

"Education Production Functions"

THE
INTERNATIONAL
ENCYCLOPEDIA OF
EDUCATION

SECOND EDITION

Volume 3

Editors-in-Chief

TORSTEN HUSÉN

University of Stockholm, Sweden

T. NEVILLE POSTLETHWAITE

University of Hamburg, Germany



PERGAMON

1994

- Howard A 1970 *Learning to be Rotuman: Enculturation in the South Pacific*. Teacher's College Press, New York
- Kidd D 1906 *Savage Childhood: A Study of Kafir Children*. A and C Black, London
- Leis P 1972 *Enculturation and Socialization in an Ijaw Village*. Holt, Rinehart and Winston, New York
- Little K 1967 *The Mende of Sierre Leone*, rev. edn. Routledge and Kegan Paul, London
- Mead M 1931 *Growing Up in New Guinea: A Comparative Study of Primitive Education*. William Morrow, New York
- Mead M 1954 *Coming of Age in Samoa: A Psychological Study of Primitive Youth for Western Civilisation*. New American Library, New York
- Mead M 1956 *New Lives for Old: A Cultural Transformation—Manus*. Gollancz, London
- Modiano N 1973 *Indian Education in the Chiapas Highlands*. Holt, Rinehart and Winston, New York
- Nadel S F 1942 *A Black Byzantium: The Kingdom of Nape in Nigeria*. Oxford University Press, London
- Nash M 1961 Education in a new nation. The village school in upper Burma. *Int. J. Comp. Sociol.* 2(3): 135-43
- Ottenberg S 1989 *Boyhood Rituals in an African Society: An Interpretation*. University of Washington Press, Seattle, Washington
- Raum O 1940 *Chagga Childhood: A Description of Indigenous Education in an East African Tribe*. Oxford University Press, London
- Read M 1959 *Children of Their Fathers: Growing up Among the Ngoni of Nyasaland*. Methuen, London
- Redfield R 1943 Culture and education in the midwestern highlands of Guatemala. *Am. J. Sociol.* 48: 640-48
- Tylor E 1871 *Primitive Culture: Researches into the Development of Mythology, Philosophy, Religion, Art and Custom*. J Murray, London
- Whiting J 1941 *Becoming a Kwoma: Teaching and Learning in a New Guinea Tribe*. Yale University Press, New Haven, Connecticut
- Williams T 1958 The structure of the socialization process in Papago Indian society. *Social Forces* 36(3): 251-56
- Wolcott H 1967 *A Kwakiutl Village and School*. Holt, Rinehart and Winston, New York

S. Juan

Education Production Functions

Although research into the determinants of students' achievement takes various approaches, one of the most appealing and useful is what economists call the "production function" approach. (In other disciplines it is known as the input-output or cost-quality approach.) In this, attention is focused primarily on the relationship between school outcomes and measurable inputs into the educational process. If the production function for schools were known, it would then be possible to predict what would happen if resources were added or subtracted, and to analyze what actions should be taken if the prices of various inputs were to change. The problem is that the production function for education is unknown and must be inferred from data on students and their schools.

1. The Coleman Report and its Influence

The origin of estimating input-output relations in schools is usually traced to the monumental United States study *Equality of Educational Opportunity* (commonly known as the Coleman Report—Coleman et al. 1966). Designed explicitly to study equity, this report was the United States Office of Education's response to a requirement of the Civil Rights Act of 1964, namely to investigate the extent of inequality (by race, religion, or national origin) in the nation's schools. The study's fundamental contribution was to direct attention to the distribution of student performance—the outputs under consideration in this entry. Instead of addressing questions of inequality simply by producing an inventory of differences among schools and teachers by race and region of the country, the Coleman Report sought to explain those differences; it delved into the relationship between inputs and outputs of schools.

The Coleman Report was widely interpreted as finding that schools have little importance in determining student achievement. Families and, to a lesser extent, peers were seen to be the primary determinants of variations in performance. The findings were clearly controversial (see critiques by Bowles and Levin 1968 and Hanushek and Kain 1972) and immediately led to an extensive research effort to compile additional evidence about the relationship between school resources and school performance. As described below, the common interpretation of the Coleman Report in fact results from a misinterpretation of the statistical findings.

The underlying model guiding the Coleman Report and most subsequent studies is straightforward. It postulates that the output of the educational process—that is, the achievement of individual students—is directly related to a series of inputs. Policymakers directly control some of these inputs; for instance, the characteristics of schools, teachers, and curricula. Other inputs—those of families and friends plus the innate endowments or learning capacities of the students—generally cannot be affected by public policy. Further, although achievement is usually measured at discrete points in time, the educational process is cumulative; past inputs affect students' current levels of achievement.

Starting with this model, statistical techniques, typically some form of regression analysis, are employed to identify the specific determinants of achievement and to make inferences about the relative importance of the various inputs into student performance. The accuracy of the analysis and the confidence the answers warrant depend crucially on a variety of issues regarding measurement and technical estimation. This summary sets aside these issues (see Hanushek 1979, 1981, 1986). Instead it highlights the overall findings and the major unanswered questions from this research.

Most studies of educational production relationships measure output by students' scores on standardized achievement tests, although significant numbers have used other quantitative measures, such as student attitudes, school attendance rates, and college continuation or dropout rates. The general interpretation is that they are all plausible indicators of future success in the labor market.

The reason for concentrating on achievement for students in school is straightforward. The policy question centers on how different teachers and school resources affect student performance. It would be generally impractical to have to wait a decade or two after observing educational inputs to measure any subsequent outcomes that will be related to those inputs. Data and analytical necessities dictate concentration on immediate measures of student performance such as test scores. Other research, however, indicates that these in-school measures are related to subsequent performance in the labor market and that they are thus reasonable proxies of economically pertinent skills.

Test measures have been included in standard models that explain earnings differences in the population. Studies of adult earnings in developed countries typically show significant direct effects of achievement. These come, however, in statistical models that also include years of schooling, and test achievement is an important determinant of continuation in schooling, implying an important additional indirect effect. The evidence on returns to different measured skills has tended to be stronger in developing countries (see review in Harbison and Hanushek 1992). Second, studies have found direct links with productivity, particularly in agriculture (Welch 1970, Jamison and Lau 1982). In short, there is reasonably broad support for the notion that school quality as measured during schooling is directly related to productivity and earnings when students enter the labor force. Thus, although most attention is focused on the ability of schools to raise students' academic performance, there is reason to interpret this in the broader context of increasing economic performance of the students and of the overall economies.

Empirical specifications of production functions have varied widely in details, but they have also had much in common. Family inputs tend to be measured by sociodemographic characteristics of families, such as parental education, income, and family size. Peer inputs, when included, are typically aggregate summaries of the sociodemographic characteristics of other students in the school. School inputs include measures of the teachers' characteristics (education level, experience, sex, race, and so forth), of the schools' organization (class sizes, facilities, administrative expenditures, and so forth), and of district or community factors (e.g., average expenditure levels). Except for the original Coleman Report,

most empirical work has relied on data, such as the normal administrative records of schools, that were constructed for other purposes.

2. Schools, Expenditures, and Achievement in the United States

The production function approach has been employed broadly to investigate the effect on school performance of the core factors that determine expenditure on education. Instructional expenditures make up about two-thirds of total school expenditures in the United States. Instructional expenditures are, in turn, determined mostly by teacher salaries and class sizes. Finally, in most United States school districts, teacher salaries are directly related to the years of teaching experience and the educational level of the teacher. Thus, the basic determinants of instructional expenditures in a district are teacher experience, teacher education, and class size. Most studies, regardless of what other school characteristics might be included, analyze the effect of these factors on outcomes. (These are also the factors most likely to be found in any given data set, especially if the data come from standard administrative records.)

Because the analyses have such common specifications, the effects of the expenditure parameters can easily be tabulated. An exhaustive search through 1988 publications uncovered 187 separate qualified studies found in 38 separate published articles or books. (Qualified studies satisfy certain minimal quality standards—being published in a book or journal, providing direct information about the effects of school resources on student performance, and providing information about the statistical significance of any findings.) These studies, while restricted to public schools, cover all regions of the United States, different grade levels, different measures of performance, and different analytical and statistical approaches. About one-third draw their data from a single school district, while the remaining two-thirds compare school performance across multiple districts. A majority of the studies (104) use individual students as the unit of analysis; the remainder rely upon aggregate school, district, or state level data. The studies are split about evenly between primary schooling (Grades 1–6) and secondary schooling (Grades 7–12). Over 70 percent of the studies measure school performance by some kind of standardized test. However, those that use nontest measures (such as dropout rates, college continuation, attitudes, or performance after school) are for obvious reasons concentrated in studies of secondary schooling. There is no indication that differences in sample and study design lead to differences in conclusions.

Table 1 summarizes the expenditure components of the 187 studies (Hanushek 1989). Since not all

Table 1
Summary of estimated expenditure parameter coefficients from 187 studies of educational production functions: United States

Input	Number of studies	Statistically significant			Statistically insignificant			Unknown sign
		Total	+	-	Total	+	-	
Teacher-pupil ratio	152	27	14	13	125	34	46	45
Teacher education	113	13	8	5	100	31	32	37
Teacher experience	140	50	40	10	90	44	31	15
Teacher salary	69	15	11	4	54	16	14	24
Expenditures/pupil	65	16	13	3	49	25	13	11
Administration	61	8	7	1	53	14	15	24
Facilities	74	12	7	5	62	17	14	31

Source: Hanushek 1989

studies include each of the expenditure parameters, the first column in the table presents the total number of studies for which an input can be tabulated. For example, 152 studies provide information about the relationship between the teacher-pupil ratio and student performance. The available studies all provide regression estimates of the partial effect of given inputs, holding constant family background and other inputs. These estimated coefficients have been tabulated according to two pieces of information: the sign and the statistical significance (5 percent level) of the estimated relationship. Statistical significance is included to indicate confidence that any estimated relationship is real and not just an artifact of the sample of data employed.

According to both conventional wisdom and generally observed school policies, each tabulated factor should have a positive effect on student achievement. More education and more experience on the part of the teacher cost more and are presumed to improve individual student learning; smaller classes (more teachers per student) are also expected to be beneficial. More spending in general, higher teacher salaries, better facilities and better administration should also lead to better student performance. The quantitative magnitudes of estimated relationships are ignored here; only the direction of any effect is analyzed.

Of the 152 estimates of the effects of class size, only 27 are statistically significant. Of these, only 14 show a statistically significant positive relationship, whereas 13 display a negative relationship. An additional 125 estimates show that class size is not significant at the 5 percent level. Nor does ignoring statistical significance help to confirm the benefits of small classes. By a margin of 46 to 36 the insignificant coefficients are negative, the wrong sign according to conventional wisdom.

The entries for teacher education tell a similar story. The statistically significant results are split between positive and negative relationships, and in a vast majority of cases (100 out of 113) the estimated

coefficients are statistically insignificant. Forgetting about statistical significance and looking just at estimated signs again does not make a case for the importance of added schooling for teachers.

Teacher experience is possibly different. A clear majority of estimated coefficients point in the expected direction, and about 29 percent of the estimated coefficients are both statistically significant and of the conventionally expected sign. These results, however, are hardly overwhelming; they only appear strong relative to the other school inputs. Moreover, they are subject to interpretative questions. Specifically, these positive correlations may result from senior teachers having the ability to locate themselves in schools and classrooms with good students. In other words, causation may run from achievement to experience and not the other way around.

Overall, the results are startlingly consistent. No compelling evidence emerges that teacher-pupil ratios, teacher education, or teacher experience have the expected positive effects on student achievement. It cannot be stated with confidence that hiring teachers with more education or having smaller classes will improve student performance. Teacher experience appears only marginally stronger in its relationship.

The remaining rows of Table 1 summarize information on other expenditure components, including administration, facilities, teacher salaries, and total expenditure per student. The quality of administration is measured in a wide variety of ways, ranging from characteristics of the principal to expenditure per pupil on noninstructional items. Similarly, the quality of facilities is identified through both spending and many specific physical characteristics. The absence of a strong relationship between these two components and performance may result in part from variations in how these factors are measured. If only because of the preponderance of positive signs among the significant coefficients, administration appears marginally stronger in its relationship than

facilities. Nevertheless, the available evidence on both again fails to support convincingly the conventional wisdom.

Finally, explicit measures of teacher salaries and expenditure per student are tabulated. These measures are less frequently available and are more difficult to interpret because they are included along with their underlying determinants. Nevertheless, it is not surprising that they do not suggest that they have a potentially important role in determining achievement. After all, the underlying components of this expenditure were themselves unrelated to achievement.

The research reveals no strong or systematic relationship between school expenditures and student performance. This is the case both when expenditure is decomposed into underlying determinants and when it is considered in the aggregate.

3. Other Inputs into Education—United States Studies

Since the publication of the Coleman Report, intense debate has surrounded the fundamental question of whether schools and teachers are important to the educational performance of students. That report has been commonly interpreted as finding that variations in school resources explain only a negligible portion of the variation in students' achievement.

A number of studies provide direct analyses of this overall question of differential effectiveness of teachers and schools (Hanushek 1971, 1992; Murnane 1975; Armor et al. 1976; Murnane and Phillips 1981). They do this by estimating differences in the average performance of each teacher's (or school's) students after allowing for differences in family backgrounds and initial achievement scores. The findings are unequivocal: teachers and schools differ dramatically in their effectiveness. The formal statistical tests employed in these studies confirm that there are striking differences in average gain in student achievement across teachers. The faulty impressions left by the Coleman Report and by a number of subsequent studies about the importance of teachers have resulted primarily from a confusion between the measures of effectiveness and true effectiveness itself.

These production function analyses have also investigated a wide variety of other school and non-school factors. Although it is difficult to be specific in any summary of other factors because the specifications are quite idiosyncratic, three generalizations are possible. First, family background is clearly very important in explaining differences in achievement. Second, while considerable attention has been given to the characteristics of peers or other students within schools, the findings are ambiguous. Finally, studies have examined many additional measures of the effects of schools, teachers, curricula, and, es-

pecially, instructional methods on achievement, but few consistent results have emerged.

4. Schooling in Developing Countries

Research on school achievement in developing countries is less extensive, less rigorous, and more difficult to interpret than that for the United States. Nevertheless, some conclusions can be drawn from school operations in developing countries from such research.

Dissimilar findings about the determinants of school performance in developing countries, as contrasted with developed countries, might be expected. The dramatic differences in the level of educational support provided by families and schools imply that the educational production process could be very different in developed and developing countries. In particular, while the effect of marginal resources on achievement may be hard to discern when average school expenditure in the United States is US\$6,000 per year per pupil, they might be much larger and more noticeable when expenditure is one tenth or one hundredth that level.

At the same time, the standards of data collection and analysis are so variable that the results from this work tend to be uncertain. Much of the analysis of input-output relationships for developing countries is not published in standard academic journals, and thus it does not have that basic level of quality control. Even more important, the data for many of these studies do not come from regular collection schemes, are difficult to check for quality, and miss key elements of the educational process.

Different researchers have attempted to summarize key aspects of these studies, frequently providing qualitative discussions of the analyses, their results, and their interpretation. Here, however, an overall quantitative summary of the available analyses will be presented, which parallels that for the United States studies. The starting point is the comprehensive review of studies by Fuller (1985). This is supplemented by additional studies that have appeared since that review or were omitted from it. There are limitations, however. Because this discussion and analysis relies chiefly on secondary materials, the reporting of results has to be accepted. Consequently, the results cannot be presented in the same depth as those for the United States. Additionally, there is virtually no control over the selection of papers (i.e., according to explicit minimal quality standards) or over the interpretation of the statistical results.

A total of 96 underlying studies form the basis for the analysis (about half the number utilized in the United States analysis). Table 2 divides the available studies into statistically significant (by sign) and statistically insignificant. (The insignificant findings, unfortunately, cannot be divided by direction of

Table 2

Summary of estimated expenditure parameter coefficients from 96 studies of educational production functions: developing countries

Input	Number of studies	Statistically significant		Statistically insignificant
		+	-	
Teacher-pupil ratio	30	8	8	14
Teacher education	63	35	2	26
Teacher experience	46	16	2	28
Teacher salary	13	4	2	7
Expenditures/pupil	12	6	0	6
Facilities	34	22	3	9

Source: Harbison and Hanushek 1992

effect.) The table is laid out similarly to that for the United States studies. It begins with the characteristics directly related to instructional expenditure per student and then treats other attributes of schools.

The studies differ from the United States studies in terms of the overall significance of the estimated school effects. Simply put, compared with the results presented in Table 1, a higher proportion of the tabulated coefficients for the 96 studies in developing countries is statistically significant. (It must be emphasized, however, that the proportion of results that are "correct"—statistically significant by conventional standards and in the right direction—never reaches two-thirds; moreover, the general conclusion of no strong evidence of a systematic relationship between these factors and performance will not change.) The relative robustness in statistical findings could reflect analysis of settings where there is either greater variation in the tabulated educational inputs or greater sensitivity to these inputs by students. Alternatively, the differences could reflect attributes and, specifically, biases of the analyses themselves.

The evidence in Table 2 from developing countries provides no support for policies of reducing class sizes. Of the 30 studies investigating teacher-pupil ratios, only eight find statistically significant results supporting smaller classes; an equal number are significant but have the opposite sign; and almost half are statistically insignificant. These findings are particularly interesting because class sizes in the studies of developing countries are considerably more varied than those in the United States studies and thus pertain to a wider set of environments.

The analysis of the effect of teacher experience yields results that are roughly similar to those in the United States studies. Although 35 percent of the studies display significant positive benefits from more teaching experience (the analogous figure for United States studies is 29 percent), the majority of the estimated coefficients still are statistically insignificant. The primary difference between the two sets of tabulations arises from the relative support implied

for the different school inputs. The United States studies are the most supportive of the conventional wisdom regarding the effects of teacher experience on performance. Similar support compared to other factors is not found in developing country studies.

The results for teacher education, on the other hand, diverge in relative terms from those seen for the United States. A majority of the studies (35 out of 63) support the conventional notion that providing more education for teachers is valuable. In the United States studies, teacher education provided the least support of all the inputs for the conventional wisdom. Although still surrounded by considerable uncertainty (since 26 estimates are insignificant and 2 display significantly negative effects), these noticeably stronger results in developing countries clearly suggest a possible differentiation by stage of development and general level of resources available.

The teacher salary findings in developing countries contain no compelling support for the notion that better teachers are systematically paid more. Since the studies aggregate findings across very different countries, school organizations, and labor markets, however, it is difficult to take these results too far. For policy purposes, it would be desirable to seek information on what happens if the entire salary schedule is altered (as opposed to simply moving along a given schedule denominated, say, in experience, education, or other attributes of teachers). It is impossible to distinguish, however, between studies reflecting differences in schedules and those reflecting movements along a schedule.

Data for total expenditure per pupil are rarely available in analyses of education in developing countries. The 12 studies in which estimates can be found are evenly split between statistically significant and statistically insignificant. Given questions about the quality of the underlying data, not too much should be inferred from the findings for direct expenditure measures.

One of the clearest divergences between the two sets of findings is for facilities, again suggesting that differences in school environments are of some

importance. The measures of facilities in developing countries (which incorporate a wide range of actual variables in specific studies) indicate more likely effects on student performance than found in United States studies. Some 22 of the 34 investigations demonstrate support for the provision of quality buildings and libraries.

In summary, the results of studies in developing countries do not make a compelling case for specific input policies. They do, however, indicate that direct school resources might be important in developing countries. Nevertheless, as in the United States research, the estimated models of educational performance undoubtedly fail to capture many of the truly important inputs to the educational process.

5. Other Factors—Developing Countries

As with the United States studies, a variety of other factors has been investigated in the course of the analyses of developing countries, including an assortment of curriculum issues, instructional methods, and teacher training programs. Many of these are difficult to assess (at least in a quantitative, comparative way) given the evidence from many countries and the probable importance of local institutions.

One intervention that has widespread endorsement, although as much for conceptual reasons as for solid empirical ones, is the provision of textbooks. The relationship of textbooks and writing materials to student performance is found with reasonable consistency to be important in developing countries, but there are relatively few studies of this (see Lockheed and Hanushek 1988, Lockheed and Verspoor 1989).

Investigations of technological or organizational differences have led to mixed results. Because of scattered settlement in many rural areas, several approaches to "distance education" have been investigated. In three extensive investigations (Nicaragua, Kenya, and Thailand), the use of interactive radio has proved effective (Lockheed and Hanushek 1988). However, this conclusion should not be generalized to all possible uses of new technology. In particular, there is little evidence at this time that the widespread introduction of computers is sensible (Lockheed and Verspoor 1989).

6. Conclusion—Implications for Policy

Somewhat surprisingly, perhaps, the available research from both developed and developing countries leads to many of the same conclusions. Two potential policy conclusions spring immediately from the overall results. First, since within the current institutional structure expenditures are not systematically related to performance, policies should not be dictated simply on the basis of expenditure.

Second, since common surrogates for teacher and school quality—class size, teachers' education, and teachers' experience being among the most important—are not systematically related to performance, policies should not be dictated simply on the basis of such surrogates.

Moreover, these results strongly suggest that policies based solely on inputs—such as general reductions in class size or uniform increases in spending—are unlikely to be successful. This underscores the importance of moving toward policies based on performance; that is, output-based policies. Such policies would emphasize the importance of performance incentives (see, e.g., Chubb and Hanushek 1990).

See also: Macroeconomics of Educational Time and Learning; Political Economy of Educational Production; Joint Production of Education; Educational Efficiency and Effectiveness. Concepts of; Microeconomics of School Production; Quality of Education in Developing Nations: Policies for Improving; Cost-Benefit Analysis

References

- Armor D et al. 1976 *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools*. R-2007-LAUSD, Rand Corporation, Santa Monica, California
- Bowles S, Levin H M 1968 The determinants of scholastic achievement: An appraisal of some recent evidence. *J. Hum. Resources* 3(1): 3-24
- Chubb J E, Hanushek E A 1990 Reforming educational reform. In: Aaron H J (ed.) 1990 *Setting National Priorities: Policy for the Nineties*. Brookings Institution, Washington, DC
- Coleman, J S et al. 1966 *Equality of Educational Opportunity*. US Government Printing Office, Washington, DC
- Fuller B 1985 *Raising School Quality in Developing Countries: What Investments Boost Learning?* Report No. ED7, Education and Training Series, World Bank, Washington, DC
- Hanushek E A 1971 Teacher characteristics and gains in student achievement: Estimation using micro-data. *Am. Econ. Rev.* 61(2): 280-88
- Hanushek E A 1979 Conceptual and empirical issues in the estimation of educational production functions. *J. Hum. Resources* 14(3): 351-88
- Hanushek E A 1981 Throwing money at schools. *Journal of Policy Analysis and Management* 1(1): 19-41
- Hanushek E A 1986 The economics of schooling: Production and efficiency in public schools. *J. Econ. Lit.* 24: 1141-77
- Hanushek E A 1989 The impact of differential expenditures on school performance. *Educ. Res.* 18(4): 45-51
- Hanushek E A 1992 The trade-off between child quantity and quality. *J. Pol. Econ.* 100(1): 84-117
- Hanushek E A, Kain J F 1972 On the value of "equality of educational opportunity" as a guide to public policy. In: Mosteller F, Moynihan D P (eds.) 1972 *On Equality of Educational Opportunity*. Random House, New York
- Harbison R W, Hanushek E A 1992 *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*. Oxford University Press, New York

- Jamison D, Lau L 1982 *Farmer Education and Farm Efficiency*. Johns Hopkins University Press, Baltimore, Maryland
- Lockheed M E, Verspoor A 1989 *Improving Primary Education in Developing Countries: A Review of Policy Options*. World Bank, Washington, DC
- Lockheed M E, Hanushek E A 1988 Improving Educational efficiency in developing countries: What do we know? *Compare* 18(1): 21-38
- Murnane R J 1975 *Impact of School Resources on the Learning of Inner City Children*. Ballinger, Cambridge, Massachusetts
- Murnane R J, Philips B 1981 What do effective teachers of inner-city children have in common? *Social Science Research* 10(1): 83-100
- Welch F 1970 Education in production. *J. Pol. Econ.* 78(1): 35-59

E. A. Hanushek

Educational Achievement: Comparative Studies

This entry describes the major international educational achievement studies undertaken since the 1960s: those conducted by the International Association for the Evaluation of Educational Achievement (IEA), those by the International Assessment of Educational Progress (IAEP), and the Michigan studies. Certain issues concerned with the undertaking of comparative studies in educational achievement are examined. These include the definition of the target populations, whether countries or systems of education should be the focus of attention, and the problem of translation.

For the most part "the earliest descriptions of foreign educational practices were generally piecemeal observations by curious and interested travellers" (Eckstein 1988 p.7). With the development of national school systems in the nineteenth century scholars began to study aspects of schooling in other countries with the twin motives of both informing their own compatriots about them and using such information to improve practice in their native countries. In particular, they examined the structures of schooling, staffing, instructional methods, and the curriculum (Brickman 1988, Postlethwaite 1988).

In the 1950s international agencies such as the United Nations Educational, Scientific, and Cultural Organization (UNESCO), in particular its International Bureau of Education (IBE), and the Organisation for Economic Co-operation and Development (OECD) intensified the exchange and accumulation of data relating to the different patterns of educational organization, curricula, and teaching methods. There was a growing awareness of the role that formal education played in promoting—or hindering—social and economic development. This was coupled with the realization that few countries had sufficient resources or manpower to satisfy the

growing demand for education. Economists looked in vain for indices of educational "productivity." They tended to use the proportion of an age group graduating from various levels of a school system as an index but recognized that what had actually been learned by these proportions of students could vary considerably from country to country.

It was Anderson (1961) who pointed out the need for some objective measure of "outcomes" using the methods of quantitative assessment developed in educational psychology.

1. First Studies of Achievement

Pidgeon (1958) conducted a study of basic achievement of 11-year olds in reading, arithmetic and non-verbal ability in England and Wales, Queensland, Australia, and Central California on samples of convenience. The findings were of interest and showed that such a study could be of value in comparative education. Pidgeon suggested that the possible factors accounting for divergences among the three systems were differences in the methods and approaches employed in teaching the different subjects.

In 1958, a group of researchers (including Anderson and Pidgeon) met to consider the possibility of undertaking a study of measured "outcomes." They decided to conduct a feasibility comparative study which was carried out in the period 1959-61. The first sentence of the report (Foshay et al. 1962 p.7) was: "If custom and law define what is educationally allowable within a nation, the educational systems beyond one's national boundaries suggest what is educationally possible." The aims of this exploratory study were:

- (a) to see whether some indications of the intellectual functioning behind responses to short-answer tests could be deduced from an examination of the patterning of such responses from many countries with different languages and cultural backgrounds;
- (b) to discover the possibilities and difficulties of a methodological and administrative kind attending a large-scale international survey.

In this study in the early 1960s, tests of reading comprehension, mathematics, science, geography, and nonverbal ability were administered to samples of 13-year olds in Belgium, England, Finland, France, the Federal Republic of Germany, Israel, Poland, Scotland, Sweden, Switzerland, the United States, and Yugoslavia. The overall findings were that:

- (a) A large-scale project, which depends on similarities in technical and philosophical assumptions in education and in measurement, can be done.