries with special reference to writing and written composition.

This analysis had two components. First, each country had provided a set of ratings indicating to what extent to which certain objectives (content areas and processes) were covered in the curriculum at each population level. These data were collected by means of a Curriculum Questionnaire. Second, each country had written a National Case Study, which dealt with the impact of the following factors on the development and state of the mother tongue curriculum in general and that of writing and written composition in particular: (a) national literary developments, (b) schools of thought with regard to mother tongue education, (c) position of teachers and education, (d) examinations and assessment procedures, (e) minimum competency movements, and (f) writing outside of school.

International Instruments

Data for the study were collected by means of a series of instruments specially prepared for the study (writing tasks, questionnaires, and some attitude measures, all of which were administered by teachers). All writing tasks were direct writing samples, that is, no indirect 'objective' tasks were used in the international component. The selection of tasks was based on a model for the specification of the domain of writing, and on the information provided by RC's concerning curricula, typical topics and writing tasks, and samples of examinations used in participating countries. There are some common writing tasks across three populations. Tasks were rotated in order to cover a wide area of appropriate writing tasks without extending testing time too much.

National background information has been obtained by means of the Interview Schedule , National Context Questionnaire and

National Case Study Questionnaire. Only a few attitude questions were used to measure students'attitudes towards school and writing. The Teacher Questionnaire was given to all the teachers who teach written composition to the students in the samples, to obtain information on teachers'qualifications, experience, teaching and feedback methods, etc. The School Questionnaire was answered by the school principals and provides data on the community and the school.

At its meeting in Finland in August 1980, the IEA General Assembly decided to proceed with the Second International Science Study (SISS), following the First International Science Study (FISS) conducted in 1970. Dr. J.P. Keeves was approached about being involved in such a study and agreed. Dr. Malcolm Rosier was appointed International Coordinator. His first major activity of the study was to prepare a comprehensive planning document: Guidelines to the Second IEA Science Study. This document in turn served to introduce the study to key persons in science education in a range of countries, leading in many cases to an invitation to participate in the study. The data for the study were collected mainly in 1984.

Aims and Expected Outcomes

The aims of the study were:

- to examine the state of science study across the world;
- to identify factors which explain differences in achievement and other outcomes of science education, with particular attention to the role of the science curriculum as an explanatory factor, and;
- to examine changes in the descriptive picture of science education and in the patterns of explanatory relationships since the early 1970s in the ten SISS countries that also participated in the first study.

The expected outcomes of the study were essentially linked to its cross-national nature. From a research viewpoint, it was expected that the cross-national study would produce generalizations about science education that apply across a wide range of countries. In addition, the study would enable policy-makers in individual countries to examine national performance in the context of the cross-national results, as well as providing data to illuminate important national questions, such as:

- What is the relative performance of students in different types of schools in the country, of different ethnic or minority groups, or of different socio-economic groups?
- What is the relative performance of students studying different curricula?
- Which important objectives in the science curriculum are poorly achieved? Which empirical data from the study can be used to improve the curricula?
- What teaching methods or other characteristics of teachers associated with higher student performance in different types of science - knowledge, practical work, understanding of processes, etc.
- What other malleable variables are linked to higher performance; for example, laboratory staff and facilities, the student-teacher ratio, amount of time spent in class on science?
- What are the differences in participation patterns, and performance between male and female students?
- Will the supply of adequatly trained science students be sufficient to meet national manpower needs for scientific and technological workers?

Research Design

Populations and Samples

The target populations for the study were similar to those adopted in the first study. Each country selected the population levels at which it wished to participate, and prepared specific national definitions for:

- *Population 1.*

All students of age 10 at the time of the testing program or all students at the year level (grade level) where most students are of age 10 years at the time of the testing program (the modal year level).

- *Population 2.*

All students of age 14 years at the time of the testing program or al students at the year level (grade level) where most students are of age 14 years at the time of the testing program.

- *Population 3.*

All students studying any science subject at the final year secondary level (the pre-tertiary year level).

- *Populations 3E/3B/3C/3D.*

Sub-populations of Population 3 consisting of all students studying the following science subjects: earth science, biology, chemistry and physics respectively.

Curriculum Analysis

One aspect of the picture of science education developed by the study was the comparative analysis of science curricula across the participating countries. The analysis contained two components. First, each country had provided a set of ratings to indicate the extent of coverage of the curriculum. The classification system used to describe the curriculum had two dimensions: content areas and process areas. Ratings were provided for each population level. Second, each country prepared a National Science Curriculum Case Study, which provided an opportunity for the country to describe the structure of the curricula and the procedures adopted for their development in more detail than could be used in the rating system.

International Instruments

Data for the study were collected by means of a series of specially prepared 'instruments' -tests, attitude scales and questionnaires. All instruments were designed for completion in writing by the respondents with the

exception of parts of the proposed Process Exercises. Instruments to be completed by students were administered in the schools, in most countries, by members of the school staff.

The International Science Tests contain a high proportion of items from the corresponding tests in the First International Science Study. The final set of items in the International Science Tests has been selected to provide due emphasis on the areas of strongest consensus in the science curriculum across the world. There were some common items to be done by all students in the sample at the given population level. For Population 1 and 2 there were additional 'rotated' tests, which permitted countries to add items covering topics of specific national interest. For Population 3 there were also specialist tests in Earth Science, Biology, Chemistry and Physics, to be given to the appropriate sub-populations.

The International Student Questionnaire sought basic information from students, including sex, age, year level (grade level), family background, time spent in class on science, time spent on science homework.

The International Attitude and Other Scales were designed to measure students'perceptions of science teaching, their attitudes to science and society, and their verbal and quantative ability.

The Proces Exercises were developed as an international option to measure the ability of students to handle equipment, design experiments, make observations, draw conclusions, etc.

Countries were encouraged to add national option components to the basic international components, so that the local testing program in any given country would be satisfying to the students and teachers, and would provide data which adress important national problems in science education.

The International Teacher Questionnaire was given to all the persons who teach science to the students in the sample, to obtain information on sex, age, qualifications, experience, etc. and to obtain ratings of the opportunity of each student to learn the content covered by the test items.

The International School Questionnaire was answered by the school principals, and obtained information on school size, instructional time, funding, etc.

INTERNATIONAL REPORTS OF IEA STUDIES

Anderson, L. W. Doris, W. Ryan, and Bernhard, J. Shapiro. (1989)
The IEA classroom Envireronment Study. Pergamon Press, Oxford, UK.

Burstein, L. (ed.) (forthcoming)
The IEA Study of Mathematics Vol III. Pergamon Press, Oxford, UK.

Carroll, J.B.,(1975)
The Teaching of French as a Foreign Language in Eight Countries. Stockholm: Almquist & Wiksell/New York: John Wiley.

Comber, L.C. and Keeves, J.P., (1973)
Science Education In Nineteen Countries: An Empirical Study. Stockholm: Almquist & Wiksell/New York: John Wiley.

Degenhart, E. (1990)
Thirty Years of IEA Rresearch .The IEA Secretariat, The Hague, The Netherlands.

Elly, W.B. (forthcoming)
Achievement in Reading Literacy in 30 Countries. The IEA Secretariat, The Hague, The Netherlands.

Elly, W.B. (forthcoming)
Reading Literacy Vol I. The IEA Secretariat, Pergamon Press, Oxford, UK.

Foshay, A.W. (Ed.), (1962)
Educational Achievement of 13-year olds in Twelve Countries. Hamburg: UNESCO Institute for Education.

Garden, R.A., Robitaille, D.F. (Eds.), (1989)
The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics. Pergamon Press, Oxford, UK.

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

Husèn, T. (Ed.), (1967)
International Study of Achievement in Mathematics: A Comparison of Twelve Countries. Volumes I and II. Stockholm: Almquist & Wiksell/New York: John Wiley.

IEA, (1988)
Science Achievement in Seventeen Countries. ergamon Press, Oxford, UK.

Keeves, J. Roiser, M. (forthcoming)
The IEA Study of Science I. Pergamon Press, Oxford, UK.

Keeves, J. (forthcoming)
The IEA Science Study III. Pergamon Press, Oxford, UK.

Lewis, E.G. and Massad, C.E., (1975.)
The Teaching of English as a Foreign Language in Ten Countries. Stockholm: Almquist & Wiksell/New York: John Wiley.

Lundberg, I. (forthcoming)
The Teaching of Reading Literacy in 30 Countries. The IEA Secretariat, The Hague, The Netherlands.

Olmsted. P. Weikart, D. (1989)
How Nations Serve Young Children: Profiles of Child Care and Education in 14 Countries. High/Scope Educational Research Foundation Press, Ypsilanti, Michigan.

Passow, A.H., Noah, H.J., Eckstein, M.A. and Mallea, J.R., (1976)
The National Case Study: An Empirical Comparative Study of Twenty-One Educational Systems. Stockholm: Almquist & Wiksell/New York: John Wiley.

Peaker, G.F. (1975)
An Empirical Study of Education in Twenty-One Countries: A Technical Report. Stockholm: Almquist & Wiksell/New York: John Wiley,

Pelgrum, H. Plomp, T. (1991)
The Use of Computers in Education Worldwide. Pergamon Press, Oxford, UK.

Pelgrum, H. Plomp, T. (forthcoming)
Computers in Education I. Pergamon Press, Oxford, UK.

Postlethwaite, N., Wiley, D. (eds) (in press)
Science Achievement in 23 Countries. Pergamon Press, Oxford, UK.

Purves, A.C. (1973)
Literature Education in Ten Countries: An Empirical Study. Stockholm: Almquist & Wiksell/New York: John Wiley

Purves, A.C. Degenhart, E. (forthcoming)
IEA Written Composition II. Pergamon Press, Oxford, UK.

Robitaille, D. Garden, R. (1989)
The IEA Study of Mathematics Vol II. Pergamon Press, Oxford, UK.

Ross, K. Postlethwaite, N., (forthcoming)
Effeective Schools in Reading: Implications for Educational Planners. The IEA Secretariat, The Hague, The Netherlands.

Ryan, D.W., Anderson, L.W. (Ed.) (1984)
Rethinking Research on Teaching: Lessons learned from an International Study: Evaluation in Education, Vol. 8.

Thorndike, R.L.,
Reading Comprehension Education in Fifteen Countries: An Empirical Study. Stockholm: Almquist & Wiksell/New York: John Wiley, 1973.

Torney, J.V., Oppenheim, A.N., and Farnen, R.F.,
Civic Education in Ten Countries: An Empirical Study. Stockholm: Almquist & Wiksell/New York: John Wiley, 1976.

Travers, K.J., Westbury, I., (Eds.),
The IEA Study of Mathematics I: International Analysis of Mathematics Curricula. Oxford: Pergamon Press, 1989.

Walker, D.A.,
The IEA Six Subject Survey: An Empirical Study of Education in Twenty-One Countries. Stockholm: Almquist & Wiksell/New York: John Wiley, 1976.