DOCUMENT RESUME

ED 321 757 IR 014 525

AUTHOR Plomp, Tjeerd
TITLE IEA: Its Role and Plans for International Comparative Research in Education.
PUB DATE Apr 90
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Comparative Analysis; *Comparative Education; *Curriculum Evaluation; *Educational Improvement; Educational Research; Foreign Countries; *International Organizations; International Studies; *Policy Formation
IDENTIFIERS *International Assn Evaluation Educ Achievement

ABSTRACT

The International Association for the Evaluation of Educational Achievement (IEA), an independent cooperative of research centers from about 45 countries, performs the type of research that provides data which can be used by policymakers as a basis for decision making about educational improvements. This paper describes: (1) IEA's mission and history; (2) the design and structure of a typical IEA achievement study, using the Second International Mathematics Study as an illustration (including discussions of the conceptual framework on which the study is built, the process by which populations and samples are selected, and the development of achievement tests); (3) some exemplary results of IEA studies which compare total test scores internationally, illustrate the relationship between achievement and other variables, and analyze a national educational system; and (4) IEA's organization, relationship to other national and international organizations, and future plans. Four recent IEA studies are cited and references are provided for forthcoming publications. (24 references) (GL)

The International Association
for the Evaluation of
Educational Achievement

IEA: ITS ROLE AND PLANS FOR INTERNATIONAL COMPARATIVE RESEARCH IN EDUCATION

by

Tjeerd Plomp, chairman IEA

Department of Education
University of Twente
P.O.BOX 217
7500 AE Enschede, the Netherlands

Paper presented at the AERA convention at Boston, MA
April 16-20, 1990.

IEA: ITS ROLE AND PLANS FOR INTERNATIONAL COMPARATIVE RESEARCH IN EDUCATION

by
Tjeerd Plomp, chairman IEA
Department of Education
University of Twente
Enschede, the Netherlands

INTRODUCTION

Why would countries participate in international comparative survey research? Would it not be more beneficial for a country to study educational achievement and educational processes in the context of their own educational system, in stead of spending much time and energy to costly international comparisons? Can fair international comparisons indeed be made? Are not the well known educational statistics on fiscal and other resources, numbers of teachers, student participation, number of graduates, etc. sufficient for policy makers' purposes?

The first question when reflecting on international comparative research is indeed the 'why' question. Of course, nobody doubts the usefulness of national or regional comparisons (state by state, district by district), or the relevance of the 'traditional' educational statistics. Many authors have answered the 'why' question of international comparative achievement research affirmatively (for example, Bathory, 1989; Postlethwaite, 1987; Purves, 1987). Their answers can be summarized in two main points.

The most important reason for a country to participate in international studies of education is to improve their understanding of their own educational system. Since there are no absolute standards of educational achievement, comparative studies are essential to tell policy makers and educators how the level of achievement in their educational system is compared to other systems. In this way these studies can contribute in setting realistic standards for an educational system, as well as in monitoring its quality.

Comparative studies may also be helpful in understanding the reasons for observed differences in performance, by exploring across nations the relations between school achievement and such factors as curricula, amount of time spent on school work, teacher training, classroom size, parental involvement and many other possible explanatory variables.

So, the purpose of international comparative achievement studies is (i) to provide nations with information about the quality of their education in relation to relevant reference groups, and (ii) to assist in understanding the reasons for observed differences between educational systems. These comparisons can be twofold (Postlethwaite, 1987a): (i) straight comparisons of effects of education, that is of total scores or subscores in international tests, which are reflecting the curricula of participating countries; and (ii) comparisons of how well a countries intended curriculum ('what should be taught in a particular grade') is implemented in the schools and achieved by pupils.

So far, the 'why' question is discussed from the perspective of possible interest of educational policy makers. But also many researchers do have their arguments for conducting international comparative research. They are interested in understanding the functioning of education by investigating and explaining differences that exist between educational systems. If these educational systems are schools, then national comparative studies across schools or school districts may be appropriate to find explanations for some differences. However, if we are concerned with national educational systems, then international comparisons are needed to explore the differential effects of independent variables on school achievement. The survey type of comparative studies, either explanatory and/or descriptive in character, takes research out of the laboratory and into the real world; it sees the world as a 'naturally existing laboratory' (Purves, 1989).

In summary, there is a common interest of educational policy makers and of educational researchers in international comparative studies of educational achievement and in possible explanations for observed differences. This common interest makes this kind of studies valuable instruments for monitoring the educational quality, that is for supporting endeavors directed at improving the quality of education.

IEA, the International Association for the Evaluation of Educational Achievement, performs the type of research that provides data for policy makers that can serve as a basis for decision making about improvements of education. The remaining of this paper describes in short IEA's mission and history, the design and structure of a typical IEA achievement study, some exemplary results of these studies, and IEA's future plans, organization, and relation to other, national and international, organizations.

IEA: MISSION AND HISTORY

The International Association for the Evaluation of Educational Achievement (IEA) is an independent international cooperative of research centers from about 45 countries. It has taken as its mission the conduct of comparative studies focusing on educational policies and practices in order to enhance learning within and across sytems of education. IEA has committed itself to a cycle of studies of learning in the basic school subjects and to additional studies of particular interest to its members.
In its studies IEA focuses not only on measuring educational achievement; but it studies also the effects on educational outcomes of groups of variables such as the curriculum, and the organization of schools and classrooms; the relationship between achievement and attitudes; the effects of certain subject matter practices, such as laboratory work in science, and time spent on in-class composition teaching in mother tongue; and different educational practices and attainment of special groups (Purves, 1987).

Founded in 1959, IEA's first international study of educational achievement (mathematics in elementary and secondary education, in 13 countries) was completed in 1966. <u>IEA was the first international organization that used the same objective cognitive tests in more than one country</u>. Other studies were the Six-Subject Survey (in science, reading, literature, civic education, and English and French as foreign languages), with a varying number of countries, from 8 (French) to 19 (science), completed in 1973-1974; the Second International Mathematics Study (20 countries, data collection in 1981); the Classroom Environment Study (10 countries, 1982); the Written Composition Study (13 countries, 1984, 1985); and the Second International Science Study (25 countries, 1983); see appendix for the recent published international volumes on some of these studies. On-going studies are the Pre-Primary project, the Computers in Education study, and Reading Literacy study; while a Third International Mathematics study is under preparation.

The IEA guidebook (Pelgrum, 1989) describes the activities, institutions, and people involved in IEA.
A special issue of Comparative Education Review (Postlethwaite, 1987b) focuses on the first 25 years of IEA.

DESIGN OF A TYPICAL IEA ACHIEVEMENT STUDY

A typical IEA study starts with developing a conceptual framework that clarifies the issues to be addressed, suggests appropriate methods of investigation, results in validated measures of educational outcomes and processes, and uses those analytic tools that can best elucidate key factors and issues (from IEA's mission statement). We will illustrate for the population of

13-years-olds of the Second International Mathematics Study (SIMS) some aspects of the design of a typical achievement study, by discussing the conceptual framework of such a study; the definition of populations and samples; the instruments typically used, especially the instrument "opportunity to learn"; and the curriculum analysis which is underlying instrument development.

Conceptual framework

IEA studies are not restricted to just surveying achievement of students. They are also designed to contribute to the explanation of differences in outcomes, and in this, they contribute to monitoring the quality of education. The latter purpose requires that the variables included in a study must be derived from a conceptual model of how educational systems work. Shavelson, McDonnell, Oakes and Carey (1987) conclude that current research about schooling is not sufficently advanced to support a strictly predictive or causal model, but that it can provide a framework that identifies the logical relationships among those features. They provide a general model of the educational system and the variables (called 'indicators' by them) for measuring each component. The model consists of inputs, processes and outputs (see Figure 1). Based on their research review they suggest how various elements of the

FIGURE 1 ABOUT HERE

system are likely to be logically or empirically related.

An example of such a framework is derived from IEA's Second International Mathematics Study (SIMS)as presented in Travers and Westbury (1989); see Figure 2. It distinguishes three types of the (mathematics) curriculum: the intended, implemented and attained curriculum.

FIGURE 2 ABOUT HERE

At the macro level of the educational system (nation, region, school district), there is the set of intentions for the curriculum; the official goals, the ideas and traditions of the mathematicians and educators. This collections of intended outcomes, together with course outlines, official syllabi, and textbooks, forms the <u>intended curriculum</u>.
On the level of the school and the classroom we have the <u>implemented curriculum</u>, as is the curriculum 'translated into reality by the teacher' (p that is the curriculum as being taught by the teachers. Finally, we have the <u>attained curriculum</u>, the student behavior, that is the knowledge, skills and attitudes acquired by the students.
An important question for monitoring the quality of education is how far the implemented curriculum matches with the intended curriculum: do teachers teach the syllabus; is their perception of the intended curriculum in agreement with the intentions at the system level; do they teach all intended topics, or are they leaving out some topics because of, for example, pressure of time? Similarly important is the question whether the students learn what they are expected to learn according to the intentions; and, if this is not the case, whether the explanation can be found in discrepancies between the different kinds of curriculum.
The different types of curricula are referring to different curricular contexts (see Figure 2). Furthermore, we have the background variables, or antecedents of the curriculum, that influence the curricular contexts, as well as the curricular contents. For example, at the system level, the wealth of a society may affect the retentivity of the school system; community and home background of students may influence the amount of mathematics taught and learned.
The model for the Second Mathematics Study, summarized in Figure 2, provided the starting point for the further design of the study. The rows indicate the level at which relevant data were collected, and hence indicate the nature of variation in the data. For example (Travers & Westbury , p.8):

Level	Source of variation
1. System	Between systems
2. School/classroom	Between schools
	Between classrooms in schools
3. Student	Within classrooms

Populations and samples

Different studies may be directed at different populations. To relate classroom and teacher variables to achievement measures, in a typical IEA study the population definition is given in terms of age and grade. For example, in SIMS the population of lower secondary education was defined as: all students in the grade in which the modal number of students has attained the age of 13.0-13.11 years by the middle of the school year.
Such a population definition allows each participating system to determine a grade level, within which classrooms, and through it teachers and students, can be selected. IEA achievement studies usually have three populations, respectively at the elementary, lower secondary, and upper secondary education level.
Careful sampling procedures are applied, striving at the same confidence intervals across participating countries. Each study has a sampling referee, a specialist who determines the actual sampling procedures, who has to approve the sampling plans submitted by the countries, and has to create acceptable solutions for practical problems. Such procedures give optimal guarantees that the data collected in participating countries are indeed providing reliable estimates of the variables on which data will be collected.

Instruments, Opportunity to Learn (OTL)

In each study a variety of instruments is applied. For example, in SIMS on the level of schools, classrooms, and students: school organization questionnaire; teacher questionnaires on background, attitudes, and teaching practices; students questionnaires on background and attitudes; and achievement tests. The achievement tests consist predominantly of multiple choice items. In some cases IEA experimented with other item formats, such as open ended mathematics items, essays in mother language, performance tests for science, and oral testing of foreign language.
Besides, teacher ratings were collected on whether the content needed to respond to each item of the achievement tests had been taught that year, in prior years, or not at all, to their students. This measure is called Opportunity To Learn (OTL). It provides us with a second measure for student outcomes, a score, adjusted for OTL, next to the raw scores on the achievement tests. Whether to choose in reporting results the raw achievement scores, or the adjusted ones, depends on the purpose. If policy makers are interested in comparing their country with some other countries, they may be interested in the raw test scores. Generally, IEA is reserved in using just raw scores, because it easily makes a study to a kind of 'Olympics' or a 'horse race'; while scores adjusted for OTL may provide more insight in the quality of the educational processes. OTL offers in principle an important tool for carrying out nuanced and relevant comparisons.
This can be illustrated with the following example of the subject French as a foreign language (from Postlethwaite, 1987a).In any foreign-language testing it is usual to test the four skills of reading comprehension, listening comprehension, writing, and speaking. It is also common that these four skills are emphasized differently by teachers according either the intended curriculum in a country, or to how well teachers have mastered each skill. Speaking, for

example, is a skill that has the greatest emphasis in some countries and the lowest in others. Therefore, it would be fair to compare the speaking achievement raw scores in Country A (with high emphasis-OTL) with the raw scores in Country B (with very low emphasis-OTL)? Or should the emphasis-OTL be used to create an achievement score reflecting how well students perform given their opportunity to learn? (p.154).

Development of achievement tests

An important aspect in making valid international comparisons of achievement is the appropriateness or fairness of the achievement tests for the students in all participating systems. This implies that the test development at least needs to reflect the intended curricula in these countries. The procedure for the development of achievement tests, which IEA follows to obtain optimal fairness, will be illustrated with the SIMS (see Travers & Westbury, 1989, chapter 2).

Starting point is that tests reflect as good as possible the intended curricula of the participating educational systems. At the other hand, tests need to have sufficient commonalities of mathematical content that meaningful comparisons between the participating systems is possible. The organizing framework for this was a content-by-behavior grid. For example, the content dimension for population A (13 year olds) consisted of 133 entries under five broad categories: Arithmetic, Algebra, Geometry, Statistics, and Measurement. The behavior dimension was divided into four levels: computation, comprehension, application, and analysis. Each participating system was asked to describe their curricula using the format of this grid. For each cell in the grid, national centers were asked to report on:

1. its universality (i.e., whether the mathematical skills and knowledge defined by the cell was part of the curriculum for "all", "some", or "none" of the students in the system);
2. the emphasis given this aspect of the curriculum;
3. the importance of this aspect of the curriculum within each systems curriculum (p.18).

National centers were also requested to supply items which they considered appropriate for ten preselected cells in each grid.

Based on the completed returns of 14 educational systems, the International Mathematics Committee of the study discussed the differences among the ratings and reached a consensus. An example of such a rating is the grid for Algebra in Table 1.

TABLE 1 ABOUT HERE

Given the differences in ratings between systems, it is clear that the curriculum of a particular systems will be reflected in the international grid more or less well, depending on the extent to which systems have unique emphases on certain aspects of mathematics. As the grid served as the basis for test construction, it was important to get some information of the "goodness of fit" of the grid to each system's curriculum. This information was provided by comparing the international grid with each systems curriculum as described in their curriculum rating. For this, each importance rating was given a numerical value:

V (very important): 3
I (important): 2
Is (important in some systems): 1
Not important or not taught: 0

On the basis of these values, for each educational system the difference between their importance rating and the international rating was calculated for each topic area (i.e., row in fig. 2). Figure 3 presents the results for Algebra. The cross-hatched regions indicate that the

FIGURE 3 ABOUT HERE

international rating for the given cell exceeded the rating provided by the National Mathematics Committee for the indicated system. The black regions indicate that the rating of the National Committee for that topic exceeded the rating in the international grid. The width of the shaded region, ranging from one to three units, reflects the magnitude of the difference in ratings. Blank cells indicate agreement between national and international Importance ratings for the corresponding topics. We see that for Algebra there is some indication that the international grid does not give enough importance to some topics in some systems; countries having notable black bars included Finland, France, Japan, the Netherlands, New Zealand and Canada-Ontario. The result of the test development based on this grid was for the population of 13-years-olds a pool of 199 items.

EXAMPLARY OUTCOMES OF IEA STUDIES

Results of several IEA studies were reported recently (see appendix). Many other publications are written on national and international results of these and previous studies; see for this the IEA bibliography (Degenhart, in press). We will restrict ourselves here to some illustrative examples, predominantly chosen from the SIMS results of the population of 13-years-olds.

Comparisons of total test scores

Until recently IEA has always been hesitant to publish the raw total test scores. The publication of this type of comparison does not do justice to the careful preparation of the studies, and takes not into account differences between curricula of countries and in opportunity to learn between students. Furthermore, such "Olympics" or "horse races" clearly do not serve research purposes. It is for this reason that, for example, Robitaille and Garden (1989), when reporting the outcomes of school mathematics do not give these raw scores. Clearly, IEA is not interested in newspaper headlines such as "Sweden = Nigeria in Math", or "USA Bottom" (Postlethwaite, 1987a).
At the other hand, policy makers are interested in some sort of yardstick, preferably a total score measure. This is clearly illustrated by President Bush of the USA, who in his last State of the Union announced as one of the six national goals for education, that "by the year 2000, U.S. students will be first in the world in science and mathematics achievement." The background document of this section of the State of the Union states explains this goal as follows: "while no international comparisons of student achievement to date are considered adequate, available measures suggest that U.S. 13-years-olds perform near the bottom in science and mathematics compared to their peers in other industrialized countries." (White House, 1990). The document further states the need for a permanent international framework for coordinating international assessments that compare the performance of U.S. students in mathematics and science to that of their counterparts in other industrialized countries. In this context reference is made to the IEA Third International Mathematics Study, which is under preparation, and will have data collection in 1993/1994.
So, if IEA does not publish the total scores on the achievement tests, others will do this as soon as the data tapes of a study have been released. For example, the total cognitive score of the 13-years-olds in SIMS are published by Suter (1989), in conjunction with Opportunity to Learn (see Table 2). It is this kind of table that is referred to by the President of the USA.

TABLE 2 ABOUT HERE

The OTL scores in Table 2, which are teacher judgements, show great variation between countries, indicating that the overlap between the test and the implemented curriculum in the classrooms (see Figure 2) of the respective systems varies a lot across systems. In other words, the OTL data suggests that the SIMS test does not perfectly reflect the implemented mathematics curriculum in every participating country, unless the careful preparation of the tests based on the intended curriculum of the participating systems. These low OTL teacher

scores suggest also that within many countries a substantial difference exists between the intended curriculum ('what is expected to be taught') and the implemented curriculum ('what actually is be taught').
The differences in curriculum coverage (OTL) are strongly related to the student achievement score; Suter (1989) found a (Pearson product moment) correlation of .52. If we take the cognitive test as a standard to which student achievement might be measured (which is an acceptable assumption, given the careful procedure of 'curriculum driven' test development), a possible interpretation of this strong associations, that in some countries, such as the USA, "students might not learn as much as in other countries because they are not presented with the information in class" (Suter, p.14). This was, in fact, the conclusion of the international mathematics committee (McKnight et al., 1987), which explains the relatively low performance of U.S. students on mathematics achievement tests in SIMS to a curriculum that moves very slowly, containing huge amounts of redundancy, and introduces, compared with other systems, fewer new topics.

Another example of how national achievement data can be interpreted in an international context is presented by Marklund (1989). When the national SIMS data were published in 1983, little attention was paid to them. The between-school variance was very low, which could be interpreted as showing that the standard of mathematics education was not much different across Swedisch schools. But when the international averages were published in 1985, showing poor results of Sweden (see Table 2), the SIMS became highly relevant in the public discussions about, for example, the reform of teacher education. A governmental committee on mathematics education proposed, amongst others, a study day in all schools compulsory for all mathematics teachers in grades 7-9; earmarked resources during four years for inservice teacher training, and for development work and curriculum changes. The low between-school variance was now politically and practically helpful, in that the measures proposed by the mathematics committee resulted in a national development program for all schools (Marklund, p.41).
Based on the Swedish analysis of the SIMS and also of the SISS results, Marklund (1989) concludes, that it is imperative that national and international results are published and discussed simultaneously. If not, the international results will remain an Olympic game with very limited possiblities of being explained by and added to the understanding of national results (Marklund, p.42). Given the earlier discussion of some USA results, we can only but fully agree with this conclusion: total test scores, in conjunction with measures of curriculum coverage like OTL, provide us with a rich source of possibilities to learn about our national curricula.

Relation between achievement and other variables

In the preceding section we have seen that there is a relation between student achievement and teacher ratings of <u>OTL</u> for mathematics. Similar results are found in other IEA studies; for example, by Comber and Keeves (1973) in the first science study. Pelgrum (1990) recently showed that the validity of these teacher ratings are in general promising, which makes this measure a powerful instrument for analysing student achievement in the context of the implemented curriculum ('what is actually taught in the classrooms'). Kifer (1989) concludes from analyses of the different findings on the relationship of achievement with other variables, that <u>what</u> is taught (thus: OTL) is related to achievement differences, but <u>how</u> students are taught is not (p.57).
<u>Amount of time</u> is another variable which is related to achievement. Carroll (1975) shows from the IEA study of French as a second language, that the longer students study a subject, the more proficient they become in it. This amount of time could be reduced if the student would have a teacher fluent in French, being taught mainly in French, and have serious aspirations to learn French.
Kifer (1989) points to different implications of these two findings. "Where Carroll's findings include implication for the allocation of time, teacher competence, and teaching methodology;

the OTL findings suggest that the relative importance of the variables places exposure to content very high on a list, and teaching variables and 'sequencing' of the curriculum rather low. ... This means that exposure to the material is the necessary condition for learning; while sequencing, pacing, and teaching variables follow in importance" (p.58/9).

Schmidt and Kifer (1989) also concluded that in SIMS for the population of 13-years-olds class size has no effect on achievement (which, of course, does not mean that there are no other arguments for reducing class size in many systems)

Looking at gender, the results of SIMS (13-years-olds) tend to confirm those of earlier research, that boys tend to perform better on mathematics than girls (Robitaille, 1989; Schmidt & Kifer, 1989). A more interesting outcome is that there are systems in which girls are outperforming boys on all mathematics subtests (Belgium-Flemish, Belgium-French, Finland, Sweden, and Thailand); while in other countries the reverse is found (for example, in Canada-Ontario, France, Hong Kong, The Netherlands, New Zealand, Scotland, and some others). The interesting lesson from such a finding is, that a pattern what may be considered as 'given' or not changeable, may seem a less unchangeable situation than many of us presuppose. A closer analysis of countries where girls outperform boys may lead to interesting suggestions of how to change mathematics practice in countries with the reverse situation.

A national system analyzed: the case of Hong Kong

The examplary findings of IEA studies, presented so far, are in fact all examples of results presented in the context of the input-process-output model (Figure 1). Main effects are discussed (based on raw scores, and on scores adjusted for OTL), as well as relations of achievement with a few process and input variables.
IEA studies may give reason for other types of analysis within an educational system, in which one tries to understand and interpret complex relations between variables within the own system with the help, and in the context of international comparisons. An example provide Brimer and Griffin (1985) in their analysis of the SIMS data of Hong Kong (as reported by Garden, 1987). They had particular interest to investigate why Hong Kong students in SIMS (population of 13-years-olds) achieved relatively high in the international comparison, while they were a year younger, on average, and given that measures on most important home background and school-related variables for Hong Kong pointed to the likelihood of low average achievement. Brimer found that various factors pointed to a systematic selective effect, having to do with the process of grade repetition and promoting by which Hong Kong students reach seondary school. At each grade level in the primary school, about 6 percent of students are retained to allow them to improve their low achievement. This results in a wide age dispersion by the time they reach seondary school, and a strong negative correlation between age and achievement. More important, at the transition stage from primary to secondary school, internal school assessments that are scaled on the basis of aptitude measures taken toward the end of primary schooling are used in determining which students shall have the first choice of secondary schools. Brimer's evidence suggested that the mathematics component of the aptitude tests played a dominant role in selection for secondary school and thus "that a bias exists in the evaluation of the mathematics component of the curriculum, even though there is no evident bias in the time of school resources devoted to it. ... The combined evidence lends credibility to the view that international comparisons are capable of highlighting systematic disorders in an education system that might otherwise have gone undetected" (from Garden, 1987; p.67).

IEA'S CYCLE OF STUDIES

From the examplary results of IEA studies we conclude that these studies can play an important role in monitoring the quality of education. The model of the educational system in Figure 1 illustrates the kind of data which are needed for monitoring activities.
Countries can often easily generate certain national educational statistics or indicators, such as numbers of students and teachers at each level of education, percentage of repeaters, numbers of graduates of the different school types, public expenditure on education, etc. These kind of data are internationally published by, for example, Unesco.
But Unesco and OECD are not collecting or publishing data on student achievement, on the curriculum and curricular change, on classroom processes, and on student background. IEA type of studies are needed for generating these kind of indicators and to allow nations and international organizations to analyze relationships between these kind of variables, and of these variables and the more 'easy to collect' educational statistics.
To meet this need, IEA has decided to plan a series of phased repeat surveys in the participating systems, with constant updating of the items and tests in the light of curricular changes, but at the same time with the possibility to study developments over time (Purves, 1987). These surveys are focused on basic school subjects in elementary and secondary education : reading (data collection in the academic year 1990-91), mathematics (1993-94), science (1996-97), and foreign languages and cultural studies (1999-2000). Next to the cycle, the possibility will remain for additional studies, like the on-going Pre-Primary study, and the study on Computers in Education.
The surveys in basic school subjects will have an increased emphasis on opportunity to learn and on classroom processes; which implies that IEA will continue to take as the important starting point for their studies that the intended and the implemented curriculum of the participating systems has to provide the contexts for interpreting student achievement (i.e. the attained curriculum). Next to this, increased attention will be paid to ways of combining quantitative and 'narrative' methodologies in order to provide potential for rich interpretations of the statistical data, and in this way providing decision makers which more comprehensive information. This 'enrichment' will already be applied in the on-going studies of Reading Literacy and Computers in Education. In the latter, as an international option, per country a case study will be conducted in two schools from the survey sample which are known as belonging to the best examples of computer using schools in their country, to get an understanding of the factors on school level which influenced positively the integration of computers in the educational processes.

Such a cycle of studies has several advantages. Members can plan ahead for participation, and can budget the national studies well in advance. They also can relate national assessment projects to IEA studies, for example by adding to the 'national' IEA studies items from national assessments.
A more substantive advantage is that in a cycle, although the subject matter focus may differ, certain variables or indicators can be studied longitudinally across the studies, such as school variables, student and teacher background variables, and classroom processes. In this way the IEA cycle enriches the possiblities of studying over time the effects of interventions in educational systems.

RELATIONS WITH OTHER INTERNATIONAL ORGANIZATIONS

Nowadays, we observe not only an increasing interest in the quality of education on national levels (see, for example, our discussion of Sweden and the USA), but this interest is also manifest and increasing on international level. Examples are Unesco, with its regular publication of educational statistics of member states, and the OECD International Indicator Project, which attracts a lot of interest and involvement of all member countries. Furthermore, in Europe the process of unification within the European Community is accelerating, which as well calls for increased interest in indicators of educational quality in the member countries. IEA has decided to take into account this interest of international organizations in the kind of indicators generated by its studies. Already in the Reading Literacy study, in a more informal way, some variables proposed by the OECD indicators study will be included in the instruments. From now on, international organizations will be invited to participate in the preparation stage of each new study; for example in invitational workshops and meetings of national research coordinators where the design of a study will be elaborated, and the variables to be included are determined. Similarly, IEA is prepared to have these organizations in some way or another involved in the discussion of the results of the studies. To be concrete, IEA will already invite representatives of OECD, Unesco, EC, and World Bank (and possibly other relevant institutions) for meetings during the fall of 1990, where the design of the Third International Mathematics Study (data collection 1993-94), and stage II of the Computer in Education study (data collection 1992) will discussed.

FUTURE PLANS AND ORGANIZATION

The cycle of studies, and the involvement international organizations as potential 'consumers' of IEA results in certain stages of the studies, are both the result of a process of reflection within IEA on how to function meaningfully in the future, while maintaining its mission. Another result of IEA's reflection on its future will be the re-establishing of the training activities. In the 1990 General Assembly plans will be discussed for a set of training seminars for beginning and advanced survey researchers, to be conducted in English, French, and Spanish. These plans are inspired not only by the need of IEA of having well qualified researchers as national research coordinators in its studies, but more generally also by the growing interest in monitoring the quality of their education, both in more developed countries as well as in less developed countries.
The announcement of these plans and the typical character of IEA studies, being collaborative efforts in the design of the study, the analysis of the curricula, and the development of instruments, sets on IEA the task of carefully monitoring its studies and other activities. For this purpose, a small executive office is established in the Hague, which is operational from March 1990. The address is:

IEA
c/o SVO
Sweelinckplein 14
2517 GK THE HAGUE
the Netherlands
tel: *3170-3469679 (also fax)
e-mail (EARN/BITNET): SURF445 @ KUB.NL (from 1 September 1990: IEAHQ @ KUB.NL).

Research volumes on recent studies:

Anderson, L.W., Ryan, D.W. and Shapiro, B.J. (Eds) (1989). The IEA classroom environment study. Oxford (UK): Pergamon Press.

Gorman, T.P., Purves, A.C. and Degenhart, R.E. (Eds) (1988). The IEA study of written composition I: the international writing tasks and scoring scales. Oxford (UK): Pergamon Press.

Robitaille, D.F. and Garden, R.A. (Eds) (1989). The IEA study of mathematics II: contexts and outcomes of school mathematics. Oxford (UK): Pergamon Press.

Travers, K.J. and Westbury, I. (Eds) (1989). The IEA study of mathematics I: analysis of mathematics curricula. Oxford (UK): Pergamon Press.

Other volumes on the Second International Mathematics Study (vol. III), the Written Composition Study (vol. II and III), and the Second International Science Study will published later on in 1990 or in 1991.

REFERENCES

Anderson, L.W., Ryan, D.W. and Shapiro, B.J. (Eds) (1989). The IEA classroom environment study. Oxford (UK): Pergamon Press.

Bathory, Z. (1989). The Hungarian experience. In Purves, A.C. (Ed), International comparisons and educational reform (pp. 45-50). Association for Supervision and Curriculum Development.

Brimer, A. and Griffin, P. (1985). Mathematics achievement in Hong Kong secondary schools. Hong Kong: University of Hong Kong, Centre of Asian Studies (presented to the 26th IEA General Assembly).

Carroll, J.B. (1975). The teaching of French as a foreign language in eight countries. Stockholm: Almqvist and Wiksell; New York: John Wiley.

Comber, L.C. and Keeves, J.P. (1973).Science education in nineteen countries. Stockholm: Almqvist and Wiksell; New York: John Wiley.

Degenhart, R.E. (in press). IEA bibliography. The Hague: IEA Headquarters.

Garden, R.A. (1987). The second IEA mathematics study. Comparative Education Review, 31 (1), 47-68.

Gorman, T.P., Purves, A.C. and Degenhart, R.E. (Eds) (1988). The IEA study of written composition I: the international writing tasks and scoring scales. Oxford (UK): Pergamon Press.

Kifer, E. (1989). What IEA studies say about curriculum and school organization. In Purves, A.C. (Ed), International comparisons and educational reform (pp. 51-72). Association for Supervision and Curriculum Development.

Marklund, I. (1989). The Swedisch experience. In Purves, A.C. (Ed), International comparisons and educational reform (pp. 35-44). Association for Supervision and Curriculum Development.

McKnight, C., Crosswhite, F.J., Dossey, J.A., Kifer, E., Swafford, J.O., Travers, K.J. and Cooney, T.J. (1987). The underachieving curriculum. Champaign (Ill., USA): Stipes.

Pelgrum, W.J. (Ed) (1989). IEA guidebook 1989: activities, institutions and people. The Hague: IEA.

Pelgrum, W.J. (1990). Educational assessment: monitoring, evaluation and the curriculum. De Lier (the Netherlands): Academisch Boeken Centrum.

Postlethwaite, T.N.P. (1987a). Comparative educational achievement research: can it be improved? Comparative Education Review, 31 (1), 150-158.

Postlethwaite, T.N.P. (Febr. 1987b). Comparative Education Review, special issue, 31 (1).

Purves, A. (1987). IEA: an agenda for the future. International Review of Education, 33, 103-107.

Purves, A. (1989). The world as an educational laboratory. A paper submitted to Fundamental studies in educational research. The Hague: to be published by SVO.

Robitaille, D.F. (1989). Students' achievement: population A. In Robitaille, D.F. and Garden, R.A. (Eds) (1989). The IEA study of mathematics II: contexts and outcomes of school mathematics (102-125). Oxford (UK): Pergamon Press.

Robitaille, D.F. and Garden, R.A. (Eds) (1989). The IEA study of mathematics II: contexts and outcomes of school mathematics. Oxford (UK): Pergamon Press.

Schmidt, W.H. and Kifer, E. (1989). Exploring relationships across population A systems: a search for patterns. In Robitaille, D.F. and Garden, R.A. (Eds) (1989). The IEA study of mathematics II: contexts and outcomes of school mathematics (209-231). Oxford (UK): Pergamon Press.

Shavelson, R., McDonnell, L., Oakes, J. and Carey, N. (1987). Indicator systems for monitoring mathematics and science education. Sant Monica (CA, USA): RAND.

Suter, L.E. (1989). An examination of country differences in mathematics achievement: a synthesis of research results from the IEA SIMS. Washington, D.C.: National Center for Education Statistics (Revised from a paper presented at the annual meeting of the American Sociological Association, 1987).

Travers, K.J. and Westbury, I. (Eds) (1989). The IEA study of mathematics I: analysis of mathematics curricula. Oxford (UK): Pergamon Press.

White House, The (1990). National educational goals. Washington, D.C.: The White House, Office of the Press Secretary.

Figure 1: A model of the educational system (from Shavelson et al., 19^7).

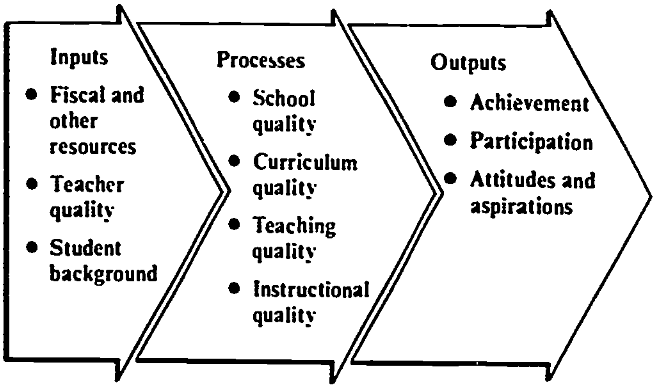

Figure 2: A model for the Second International Mathematics Study, SIMS (from Travers & Westbury)

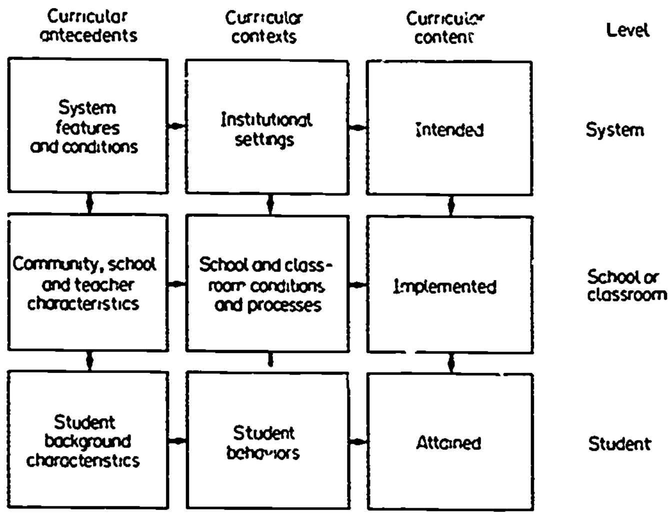

Table 1: Importance for instrument construction of content topics and behavioral categories for Algebra (population 13-years-olds) (from Travers & Westbury)

Content topics	Behavioral Categories* Computation	Comprehension	Application	Analysis
Algebra				
101 Integers	V	V	I	I
102 Rationals	I	I	I	I
103 Integer exponents	Is	–	–	–
104 Formulas and algebraic expressions	I	I	I	I
105 Polynomials and rational expressions	I	Is	–	–
106 Equations and inequations (linear only)	V	I	I	Is
107 Relations and functions	I	I	I	–
108 Systems of linear equations	–	–	–	–
109 Finite systems	–	–	–	–
110 Finite sets	I	I	I	–
111 Flowcharts and programming	–	–	–	–
112 Real numbers	–	–	–	–

*Rating scale: V = very important; I = important; Is = important for some systems. A dash (–) = not important

Figure 3: The match between the international grid and the national grids for Algebra (population 13-years-olds) (from Travers & Westbury, 1989)

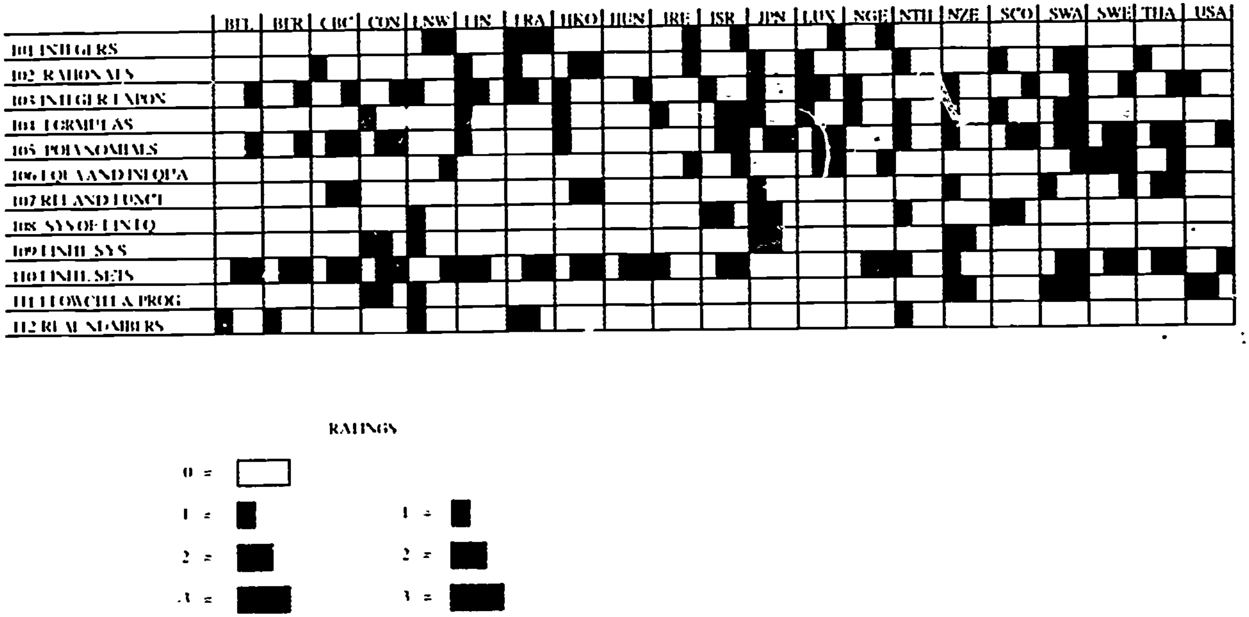

Table 2: Student Cognitive Achievement and Teacher Opportunity to Learn (OTL) for Eighth Grade Level Mathematics Students in Fifteen Countries: Rank Order by Total Cognitive Score (from Suter, 1989)

(Average percent of items answered correctly)

Country	All 157 items			30 Algebra items		
	Cognitive	OTL	Learning efficiency (2/3)	Cognitive	OTL	Learning (5/6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Japan	62.1	76.7	81	60.3	83.4	72
Netherlands	57.1	70.5	81	51.3	73.1	70
Hungary	56.0	90.1	62	50.4	90.9	55
Belgium (Flem)	53.2	60.6	88	52.9	70.9	75
France	52.5	72.3	73	55.0	86.1	64
Canada (B.C.)	51.6	72.0	72	47.9	84.6	57
Belgium (Fr)	51.3	-	-	49.1	-	-
Canada (Ont.)	49.0	71.7	68	42.0	71.0	59
Scotland	48.4	-	-	42.9	-	-
England @ Wales	47.3	68.4	69	40.1	63.1	64
Finland	46.8	61.0	77	43.6	68.5	64
New Zealand	45.5	64.1	71	39.4	62.7	63
United States	45.3	68.6	66	42.1	68.9	61
Thailand	42.2	75.2	56	37.7	83.0	45
Sweden	41.8	53.8	78	32.3	49.8	65
15 Country ave.	50.0	69.6	72	45.8	73.5	62

Note: OTL is defined in this table as Opportunity to Learn, as reported by the teacher