

# THE CONTRIBUTION OF IEA TO RESEARCH AND EDUCATION

*Richard M. Wolf*  
Teachers College, Columbia University, USA

I want to thank the organizers of this conference for their kind invitation to me to present what I consider the important contributions that IEA has made to both research and education over the years. I have been a part of IEA and IEA has been a part of my and my family's life for over forty years now. It has been an exciting ride, to say the least. Despite my continuous connection with IEA over this period, trying to identify its contributions is no easy task. The reason for this difficulty stems from the fact that isolating the contribution of a specific activity or organization to some larger social structure such as education is risky. One is not sure how much to ascribe to the specific activity and how much to other factors in a society. What I am offering here are my best guesses as to what IEA has contributed to research and education. It is not inconsiderable, I believe, and I shall try to demonstrate this in this paper.

I would like to begin by referring to the 1950s. There was a great deal of ferment in education at that time. As one example, two books were published in the United States that became best sellers with the general public. The first was published in 1955 by Rudolph Flesch and was titled, "Why Johnny Can't Read" and the second was published in 1958 by Admiral Hyman Rickover and was titled, "Swiss Schools are Better Than Ours". Besides being best sellers, these two books had two other things in common. First, each was highly critical of primary and secondary schooling in the United States. Second, neither book presented any systematically gathered evidence to support the allegations that were made. All one had to rely on were some scattered anecdotes throughout each tome. (Incidentally, it is not at all clear how a United States Navy admiral could become an expert on the educational system of a totally land-locked country). The second source of ferment in education was the launching of Sputnik I in 1957. It raised a number of questions about the state of American education.

Soon after the publication of these books, the launch of Sputnik I, and other similar books and reports in other countries, a group of educators began to meet in 1955. One product of that early meeting was a book on evaluation by F. Hoyt of Belgium. William Wall who was the Director of the National Foundation for Educational Research in England and Wales and the first chair of IEA (1958-1962) pushed the group to conduct a pilot study to test the feasibility of conducting multi-national studies of educational performance. The first formal meeting of the group was held in Eltham, England in 1958. Researchers attended it from England, the United States and several other countries. The group began to discuss ways in which research could contribute to the improvement of education not only in their countries but in other countries as well. It was at that meeting that the proposal that an exploratory study be undertaken to determine the feasibility of carrying out an empirical study of student performance in various school subjects was considered and decided.

Several work groups were formed to develop short tests in several subjects that could be administered to students (in their own language) to measure their performance in these subjects. Since there was no real funding for the study, it was decided that a representative from each country would take responsibility for translating the tests into their own native language, identify small judgment samples of students who would take the tests, administer them, score them and bring the results back to the main group. While the researchers were well aware of the limitations of their study, their prime concern was in testing the feasibility of such a study and how well researchers from differing systems of education could work together. It was also hoped that the researchers could begin to learn something about the educational systems of the other countries.

While the study was actually a feasibility exercise, there were two notions that were of concern to the researchers: (1) what could countries learn about their own educational system from such an exercise, and (2) what are the antecedents and correlates of achievement and how similar were these relationships across countries. A third notion that emerged a bit later was that there were certain variables that could only be tested in a multi-national study, e.g., the influence of age of starting school on subsequent performance. Some countries began schooling at age 5, others at age 6 and others at age 7. Testing the influence of these different starting ages on later achievement could not be carried out within the borders of a single country. It required a multi-national study. The ultimate goal of the work was to identify ways in which education could be improved throughout the world.

The group called itself IEA beginning in 1960 but was not legally incorporated until 1966. It carried out the study, which had been designed, and instruments developed over a couple of days. As a piece of research, it would never win any awards for scientific rigor, but as a feasibility study, it was successful enough to encourage the researchers to undertake what the English would call “a proper study”.

The researchers decided to conduct a study in the area of mathematics, believing that this would be the easiest subject in which to test because of the use of numbers instead of relying heavily on language. While the investigators were probably correct in their choice of mathematics, there were still enormous problems to overcome. A few additional educational systems were invited to participate and the first large-scale study of mathematics achievement was begun. (I use the term “educational systems” instead of countries because from the very outset of IEA some countries had two different systems of education, notably French and Flemish Belgium and England and Scotland in the United Kingdom).

It was decided that, along with the mathematics tests, attitude scales for students and questionnaires for students, teachers and school heads would be developed and administered. The levels to be tested were: (1) 13 year olds, (2) students in the eighth grade of school (largely 13 year olds), and (3) two groups of students in the last year of secondary school, students specializing in mathematics and students not specializing in mathematics. Countries were given the option of testing a group midway between these two grade levels. Furthermore, it was decided that probability sampling would be undertaken to obtain representative samples of each defined population and the samples should be large enough for sufficiently precise populations estimates and for statistical testing.

The researchers were strongly committed to carrying out a high quality research study. However, few, if any, of the researchers realized how large and complex a job they had taken on. This was only slowly discovered over the next three years. The fact that the study was carried out and completed as well as it was is due to the heroic efforts of a number of people. The report of the study received major coverage by the press all over the world. For the very first time there was evidence on which to make judgements and improvements in education rather than relying on mythology, unsupported allegations, and the whims of educational or political officials. This is perhaps the major contribution that IEA has made to research and education. Although there have been many misguided attempts to use evidence to make educational decisions, the notion that evidence should be used in making educational decisions is well established.

This, however, is not IEA’s only contribution to research and education. Let me cite what I regard as three other major contributions and some minor, but I believe important ones.

The three major contributions are easily remembered since they all begin with the letter “C”. They are competence, cooperation, and computers. First, there is competence. While IEA began as an informal association of researchers (IEA was not formally incorporated until 1966). It was able to enlist an impressive roster of specialists to work on IEA studies. The group included, among others, Ralph Tyler, John Tukey, Julian Stanley, James Coleman, John B. Carroll, Sten Henrysson and Herman Wold. As researchers, they had rather similar training and backgrounds that easily overcame differing national views. In addition to the IEA researchers, IEA was able to draw on many of the best minds in the world in the areas of education, psychology, sociology, economics, political science, and statistics as consultants. These

specialists were quite happy to work with IEA, usually for no payment. The studies that IEA was planning or carrying out engaged their interest and IEA has benefited greatly from their expertise. This was over and above the expertise that already existed in IEA which was considerable. The initial research workers in IEA included Benjamin Bloom, Robert Thorndike, C. Arnold Anderson, Douglas Pidgeon, Gilbert Peaker, David Walker, Torsten Husen, Sven Hilding, and Neville Postlethwaite among others.

The second major contribution of IEA was the promotion of international cooperation. While nations had differences, this did not spill over into IEA work. I believe that the reason for this is that the IEA members in the 1960s and 1970s were, for the most part, researchers and were working towards a common set of purposes. They had similar training and spoke a common language, the language of science and research. This is not to say that there weren't disagreements. There were, in fact, many of them from the very outset. Two examples that spring to mind are: (1) the discussion over the use of multiple-choice vs. open-ended questions in the mathematics tests for the first major IEA study, and the use of: (1) analysis of variance or (2) multiple regression analysis for data analysis. It wasn't until later that we found out that the two are mathematically identical. However, both were actually used in the data analysis of the first mathematics study. Another issue that continued to vex the IEA researchers over a substantial period of time was whether to define populations of students in terms of age or grade in school. There are good reasons to support either approach but IEA has found a way to deal with this by testing samples in adjacent grades that contain the great majority of an age group. Data can then be analyzed either by age or grade.

The third major contribution of IEA has been in the area of what can loosely be described as computers but includes considerably more than hardware. The first IEA mathematics study tested over 250,000 students in over 10,000 schools in twelve different educational systems. Analyzing that amount of data would be impossible without high speed, high capacity computers. It was a stroke of good fortune that such computers became generally available just as IEA was launching its first mathematics study. In addition, optical mark readers also became commercially available at the same time. It was thus possible to design a single answer sheet for each tested population and that could be used in all twelve participating systems of education. The answer sheet not only contained response areas for multiple choice mathematics items, but room for open-ended responses and responses to sixty-five attitudinal items. Questionnaires for students, teachers, and school principals were asked in paper and pencil form and had to be coded at each national center, entered onto punch cards and transmitted to the international data center for entry into computer.

IEA's use of the best computer hardware available should not mask the fact that the software developed for IEA studies is clearly at the forefront of research. The Data Entry Manager, developed for the IEA Reading Literacy Study of the early 1990s, is one of the keys to the success of all IEA studies since that time. As a testament to the success of IEA developed software, OECD's WEI and PISA projects make use of some of this software along with UNESCO's SACMEQ project.

IEA today is a far different organization than it was when its first began. It is an organization that is incorporated under Belgian law, has over 70 member systems of education, has a headquarters in Amsterdam with a permanent staff, produces newsletters and bulletins and has a process for pre-publication review of study reports. This last item arose from the fact that early IEA study reports were strongly criticized for being written in "researchese" and almost unreadable to all but a few researchers.

IEA has changed considerably from its earliest days in terms of the composition of its General Assembly, the main governing body of IEA. As previously noted, IEA originally consisted of a group of researchers. Currently, the IEA General Assembly consists largely of officials of ministries of education along with some researchers. This is good in the sense that it improves the chances for funding IEA studies and the organization as a whole through the payment of dues. Having officials from ministries of education on the General Assembly also provides a way for the various systems of education to decide what studies would be most useful to their work and what questions they want answered. The drawback to this is that some of the ministry officials sitting as IEA General Assembly members are not in a good position to judge research

proposals and have to rely on the expertise of IEA staff and other consultants to ensure good research. So far, this has worked rather well.

IEA has conducted a substantial number of studies over the past forty years. Some have had considerable impact internationally while others have had great impact in some countries but not in others. In viewing all IEA studies, it seems that one of the major outcomes of IEA research work is their contribution to what Torsten Husen, the second Chair of IEA, refers to as the “massification” of education. Prior to the IEA studies, a widely held view in many countries was that the pool of talent was limited and a selection system was needed to identify those students who were considered capable of completing a secondary education. There was a deep-rooted view in a number of countries where an economics of scarcity gave rise to a psychology of scarcity. According to this view, the pool of talent was regarded as being limited (quite limited in the view of some countries). The IEA studies of the 1960s through the 1990s helped destroy this myth. Those studies showed that many countries had underestimated their talent pool. IEA studies showed that many more students could be educated to the completion of secondary school without causing any harm to the performance of the highest ability. In other words, the yield of an educational system could be increased markedly by retaining a much greater percentage of an age group in school. To see evidence of this “massification” of education, one only has to turn to enrolment figures in a number of countries. The percentage of an age group completing a secondary education has increased markedly over the past 30 years, as shown in the IEA studies, with no loss in the achievement levels of the highest level of students.

Another major contribution of IEA studies was the development and use of measures of student achievement as outcomes of an educational system instead of previous work that usually relied on just enrollment figures. Furthermore, IEA introduced the notion of measuring achievement at different points in an educational system instead of just at a single point, usually the end of secondary education. This was a distinctly novel contribution since it allowed a country to track achievement throughout the levels of an educational system.

On the other hand, despite forty years of effort, IEA has not succeeded in developing suitable ways of measuring student achievement at the end of secondary school for comparative purposes. IEA has finally learned that there are so many differences in the various systems of education that any attempt at comparison is doomed to fail. The major differences between systems are: (1) the age of students in the last year of secondary education (the range is from age 16 to age 20), (2) the percentage of students in an age group still enrolled in school (over the years this has shown a range from 9% to over 80%), and (3) the amount of specialization that occurs in a number of systems (in science, for example, some students at the end of secondary school may have taken only one course in physics while in other educational systems student may have taken four or five courses in physics). Reluctantly, IEA has had to confine its studies to measuring student achievement during the years of compulsory education. This is still a notable accomplishment.

Sometimes overlooked is IEA’s development of tests that could be used fairly in a large number of educational systems. This involved the development of a newer technology of testing, one with an international dimension. The use of new technologies in testing such as item response methods, notably Rasch scaling, and imputation procedures that allow one to estimate a student’s performance, despite the fact that a student takes only a limited number of items, have worked well. These procedures have allowed for a common test to be used in large scale surveys, the use of a much larger pool of test questions to be used to cover a domain, and ways of linking tests at different levels within a study and with tests used in previous IEA studies. This allows for achievement comparisons over time. The recent repeat of the IEA Reading Literacy Study of the early 1990s is a case in point as well as the TIMSS Repeat study.

IEA has also done pioneering work in the use of complex sampling procedures in the conduct of its studies. These procedures have allowed IEA researchers to obtain unbiased estimates of student achievement and other variables with a high level of precision at a reasonable cost. Subsequent research projects such as

PISA and SACMEQ have benefited from IEA work in this area. A notable contributor to this work was the late Gilbert Peaker.

The idea of using research results to make recommendations about systems of education is hardly new. However, the IEA studies did provide an impetus to using such results since they contained pertinent information about other systems of education that could be used for comparison purposes.

Going further, some educational systems have used item analysis results from IEA tests to provide information to curriculum centers so that modifications could be made to curricula. Hungary has been a leader in carrying out this type of work.

IEA provided the first empirical evidence in support of C. P. Snow's construct of a two-culture theory. In its six-subject survey of the 1970s, IEA tested the same students in both science and literature at the 14 year old level and at the end of secondary school.. This provided a way of testing Snow's theory. The evidence clearly supported Snow's theory in the countries where such a phenomenon was hypothesized to exist and showed that it was not a universal phenomenon since several countries showed no virtually no evidence of the existence of two cultures, one scientific and one literary. However, in education, tradition frequently takes precedence over evidence. Countries that have a highly developed two cultures system continue to maintain it, despite Snow's warnings that it can hurt a society in the long run.

A contribution of IEA to educational research that has often been overlooked is its training function. The seminars held in Granna, Sweden during the 1970s has been the most visible aspect of IEA's training work. Probably more important, however, has been the hands-on training provided to national research coordinators as part of IEA's studies. The research coordinators who are charged with carrying out the work of an IEA study in their country often come to the task with limited training and experience. This was recognized early on and training sessions have been provided at each research coordinator's meeting so that they would be able to carry out the work of the study when they returned home. This training has increased the research capacity in many countries, often under the leadership of these research coordinators. Some have risen to prominent positions in research in their countries. An argument could be made that this may be IEA greatest contribution to research in education.

Identifying IEA's substantive contributions to education is not an easy task for two reasons. First, there have been so many detailed results from IEA studies that it is impossible to catalog them all. Second, isolating the contribution of results from IEA studies from other research studies is simply impossible. The best thing that can be said with regard to this is that when the results of IEA studies are in agreement with the results of other studies, they add to the strength of the results and their likely acceptance by the education community. In some cases, however, IEA results have had a considerable influence on educational policy and practice.

While the identification of specific IEA results to education is extremely difficult, if not impossible, some broad contributions can be identified. From the very first study in mathematics, IEA researchers have sought to determine the relationships between out of school variables such as home and community background with achievement. These relationships have been found to be rather substantial. Subsequent studies have confirmed this finding although the level of relationship varies across school subjects, ages and systems of education. The relationships are generally the highest in the area of reading and somewhat lower in mathematics. They are still significant and meaningful, however,

The importance of out-of-school factors is still being studied. In the TIMSS study, students were asked to report on whether they received extra instruction in mathematics and science (outside of the regular school program). This phenomenon has received different names from different investigators. These include extra tuition, extra-class instruction, and shadow education. It is only now starting to be studied in a systematic

and detailed way although it has been going on for years in various places. In Japan, it is called the Juku but has different names in different countries. For example, Japan has carried out several impressive studies in this area. We are only beginning to understand this phenomenon and its possible effect on achievement. We do know that from about 20% of students to a very high percentage of students receive extra school instruction, often at considerable cost to their families. In general, lower performing students report a higher level of such extra-school instruction than higher performing students. This has only been studied in mathematics and science so far, but the study of extra-school instruction can be expanded considerably to increase the number of subjects and frequency of such instruction. As noted previously, Japan has already conducted some interesting studies in this area but not as a result of IEA research.

IEA has added considerably to our knowledge about gender differences in achievement. Unfortunately, the picture that emerges is not that all clear cut. In some subjects at some age or grade levels, there are considerable differences in the achievement of boys and girls while in other subjects there are virtually no differences. Furthermore the differences are often inconsistent. Also, in some educational systems, there are boy/girl differences in achievement but not in all countries. It is a somewhat confusing picture. However, it does demonstrate that there are no inherent differences between boys and girls achievements in different school at different age or grade levels.

One possible explanation for these mixed results that has been put forward for gender differences in some subjects at some age or grade levels but not in others is that where a school subject is learned individually, e.g., at one's desk, especially at younger levels, gender differences are small or non-existent. This appears to hold for both reading and mathematics. However, when students at these lower levels work in groups, boys tend to outperform girls. In mixed groups, boys tend to take over the work from the girls, often pushing them aside or giving them menial tasks to perform. This was quite noticeable in the first computers in education stud where students worked in groups around a single computer. It is also noticeable, to some extent in science studies that involve laboratory work. To be sure, such behavior can stem from defined gender roles in different countries. Despite analysis of gender differences in virtually every study IEA has conducted over the past forty years, the issue of gender differences (or the lack of them in a number of instances) has not been adequately explained. We do know, from all the conflicting evidence, that gender differences in achievement appear to be more a cultural phenomena than a biological one.

While IEA was never regarded as an educational Olympics, the league tables of achievement that have resulted from IEA studies have received widespread attention; perhaps more attention than anything else in the IEA studies. These results have been interpreted (and misinterpreted) in many places. Poor results have often been a spur to reform an entire educational system. In the United States, for example, the results of the IEA six subject survey served as a basis for a call to reform U.S. education in the 1983 government publication, "A Nation at Risk".

Finally, consider for a moment the difference between the 1950s when all sorts of pronouncements and allegations were made about the state of education in many countries without any supporting evidence and the 1980s onwards where people have come to expect evidence to support pronouncements about the state of an educational system. I believe that IEA has contributed substantially to this change. In fact, it may be IEA's great contribution to education. One hopes that the educative function of IEA will continue to guide the process of basing educational policies on a foundation of evidence and not on whims or simple beliefs that have no evidential base.