DOCUMENT RESUME

ED 055 321 EA 003 653

AUTHOR Kravetz, Nathan
TITLE The Evaluation of Educational System Outputs: An Exploratory Study. A Research Project.
INSTITUTION United Nations Educational, Scientific, and Cultural Organization, Paris (France). International Inst. for Educational Planning.
REPORT NO IIEP-RP-8
PUB DATE 19 Oct 70
NOTE 107p.
AVAILABLE FROM Publications Officer, International Institute for Educational Planning, 7, rue Eugene-Delacroix, Paris 16e (France) ($.75)

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS Curriculum Evaluation; Educational Needs; Educational Objectives; Educational Planning; *Educational Research; Educational Testing; *Evaluation; *Evaluation Criteria; *Evaluation Techniques; Input Output Analysis; Performance Specifications; Problem Solving; Research Proposals; Student Development

ABSTRACT

This 4-part document provides an overview of educational evaluation. Part I outlines the problems in educational evaluation, distinguishing between evaluations of student performance (objective attainment) and those of system operation (cost-effectiveness). Part II reviews existing practices in educational evaluation, focusing on the evaluation of students. The purposes, frequency, uses, some implications of testing, and the types of instruments used are also described. Part III indicates some national and regional agencies, international services, and individuals who have been working on the problems in educational evaluation. Part IV suggests areas for potentially fruitful research, noting the varities of possible studies and the directions they could take. (RA)

IIEP/RP/8
Paris, 19 October 1970
Original : English

INTERNATIONAL INSTITUTE FOR EDUCATIONAL PLANNING
(established by Unesco)
9, rue Eugène-Delacroix, Paris 16e

A Research Project on

THE EVALUATION OF EDUCATIONAL SYSTEM OUTPUTS: AN EXPLORATORY STUDY

by

Professor Nathan Kravetz

Acknowledgements:

I wish to express my appreciation and thanks for the efforts of Miss Beate Frank, my research assistant, who organized and annotated much of the material in this report, and to Mrs. Jacqueline Schwab, Institute librarian, who was particularly helpful in obtaining materials.

NK

C O N T E N T S

		Page
PART I :	THE PROBLEMS OF EDUCATIONAL SYSTEM EVALUATION	1
	Introduction	1
	Purpose of the study	1
	The formal versus the 'informal' system	4
	Levels of evaluation	6
	Evaluative targets and decision-making	8
	Criteria for study of evaluation	10
	Goal-setting and goal characteristics	14
	Goals become operational as curricula	17
	The relevance of curricula	20
	Evaluation of goal achievement: cognitive skills	21
	Testing for incremental gains: the value-added concept	24
	Evaluation of goal achievement: non-cognitive skills	26
	Student participation as an indication of non-cognitive development	28
	Student creativity as an indicator of non-cognitive development	29
	Indicators for system evaluation	30
	Summary	35
PART II :	CURRENT PROCEDURES FOR EDUCATIONAL EVALUATION	39
	Introduction	39
	Existing conditions of practice in the evaluation of school system outputs	39
PART III :	CURRENT RESEARCH ON EVALUATION OF EDUCATIONAL OUTPUTS	53
	Research presently being carried out	53

	Page
Indications of developing research	77
PART IV : PROPOSALS FOR FURTHER STUDY	81
BIBLIOGRAPHY	83

PART I

THE PROBLEMS OF EDUCATIONAL SYSTEM EVALUATION

Introduction

The following survey report was undertaken to serve as a foundation for further study and more intensive research on the evaluation of educational system outputs.

It was not intended that this report should provide educational planners and decision-makers with a resolution of problems. Rather it is a discussion of the situations from which the question of system evaluation arises. It is also a broad examination of the points of view which have been expressed and of the means which have been proposed and taken to deal with educational evaluation.

The intent is to turn a page in the story of educational development planning by focusing upon the pressing need to know more about the value of efforts which have been made and which are continually in the process of re-consideration for further decisions. This report attempts to ascertain the potential for further study of the basic problems of evaluation in education and most specifically of the quality and nature of the outputs of school systems.

There is continual popular demand for change and innovation in education, as if these were either magic words in themselves, or as if they represented processes which can be immediately applied for instant results. Given the current considerable demand also for more education and for more 'relevant' education, it is ever more important that evaluation be considered as a sine qua non of educational planning. The introduction of change in itself is only a part of planning, and the incorporation, or appending, of evaluation processes to educational procedures is therefore a basic requirement.

Purpose of the study

The following study will consider the demands for evaluation as they refer to individual needs and to national requirements. We will review the motivation for such demands and the means now in use and those proposed for educational evaluation on a system basis. Relevant research and research proposals will be reviewed.

We will also indicate the national and regional agencies, the research centres, the international services, and the leading individuals who have been working on the problems of evaluation of school systems.

Finally, we will suggest areas for potentially fruitful research on this subject, noting the varieties of the studies which might be made and the directions which such work could take.

The urgent interest in evaluating the results, or outputs, of education stems from the basic fact of the costs of education. In every country, developed or developing, costs for education represent large, if not major, portions of national budgets. The efforts and time of increasing numbers of people are represented both by the student population and by the teachers and officials who are engaged in school system service.

> "The concept of internal efficiency requires some assumptions about the objectives or measurable achievements of an educational system ... Results or objectives are related to the measurable efforts needed to achieve them: for example, teachers' time, students' time, use of facilities and other resources. These inputs can be measured as expenditures or costs ... The expenditure or cost of a certain course of learning can then be related to the attainment or objective."[1]

Economically, then, the importance of the education investment is significant, and is so recognized by economists in spite of the relative difficulty they have in determining linear connexions between increased investment in education and increased economic development.

---

1 Friedrich Edding, "Educational Resources and Productivity", in George Z.F. Bereday, (editor), Essays on World Education: The Crisis of Supply and Demand, New York: Oxford University Press, 1969, pp. 22-23.

However, for national populations education and its presumed effectiveness promises much. There is the direct certification of individuals for useful and well-paying occupations. There is the direct relationship to social status which is recognized as achieved by individuals who hold certificates and degrees. There is the overall tendency of an educated population to meet the requirements of manpower in a developing and changing economic picture. Also, a well-educated, broadly developing population, well-employed, provides a national sense of well-being and of constructive forward movement. A country of people who are receiving new educational opportunity and who find an open system is seen as striving, developing, and, perhaps, even succeeding.

Yet in all systems, in all countries, there are evidences of lack of success. Within this context, questions begin to be raised. When large investments are made in education, drop-outs and repeaters must represent a cost. We have the suggestion of Rist that:

> "The success of an educational institution ... should not be measured by the treatment of the high-achieving students, but rather by the treatment of those not achieving."[1]

When certificates and degrees are awarded, each recipient, though a 'winner', must be somewhere on a continuum from the very best level down to the merely passing. The question of quality arises, since sooner or later some use is to be made of such certificates and degrees. Whether used in the economy, in teaching, or in continued education, they presume to indicate the educational and intellectual achievements of their holders.

---

1 Ray C. Rist, "Student Social Class and Teacher Expectations", Harvard Educational Review, Vol. 40, No. 3, August 1970, p. 448.

Further questions arise from parents and teachers as to how much effort should be made to reduce drop-outs and repeaters. Proposals are made for varieties of new or re-discovered procedures in education: remedial courses, smaller classes, use of audio-visual (and tele-visual) aids, use of programmed learning devices (books and machines) and the relaxation of required courses so as to establish new standards by which individual success in education can be determined. Current demands for relevance in curricula, for courses which matter 'here and now', and for new learning modes to go with the realities of the 'here and now', are raised because there is general agreement that when students are motivated by the curriculum and by the methods of instruction they will succeed in reaching their educational goals.

The formal versus the 'informal' system

The development of demands for educational reform has come principally in developed countries where the evidence of the existence of two, competing, educational systems has been quite strong. On the one hand, there is the formal structure with its sequences of levels, its examination system, its screening-out processes, its concentration on measurable evidence of educational success. The national system, or establishment, provides forward movement to some, eliminates others, and makes a number of choices available to still others, while operating on the continuing assumption that it is perpetuating national traditions and culture and establishing and maintaining ethics and morals.

On the other hand, there is a quite 'informal' educative apparatus making itself felt with varying effectiveness upon the population, particularly that portion which is of school age. We refer to the mass media, newspapers, magazines, books, radio, moving pictures, television, all of which are recognized as affecting the lives of those whom they reach. This is not to say that such media are totally

effective, since if they were we would not have the evidence that such devices can be and are either physically turned off or emotionally 'tuned-out' even in authoritarian States.

Nevertheless, in countries where such media are operative an abundance of impressions is made in contrast perhaps with those made or desired by the school system as it is formally constituted. It is particularly in the broad areas of religion, sexual behaviour, war, and personal ethics that such divergences are most characterized. It is also in such contexts that the objectives of national systems are most tenuous, undefined, and hesitant, and where, necessarily, evaluation of effectiveness is least in evidence. This condition is likely to be as true in developing countries as it is in the developed ones.

Yet another aspect of the informal educational system is that represented by the apparent relationship between the school-age population and their elders. In this regard what the school system teaches about morality is often negated in practice by the 'older' generation, more accurately designated as the 'holders of power'. Generalizations about human dignity are shown to be ephemeral when the treatment of labour, or of minority groups, or of women, is examined. Such clear contrasts between reality in the behaviour of controlling adults and the precepts of the organized educational system provide a simple, if not simplistic, demonstration of the real purposes and effectiveness of that system. Such demonstration is, in effect, an evaluation based on empirical evidence.

In all such aspects, the consequent demoralization of the youth, a recurring phenomenon surely, takes on greater force since many of the members of the educational establishment are themselves coming to reject their assigned roles and to join in the clamour for change.

Levels of evaluation

Thus, the problems of educational evaluation may be expressed on several planes. On the one hand, educational evaluation within an established system has tended to focus upon the achievement of cognitive skills by individuals. This has been a clearly-defined process designed to satisfy some clearly-defined goals as agreed upon by the educational community and its clientele.

In this regard the efficiency of a system may be determined by its lessening number of drop-outs and repeaters, by its continually expanding avenues of access to those desiring its services, and by the best return, in the above terms, to the investments which have been made.

On a completely different level is the determination that educational systems products (students) are lacking in certain characteristics, or curricular inputs, which they may be finding in the 'informal' system. If the formal system is engaging in certain practices which represent costs, personnel and time, such as teaching about 'culture', the 'arts', 'history', offering 'personal guidance', 'health education', and so on, these too must be subject to evaluation. Since the effort is being made, the effectiveness like that of reading or spelling, or knowing the dates of famous battles, must also be put to the test.

On still another plane is the question of whether the educational system can be evaluated in terms of its ability to remedy or compensate for the debilities of the social and economic systems. With the growing expansion of school systems, the entry of 'disadvantaged' urban children, poor rural children, 'culturally-deprived' children, and others lacking in educational 'readiness', has offered a challenge which educators world-wide have been required to face. Some have welcomed the challenge and sought to offer revised methods and compensatory curricula, 'head starts', earlier admissions, and the inclusion of indigenous (though non-professional) personnel into teaching situations.

Others have faced the same challenge and have clung to the concepts that true merit will come to the fore, that equality of opportunity once having been offered, the challenge is now to the individual and to his family to make the most of it. Under these conditions the rigours of examinations and of customary evaluative procedures have maintained and exacerbated the eliminatory processes, the early cut-offs, the dumping of natural resources (human), and the development of rigid class distinctions. In both kinds of responses to existing conditions, the problems of educational evaluation come to the fore.

The question does remain however as to the role of education in making amends for social or cultural 'slings and arrows'. If, as the answer is often made, what else is education for, then the tasks require that education be well-armed, well-stocked with both talent and money, and provided with the power that comes from working on equal levels with those that hold the strings of the purse and the reins of decision. Indeed, if the answer is affirmative to the question about education's role as a change agent within the social order, then surely the means to evaluate its effectiveness in such a task must be sought tirelessly and objectively.

Evaluation begins with the individual. The results of individual testing (not teacher reporting, which is considered as too unreliable statistically) are compiled so that groups of individuals are seen as representing large-scale tendencies. Their grouped scores become the bases by which systems are evaluated.

Thus, when in New York City (for example), in the 1960's, large groups of children showed low scores in reading, the system was declared to be ineffective in achieving its goals. This was without regard for the possible intervening elements in the social situation, in the state of children's health, or in the aspects of home life.

In the recent IEA studies of Husén and his associates[1], grouped scores, cross-sectional, were related (or relations were sought) to conditions of family life, of the educational environment, and of expressed attitudes and opinions.

The difficulties are expressed succinctly by Porath who writes:

> "The final educational output is not produced only in the schools but also at home and elsewhere, before, during and after the period of formal schooling ... "[2]

Since the evaluation of informal system effectiveness is still in a marginal and relatively undefinable condition, it would be more useful to focus upon the formal educational structure. Further, the investment of costs, personnel and time in this system are explicitly dedicated to the purposes of education.

Evaluative targets and decision-making

The bulk of evaluation of system effectiveness has been in the achievement of individuals on set tests, with the data accumulated and grouped. Let us not, however, overlook those evaluation practices which take place in the home, in the market place, and often in the voting centres, when the ordinary citizen is enabled to express his evaluation of education by criteria of his own choice.

In such circumstances, as worker, parent, employer, or general observer, the criteria of the common man regarding the success of his educational establishment are likely to be in close relation to his own goals and his own experiences.

---

1 T. Husén, (editor), International Study of Achievement in Mathematics, Volume II, New York: John Wiley and Sons, 1967.

2 Yoram Ben-Porath, "Aggregate Costs, Output and School Achievement", in Donald E. Super (editor), Toward a Cross-National Model of Educational Achievement in a National Economy, New York: Teachers College, Columbia University, 1967, p. VII-1.

It is not facetious to note that politicians have, per se, no built-in or objective criteria for determining their attitudes toward educational costs, changes, and development other than those they have gained as perceptive laymen. Initially, therefore, social and economic decision-makers look at school systems through the eyes of the man in the street.

Thus the view is taken that open access to education is desirable and should be available to all. It therefore follows that once in the system, the processes of education to which a student is exposed should make of him something which he formerly was not. That is, he should be cultured, literate, skilled, communicative, loyal to traditions and accepted norms, and, as important as any of these, employable and paid on a living standard not available to one with less education or none.

If this process of evaluating education is an empirical one, it is surely practiced by political decision-makers with a keen sense of constituents' demands, by parents with a real knowledge of opportunity costs expended (actual or with regard to young children nearing school age), and with most students who are mature enough to seek their own advantage actively.

The difficulty with this form of evaluation is that it is not only practiced in countries which are developed and which are developing, but it is almost universally practiced within each country. And it is beginning to be linked, beyond the concepts of curricular and methodological relevancy, to concepts of accountability. People are beginning to apply their empirical concepts of educational system effectiveness to what they are experiencing, to what their school-age children are experiencing, and to what the post-school population, the outputs (drop-outs and put-outs), are experiencing. With the focus on accountability, when all is not well, the clamour arises for someone or something to be looked at and studied, and for relationships between the educational inputs and the outputs to be clearly disclosed.

Criteria for study of evaluation

But accountability, while a reasonable and desirable concept to foster, cannot be properly determined nor - and this is even more important in our view than accountability alone - can appropriate reforms and corrections be applied unless certain more objective steps are taken. Evaluation must be based on specific data, quantified information, and objective procedures. Biases, hunches, sudden enlightenment must be used sparingly and, when used, carefully labelled in educational evaluation. We prefer, rather, to point out that our approach to evaluation is based on goals, processes and results.

We must, therefore ask three questions:

(1) what are the goals of the educational system which have been expressed, accepted as appropriate and feasible, and, transmuted into educational curricula?

(2) what are the curricula which are combining content and methodology and which will provide specific experiences for students so as to achieve the goals of the system?

(3) to what extent is the measurable output of the system in conformity with the purposes for which the system has been operating?

Subsumed within each question above are several other questions, each of which is necessary to consider in the further development of educational evaluation.

(1) Thus, with regard to goals:

(a) how are school system goals determined?

The formulation of goals is proposed by Goodlad in a very specific manner: "We need a national body of leading citizens whose prime purpose is to give continued attention to the formulation of educational aims ..."[1]

---

1 John Goodlad, School Curriculum Reform in the United States, New York: The Fund for the Advancement of Education, 1964, p. 81.

With somewhat different emphasis, Smith states that: "the problem of objectives for American schooling is a problem calling for highly competent professional resolution ... The determination of objectives that will give curricular and other operational meaning to the central purpose in the years ahead calls for sophisticated theoretical, technical and administrative decisions."[1]

(b) to what extent do goals reflect realistic national needs?

(c) how are goals translated into operational school processes?

(d) what can be done to make goals explicit and amenable to both curricular experiences and evaluation?

The relationship of objectives to evaluation is stressed by Goodlad, when he says: "Needed are evaluative criteria in the form of educational objectives which have been agreed on ... It is recommended that curriculum investigators ... be required to submit statements of such objectives as well as plans for evaluating their attainment. These must not be simply statements of purposes for a given project but, rather, statements of the kind of behaviour sought in the students."[2]

---

1 Philip G. Smith, "Objectives for American Education", in Stanley Elam and Gordon Swanson, (editors), Educational Planning in the United States, Itasca, Illinois: Peacock Publishers, 1969, p. 9.

2 Goodlad, op.cit., p. 82.

(2) With regard to curricula:

(a) which disciplines in the primary and secondary levels are directly responsive to prescribed system goals?

(b) how can the cognitive-oriented goals be translated into curriculum sequence so as to provide the best methodological practices to varieties of students?

(c) how can goals which are non-cognitive be achieved in curricular experiences which are basically required to be explicit including: learning sequences, prepared syllabi, teaching aids, and trained teachers? Edding refers to this question:

"Other unsolved problems lie in the difficulties of measuring achievements other than knowledge and vocational skills. Yields of education rarely considered in examinations are, for instance, knowledge of how to learn and eagerness to continue learning; ability to co-operate, to take initiative, to make decisions under conditions of uncertainty; and virtues like tolerance, honesty, self-control, and creativity. These are possibly the most important results of education, but they are harder to measure than the performances usually judged and marked in certificates."[1]

(d) how can reasonable goals which are operating in school curricula become feasible in countries where resources are scarce, teachers unavailable or untrained, and the wastage rate too high to expose a school-age population long enough?

(e) to what extent can curricula deal with the large variety of system goals, often including skills,

---

1 Edding, op.cit., p. 24.

knowledge, attitudes, citizenship, personal adjustment, creativity, and self-realizations, in view of the limited time and resources of the system?

(3) With regard to evaluation of outputs:

(a) how can system goals be used as criteria for evaluation of the effectiveness of system processes?

(b) what are the most appropriate points at which to measure the accomplishment of students within a system?

(c) what instruments measure student achievement, with validity and reliability, as a function of their experience within the school system?

(d) what characteristics of students are:

(i) measured objectively and relate to system effectiveness?

(ii) not measured at all but should be in view of system objectives and curricula?

(iii) not measured and need not be since they do not have anything to do with the purposes of the educational system?

(e) how can various evaluation procedures of students during their school enrolment, at terminal points after each level, and at long-term, be considered as clearly indicative of system effectiveness?

(f) how can estimates of system effectiveness be utilized for reform of curriculum, administrative practices, and system goals?

The foregoing considerations have been developed out of a study of the problem of evaluation as it is treated both in existing practice and current research. As the needs of school system clientele, critics, and technical 'assisters' become more clearly interlocked, the need to approach evaluation of school systems becomes more urgent.

It is not sufficient for empirically-drawn conclusions to govern the action of educational planners and decision-makers. Nor can cross-sectional 'cuts' of several systems provide us with more than 'still-life' portraits of the systems. Data obtained from such practices are of interest to us when they relate scores of students to internal system elements which may or may not represent cost inputs or curricular processes. This does not, however, yield evaluation of systems since there are no criteria for evaluations, no indications of pupils' initial status and subsequent gains as a result of system processes, and no intra-system comparisons of reliably-maintained controls.

Goal-setting and goal characteristics

Evaluation is needed to determine how well system goals have been achieved. This has been stated before and is the fundamental basis for research in evaluation. We must, of necessity call attention to the problem of goals. In the continuing search for criteria of educational effectiveness, it is reasonable to ask about the nature of the goals which are found either explicitly in educational systems' declarations of purpose, or implicitly in the nature of the evaluations which are undertaken.

In most countries, educational goals tend to be set and are characterized as follows:

(a) they represent the least common denomination of educational purposes, i.e., purposes about which almost no one would quarrel.

An illustrative statement by the then Swedish Minister of Education (4 December 1968), Mr. Palme describes his country's basic goals:

"The goals cannot be different for different parts of our education ... The individual is given opportunities to fulfil himself. Education is an absolute necessity for a living democracy. The goal of education should be to

create increased equality. The school should break down social distinctions. The school is one of the most important means of changing society."[1]

A similar general statement is made by Smith regarding educational goals in the United States:

"There are three sets of factors that constitute the basis for control over experience ... organized human knowledge ... , individual qualities of effective intelligence and personality, ... and the cultural values and customs of our society. The central purpose of American schooling should be to teach in each of these areas, the skills, habits, understandings, and attitudes that are most educative, that is, those which most enhance control and thus make possible effective choice and hence individual and social freedom."[2]

(b) they tend to be quite explicit in relation to the cognitive domain: the communications skills (reading, writing, speaking, spelling, grammar), the mathematical skills, and all others which deal with data, repetitive-observable processes, and mass-developed behaviour patterns.[3]

---

1 Quoted in: Bengt Jacobsson, School Reforms in Sweden (duplicated), International Working Party on Educational Technology and the Learning Process, Geneva: May 1970.

2 Smith, op.cit., p. 4.

3 A major contribution has been the classification of educational goals (with illustrative test items for evaluating their achievement) by Benjamin S. Bloom (editor), in Taxonomy of Educational Objectives, Handbook I: Cognitive Domain; New York: David McKay Co., 1956. The major classifications offered are: "Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation", pp. 62-200.

(c) they tend to be those which are either most measurable objectively, or most readily observed at any selected time of testing.

(d) they are usually quite general and not clearly-defined when they refer to non-cognitive development: attitudes toward home and school, loyalty toward community and country, honesty, strength of character, self-motivation toward learning, respect for scientific concepts, and ability to play a role sufficiently independent for personal happiness and sufficiently group-oriented for social success.[1] Such goals are frequently to be found in ministry documents or in the rhetoric of parents, politicians, teachers, and, often, the students themselves. A Unesco document which shows similar characteristics calls for such main goals, among others, as 'values of the society' which are interpreted as:
"value attached to learning, interaction with others, adult/child relations, sex and role differentiation, international understanding".
It further urges:
"personality development, participation in a variety of activities, health and hygiene."[2]

---

1 The following classification of non-cognitive skills is given in David Krathwohl, Benjamin Bloom, Bertram Masia, Taxonomy of Educational Objectives, Handbook II: Affective Domain, New York: David McKay Co., 1964: "Receiving (Attending), Responding, Valuing, Organization, Characterization by a value or value complex", pp. 95-175.

2 Unesco, Final Report, Meeting of Experts on Curriculum of General Education, Moscow, 16-23 January 1968, Paris, pp. 7-9.

(e) such goals as have been described in (d) above, are usually not readily amenable for use as evaluative criteria. testing devices for system effectiveness (formal or informal) are quite unreliable,

(f) a number of system goals refer to the conditions of educational opportunity rather than to the gains made by individual students. Thus, goals are expressed in terms of 'full and universal' enrolments, 'open access' to further education beyond the primary level. free education to all students to the extent that each individual can benefit from his participation;

(g) in a number of school systems where planning for educational development is operative to any extent, the goals of the system are often described in terms of manpower to be produced. Thus, we may find explicit numbers of specified positions for which trained persons will be produced, or we may note the more general intention that more 'X' will be trained for the needs of the economy without exact quantities being indicated.

Goals become operational as curricula

The characteristics which are described above tend to be operative as the result of general agreement that they will be given adequate funding with implementation as curriculum processes. It has been, therefore, the task of educators to translate into functional programmes the determined goals of the system. And, as the curricula become sequences of learning experiences, they are systematically recreated as courses, syllabi, term and quarterly 'units', and eventually, as 'lessons', 'lectures', 'practical exercises', etc., in daily class activities.

It is in this basic form that goals become most tangible. Their definition in explicit (or implicit) form becomes feasible, and the relevance of goals to conditions of need and reality may be perceived.

The basic, concretized 'courses' within the curriculum enable us to note whether there may be goal irrelevancies in the forms of 'dead' languages, extinct 'sciences', and other uses of educational investment of little consequence in the movement toward national development. Certainly, such determination may be sought, and found with more or less difficulty. An important aspect of evaluation ought to concern itself with the gaps between goals and relevancy, as well as between goals and goal-achievement. There is, however, no reason to exclude from definitions of relevancy those cultural and artistic experiences which contribute to personal growth as well as to national development.

It is within the curriculum also that teachers and school system clients begin to evaluate empirically the achievement at near-term of system goals. Teachers practice informal testing and individualized observation. At the higher levels examinations and practical exercises are used to determine if goals are being achieved. Pupils, parents, school officials, and employers of school leavers also make various evaluative attempts, indicating their satisfactions or dissatisfactions as has been noted earlier.

Yet, with the start of evaluation processes, no matter how crude, within the system itself, we note that goals are in need of continued study. The following circumstances are generally prevalent:

(a) some system goals which are stated explicitly are measured more or less objectively.

Bloom states the case directly:

"If education is to be open, public, and examinable, the specifications for it must explicit, and either the process of education or the outcomes of the process must be examinable in relation to such specifications."[1]

---

1 Benjamin S. Bloom, "Some Theoretical Issues Relating to Educational Evaluation", in Ralph W. Tyler (editor), Educational Evaluation: New Roles, New Means. 68th Yearbook of the National Society for the Study of Education, Chicago: University of Chicago Press, 1969, p. 29.

We would amend this by suggesting that both process and outcomes must be 'examinable',

(b) some stated goals of the system are not measured objectively, but by teacher observation without standardized criteria for judgment,

(c) the achievement of some expressed goals is not measured at all, but is found to be importantly operative in the curriculum. We note as an example, pupils' involvement in the 'cultural' areas such as music, the arts, physical education, etc. These activities usually are not measured as to system effectiveness, although pupils with special aptitudes in these areas are noted and encouraged;

(d) the achievement of system goals is often measured at specifically determined points and becomes a factor in the elimination of pupils from the system;

(e) system goals are often explicitly stated, but in reality are negated from the start by: insufficient primary or secondary education facilities in rural areas, poorly trained (and not enough) teachers in most of the primary schools; failure to enforce such requirements as compulsory attendance and secondary pupil vocational guidance; operation of non-functional, non-sustaining literacy programmes for all citizens needing them.

"There has always been, and no doubt will always be, a great gap between the rhetoric and the deed ... This does not say that the announced aims were a contrived deception. Rather it is a sign of how long it takes to alter the course of an ongoing educational system, to redeploy its energies in new directions and to marshal new energies and resources to take on new tasks and to do old ones better."[1]

---

1 Philip H. Coombs, The World Educational Crisis: A Systems Analysis, New York: Oxford University Press, 1968, p. 100.

Simply stated, under certain circumstances system goals are enumerated but not acted upon, there is, therefore, no educational output (or achievement) to reflect such goals.

The relevance of curricula

Goals cannot be realistic nor can system output in terms of student achievement reflect the accomplishment of goals unless the curriculum makes such conditions possible. We consider that curricular processes, the nature of curricular materials and their use, the depth and breadth of student experiences, and the procedures of teaching all have their role in achieving desired educational outcomes.

Without going too deeply into detail (since this is not a curricular study), we point to a few examples related to curricula and goals which may determine effectiveness.

(a) When the goal is generalizable learning of cognitive skills, such learning is not enhanced by rote, memorization, or repetitive exercises of non-meaningful operations.

(b) When the goal is to awaken love for literature, traditional culture, and the arts, such accomplishment is not encouraged by processes in which expression is frustrated, where emotional involvement is suppressed, where the arts are presented without imagination, and where cultural, literary and artistic participation are absent or are undertaken in a primitive or coercive context.

(c) When the goal is better understanding of natural laws, scientific concepts, and ecological balance, achievement is hindered by strict adherence to textbooks, lack of student involvement with their own immediate environment, and lack of a generalized application to school life of the concepts of logic, reason, and the scientific method.

(d) When the goals are active citizenship, pupil responsibility, and the development of ethical behaviour, these goals tend

> to be frustrated when it is not recognized that citizenship experiences begin within the government of classes and schools, that communities in which students and their families live are dynamic textbooks for the study of how people live together for their mutual benefit, and that individuals are the units that make up all systems of government. We cannot expect that coercive, dictatorial classrooms will produce citizens who are competent to participate in developing democracies.

While the above discussion is focused upon the curriculum and educational methodologies, the nature of school systems output is determined by such processes. A major problem is to ascertain the quality (and quantity) of such goal achievement as has been cited here.

When we review the two aspects of goals, the cognitive and the non-cognitive, we can discuss the means used to evaluate student achievement in each of the areas.

Evaluation of goal achievement: cognitive skills

First, as to cognitive goals, the general approach has been through testing procedures which begin with the periodic observation by teachers of the nature of each pupil's ability to satisfy certain individually-determined requirements. These include: reporting back to the teacher what the teacher stated at some earlier time; reporting to the teacher whas was found in school textbooks; carrying out assignments (homework) as related to the textbook or the presentation by the teacher.

In determining pupil accomplishment of cognitive skills teacher-developed, or school-developed, tests are usually a first phase, taking into account the specific experiences of pupils in the particular classes.

Much of such testing deals with short-term experience, basically covering several days or weeks of schooling. In many ways such testing provides school officials (including teachers) with current data about

their pupils, about teachers' relative effectiveness, and about the educational needs which exist in the localized situation. As may be determined by perceptive educators, methodologies used and curriculum processes may receive careful study, remedial courses may be instituted, and better use may be made of existing materials. Thelen describes the latter use of evaluation in referring to 'diagnosis and trouble-shooting' activities of evaluators and teachers:

> "Trouble-shooting is built-in feedback that enables the teacher and the group to see how they need to modify their activity ... "[1]

On a somewhat higher level, eliminating the dangers of bias for or against a given school, class or pupil, are the examinations prepared by external examiners on the assumption that there is a standard body of knowledge to be taught and learned, regardless of location of school or teacher qualification.

This testing may also yield feedback information which will then encourage educational change. It may, however, force teachers into establishing ever-more uniform curricula and teaching methods. We see the potential for a paradoxical situation, therefore: for schools of poor quality and inadequate resources, such examinations may point the way to up-grading of teachers and curricula. For schools of relatively competent teachers and psychologically-sound (if unorthodox) methods, such testing may serve to enforce conformity to outmoded or undesirable procedures.

To review further the procedures for evaluating achievement of cognitive goals, we find that testing encompasses not just individual schools, but entire systems. There are generally, again, two aspects of such testing. First, is the 'summative' which is prepared year

---

1 Herbert A. Thelen, "The Evaluation of Group Instruction", in Ralph Tyler (editor) op.cit., p. 148. Also pp. 124-127 and pp. 151-155.

after year by specialized educators to be used at the completion of specified blocks of school time. This may take the form of examinations like the English 'Eleven-plus' and General Certificate of Education, the French Baccalaureat, the New York State Regents examination, or other certificate-earning, school-leaving tests.

While such testing procedures are often seen as gauging the effectiveness of school systems, they actually do no more than assess pupil status at a given point in time. There is considerable evidence, in fact, that such testing cannot reliably evaluate system effectiveness since pupil achievement is determined by their experiences in school, their experiences out of school, as well as by numerous social factors, economic aspects, and motivational status, all generally unrelated to school functions. The work of Coleman, Husén, and others tends to bring these aspects into prominence.

The second aspect of systems-wide testing is that commonly referred to as 'standardized'. Simply stated, through successive application of tests (and sub-tests) it is determined that students of average competence in a specific skill will receive an average score on the resultant test. Thus, a standardized test measures achievement as it might be demonstrated by normal population of students, assuming a common (not necessarily uniform) curriculum and competent teachers. Standardized tests have been used throughout the United States, and with modifications in other western hemisphere countries. The results of such 'formative' tests have been used both to determine the achievement of individual students (and their remedial needs) and the presumed effectiveness of school systems.

Such 'effectiveness', while evident in terms of pupil status, cannot be realistically used in terms of the system itself when the existence of intervening factors which inhibit or restrain pupil progress is acknowledged.

In the various types of testing procedures described above, we note that:

(a) they are basically related to individual pupil status, though as aggregates they are used to define system effectiveness;

(b) they are indicative of specific curriculum content, generally rote-learned, and non-generalizable;

(c) they are often used as guidelines to support policies of pupil repetition or elimination from the school system;

(d) they are usually quantifiable so that scores may be assigned, though the closer to the teacher and the school the testing is, the less objective may be the quantification;

(e) they do not take into account the influences upon pupil achievement of the 'informal' educational system which makes it difficult if not impossible to evaluate with such procedures the effectiveness of the formal system.

Testing for incremental gains: the value-added concept

It seems to us that the common approaches to measurement of student achievement of cognitive skills have two basic values: first, they indicate the status of groups of students at determined points, and secondly, they make it possible to attempt some form of assistance to individual students. We do not believe, however, that such terminal or cross-sectional measures can reasonably be applied to the effectiveness of school systems, or even of sub-units within the systems. There are too many intervening variables.

However, we find numerous instances, particularly in the United States, of the adaptation and use of achievement measures on a pre-test, post-test basis. Appropriate tests are administered to a given student population at an early point in the programme to establish a base-line, or first level of status indication. Following a period of instruction, a comparable test is administered and the differences of achievement are noted. Flanagan describes the use of such procedures

for testing student progress with newly-developed instructional methods and materials.[1]

It appears probable that the gains (or losses) in achievement as demonstrated by aggregates of students would be reliable indicators of the effectiveness of their instruction. Bloom has pointed out:

> " ... unless the criteria of effectiveness are related to changes (our italics) in students, the researcher has avoided the primary criterion ... ".[2]

This is not to say that social and economic variables may not suppress or enhance student achievement apart from the effects of schooling. Such variables can, however, be held constant and allowed for with large numbers of students in common circumstances, their exposure to common instructional and administrative processes, and the pre-determined short-term between pre-testing and post-testing.

With growing demand for school accountability, the calculation of educational value-added by such procedures can give indications of the results of schooling. Such calculations can be more readily analysed within known contexts of pupils' socio-economic levels or prior school experience. Given the establishment of a base-line achievement level and the application of education processes, the forward movement by individuals and groups (or lack of movement) may be seen as a function of the school and the system. Thus, with certain cognitive

---

1 John Flanagan, "The Uses of Educational Evaluation in the Development of Programs, Courses, Instructional Materials, and Equipment, Instructional and Learning Procedures, and Administrative Arrangements", in Ralph W. Tyler (editor), op.cit., pp. 225-228.

2 Benjamin Bloom, "Testing Cognitive Ability and Achievement", in N.L. Gage, (editor); Handbook of Research on Teaching, Rand, McNally and Co., Chicago, 1963, p. 379.

skills, it is feasible to attribute pupil progress to system effectiveness. It would be difficult to ascribe effectiveness with similar confidence to the results of such end-of-sequence testing as 'Eleven-plus', Baccalaureat, or other procedures which do not permit a value-added determination.

Evaluation of goal achievement: non-cognitive skills

As we have noted earlier, it has generally been in the area of specific, cognitive skills (knowledge, problem-solving, higher mental processes) that efforts in evaluation are quantifiable and objective. From such testing the concepts of student success and of system effectiveness have been developed.

Mention has been made above of the areas of educational goals which are non-cognitive, i.e., referring to attitudes, values, appreciations, interests, and personality characteristics. These, too, are expressed as goals for educational effort.

Although there is even greater controversy over the formation and development of such behaviour components in school environments, school personnel have generally tended to accept this task within their roles. Psychologists, sociologists and others interested in the development of human attributes have generally indicated that these traits are socially formed, starting with the home and continuing within the culture of the family and community. Further, they contend that even within the school, there exist peer-group cultures and teacher orientations which are non-curricular and which affect the development and expressions of non-cognitive patterns of behaviour.

This question is considered from another standpoint by Krathwohl, Bloom and Masia:

> "While the psychologist and philosopher may have views on what is desirable and even necessary in the affective domain, there is still the question of what affective objectives society will permit and even encourage. Our own society has fluctuated as to the affective objectives it will permit the school to

develop. Political and social forces are constantly at work, pressing the schools for some affective objectives and just as constantly placing restrictions on the school with regard to others."[1]

Yet, since the school system is orientated toward the development of both intellectual and non-cognitive skills, several problems of process and evaluation should be noted.

In the acceptance of educational goals such as improved attitudes, loyalty, character, appreciations, etc., there have been few instances of curricula which are specifically designed to 'teach' such skills. In some countries, courses in civics, national history, and national culture attempt to deal with or inculcate the desired behaviours and practices through specific activities. In general, however, the development of such skills is dealt with in the context of the existing disciplines (the cognitive areas) and in terms of the expressed behaviour of the students.

The evaluation of student achievement in the non-cognitive areas is generally through the awareness and recognition of teachers and other competent observers. Attention is paid to the behaviours of students in their classes and in school-wide activities, as well as in such specific aspects as participation and creativity.

Teachers and other observers take into account students' demonstrated interests in learning, their motivation to study and go forward in the direction indicated by the teacher, and their persistence in applying themselves to required procedures. Teachers may designate such behaviour as 'good character', 'favourable attitude toward school', or 'self-actualization' within a specific environment. These observations are clearly subjective, generally expressed in terms of teacher expectations and needs, and without reference to school processes. It does seem feasible, however, to assume that students who function

---

1 Krathwol, Bloom and Masia, op.cit. p. 90.

in these ways within an educational system are highly compatible with the system as to expected behaviours and indications of success. They and the system may be seen as 'effective'.

Since such evaluation is difficult to quantify and generalize, we can describe student status in this way, but further study must be done to ascertain the processes within the system which may have produced that status. Krathwohl, Bloom and Masia suggest that:

> "... if specific changes are to take place in the learners, ... the learning experiences must be of a two-way nature in which both students and teacher are involved in an interactive manner, rather than having one present something to be 'learned' by the other."[1]

We may need to look, with this regard also, for the effects of non-system variables before we can claim system success for student adaptability.

Student participation as an indication of non-cognitive development

Student participation may be seen as an indicator of system effectiveness under specific circumstances. The level of student responsiveness within classes, student feedback to teachers, and the openness of student participation in discussion all may indicate positive attitudes, character development, and so forth, provided that such student involvement and expression are given opportunity within a school system. Student achievement in the non-cognitive skills can be observed in 'open' systems where teacher attitudes are receptive and encouraging. In such circumstances procedures can be developed for quantifying the nature of student participation as has been described. Observational techniques coupled with rating scales, case studies, and sccio-metric methods may be applied to determine student status; and, inferentially, to establish school and system

---

1 Ibid., pp. 81-82.

status in this area. With the further inclusion of pre-test, post-test concepts, such procedures may be relatively effective in noting school-determined gains.

It is, of course, conceivable that student participation in class inter-action may be usefully carried forward to collaboration in school operation, and ultimately to system evaluation, modification and reform.

Student creativity as an indicator of non-cognitive development

Student creativity, while initially requiring the achievement of skills in manipulation of specific communications tools, may be seen, when evident, as an indicator of non-cognitive achievement. Thus, when we recognize that the basic approach to literary expression is in the mastering of the cognitive skills of reading, writing, spelling, we find that the opportunity for and encouragement of such expression may indicate system achievement. Similarly, the involvement and exposure of students to graphic and musical arts within curricular experiences are based upon system goals and require evaluation as to effectiveness. Such effectiveness may be determined by careful study of the conditions and nature of student expression in the arts.

We recognize that within given systems there may be no curriculum for 'literary' creativity, or for graphic or musical expression on a creative level. When students produce such works, therefore, we are almost required to search in the school processes for incentives, motivations, and positive attitudes in these areas. Ultimately, we may determine that this is such an individualized matter that the educational system, or any of its elements, plays no role in this development.

On the other hand, when we note full courses, specialized teachers, and available materials, we may look for evidence of both appreciation and specific creative expression in some form of output. As to appreciation, some common (if depressing) practices include formalized testing for recognition and memorization of works, artists,

themes, and so on. Further evaluation may take note of self-development in leisure time, selection of undirected reading choices, and post-school interests and activities.

Indicators for system evaluation

In the developed countries of the world, evaluation of school systems is carried out on two bases: first, on the qualifications of school leavers in terms of years completed, evident literacy, and certificates and degrees held; secondly, on the internal system evidence of achievement scores coupled with repetition rates and drop-outs.

Little use is made thus far of the less quantifiable forms of system evaluation when they are voiced by school clientele, including students. The concepts of accountability are still rather limited to personalized or empirical determination except in rare cases. These occur following discussions of 'out-moded' curricula, lack of concern for individuals, and excessive stress on competition, elimination, or discriminatory tracking (streaming).

An interesting concept of accountability has developed in the United States, as reported in a recent issue of the Phi Delta Kappan.[1] Private companies are given contracts to operate educational programmes within schools. Such 'performance contracts' provide a sliding scale of payments in accordance with the rise in pupil achievement levels. The programmes are supported with national funds to the school districts which then arrange for the contracts. Companies are responsible for providing teaching personnel, teaching materials, and for determining the instructional methods. Thus accountability becomes a matter of profit or loss to the contractors. They are indeed accountable, since non-achievement by pupils to contracted-for levels may result in no payment. Other accountability projects will range from trading stamps (exchangable for gifts) as regards to students to cash bonuses for teachers.

---

[1] Vol. 52, No. 1, September 1970, p. 63.

A Canadian experience with accountability dates back to the late 19th century when financial grants to high schools in Ontario were made dependent on the number of pupils who passed an examination after a year or two of attendance. Phillips reports that:

> "under full pressure of payment by results, teachers and pupils began rapidly to measure up to the requirements ... One reason for this amazing improvement in efficiency was that trustees were aroused from lethargy. When their school failed to earn grants, their response was to blame the senior master and to dismiss him ...
>
> "No doubt some of the teachers dismissed were lazy or inefficient. But nearly all were shrewd enough to take advantage of every new means that was offered to get results. Teachers' professional journals were filled with sample examination papers, model answers, and advertisements of little books containing notes on various subjects, the memorization of which would ensure success on the examinations. History teaching became the application of a system of mnemonics and the teaching of literature little more.
>
> "Experience with payment by results in Ontario proved that it is possible to raise standards quickly if the criterion is defined as mastery of prescribed content. But there was a storm of protest against the sacrifice of all other educational values for the attainment of this end. In 1883 payment by results was abandoned in the province."[1]

In developed countries there are serious moves toward the rethinking of system goals. This is evident, as well, in the United States, Germany, and the United Kingdom where the varied centralization would not seem to encourage national, or system-wide agreement on goals.

---

1 Charles E. Phillips, The Development of Education in Canada, Toronto: W.J. Gage and Co., 1957, pp. 513-514.

The discussions have generally led to consideration of curricular processes and, equally important, to attempts at determining indicators of the effectiveness of educational systems.

In developed countries such indicators have been designated as: (a) relating to system accessibility; (b) relating to system processes; (c) relating to the educational product.

(a) As regards system accessibility, we would consider as valid the following indicators:

(i) a truly open and available primary system at no cost and with appropriate subventions to those families for whom opportunity costs were to be determined as excessive;

(ii) a clear-cut policy on promotion so that failure to achieve according to determined standards would not result in repetition, but would bring about compensatory (and remedial) services and modification of curriculum;

(iii) the development of educational and social procedures which would reduce the disparity at the start between poor children and middle class or others. Such procedures could provide earlier educational experience, increase family literacy and motivation, relieve mothers of excessive child-care burdens by providing ancillary nursery or other services;

(iv) the establishment of close relationships between primary education and the next level through information, guidance and orientation of pupils and their families. This would enable families to have specific understandings of the possibilities for their children in continuing their education. It would enable personnel in the primary schools to convey information to their students and to encourage appropriate further development. It would enable secondary school personnel to know and understand the conditions of primary education, to appreciate the status of

incoming pupils, and to restructure their own curricular and guidance procedures,

(v) the determination of a socially appropriate and economically feasible compulsory school age so that students who wish to do so may leave the school. This could only be established together with provisions to counsel students before they drop out, for incentives and reasonable conditions for remaining in school, to serve them with employment advice and liaison, and provisions for further education when desired.

(b) As regards system processes, indicators of effectiveness would include:

(i) content in curricula, courses and programmes which is developed in terms of agreed-upon goals reflecting national and individual needs. This indicator includes relevance for students and their families, acquisition of skills which are in keeping with concepts of self-development as well as with the manpower market, incorporation of materials which are reflective of national as well as community interests as regards tradition, culture, and options for future development. With curriculum content as an indicator of effectiveness, we can extend our study to relationships between system organizational echelons and the actual implementation of system goals;

(ii) teaching practices in the curriculum which are oriented toward student status and conditions of learning. It is unrealistic to consider curriculum content as an indicator of effectiveness without referring to methodology as well. The translation of educational goals into content can only be made concrete through student experience.

Methodology involves teaching practices, uses of materials, as well as general aspects of relationships between the teacher and the taught.

We would stress that the development of teaching practices to convey curricular content would need to make full use of relevant psychological factors. The acquired body of knowledge we now possess about how learning takes place must serve as a foundation for methodological development. Thus the improved procedures for teaching and learning may serve as indicators of system effectiveness when related to goals, curriculum content and to measured output.

(c) As concerns the evaluation of the educational product as indicators of system effectiveness, we would look to the following:

(i) establishment of procedures to reduce to a minimum the number of school leavers prior to the basic certificate stage as may be determined by national policy;

(ii) the utilization of interventions for measurement of educational value added within the school system at numerous points. The 'product' therefore would be measurable in terms of gains over initial, base-line status;

(iii) the development of quality criteria for the evaluation of manpower output. The production of engineers, machinists, nurses, or personnel of any profession is now seen in numerical terms, i.e., how many were produced. An important indicator of effectiveness is the quality of the production, not only based upon courses taken, years spent, theses written, or practical activities pursued, but the varying levels of efficiency of the certificate or degree holders at the end of their preparation. As we have noted earlier, the production of qualified personnel involves a minimal passing group with numerous

individuals on a quality continuum from 'passing' to 'outstanding'. The evaluation of such differences would prove a meaningful and useful indicator of system effectiveness, allowing for returns to the curricula for modification or reform;

(iv) the organization and operation of research and development units which would undertake needed studies, provide recommendations, and participate in implementation. A typical task of an R and D unit would be the preparation of long and short term follow-up studies of school leavers, including graduates. Such studies should attempt to determine the effects of schooling upon personal mobility and professional effectiveness.

Naturally, we assume that the longer the term of study, the more impressive would be the effects of non-school variables, including chance, war, revolution, and economic movements. Yet, with such studies, and the development of controls, we may determine the critical factors of school experiences which have determining influence in spite of outside variables.

Summary

In the foregoing discussion, we have attempted to establish directions which might be taken to evaluate school system effectiveness. We have not mentioned the elements of cost benefits or cost effectiveness since such considerations are limited to economic and financial aspects. We are interested in the possibilities of establishing criteria for evaluating system effectiveness which may, when coupled with such economic studies, be useful to educational planners throughout the system. We include with them those who help establish goals as well as those who plan and organize administration of schools, school curricula, teacher education programmes, and learning procedures.

We have stressed in this discussion the potential for the evaluation of school effectiveness indicators in developed countries primarily because of existing organizational strengths and available financial resources. In our view the developed countries are generally nearer to the points of universal education and qualified teaching staff so as to begin to implement the concepts expressed. Finally, we note that the research and development capabilities in developed countries make possible feasible involvement in such activities. They possess personnel and research facilities which can be engaged in the types of studies, programme development, and heuristic processes which are required.

For developing countries, we suggest that such considerations may also be valid and realistic. Given the existing conditions of insufficient resources, low enrolment rates, high wastage rates, poorly qualified personnel, and inappropriate curricula and methodologies, it is desirable that study be usefully applied in specific areas. Technical assistance teams for research and development can offer studies to establish relevant priorities. Questions of goal-setting should be resolved so that ensuing development of curricula and personnel training can be undertaken within the recognized constraints.

When efforts were made by developed countries, too often such practices and techniques were lifted almost in toto and set up for operation in developing countries. This is now almost universally recognized as ineffectual, if not counter-productive. We look instead to the establishment of improved procedures in educational systems which will serve as 're-design' models for developing countries. Such models would be oriented toward the generalizable conditions of developing countries and would enable the reconstruction of system elements if not the creation of new ones. Thus, we might find reconstruction of planning departments, of teacher training sub-systems, of primary education curriculum and methods, etc., all

amenable to modification as parts, yet within a total system configuration. The kinds of changes we have conceived of in developed countries could serve as an aid to developing ones so that repetition of errors or malfunctions could be avoided and some 'giant-steps' could be taken.

PART II

CURRENT PROCEDURES FOR EDUCATIONAL EVALUATION

Introduction

In the preceding discussion we presented a synthesis of the activities and problems concerning the evaluation of educational system outputs. We described the nature of current considerations and the developing pressures which are affecting the thoughts of researchers in the field.

It is important to note that both economists and educators are seeking to determine the differences between the evaluation of students' achievement and the evaluation of systems as they enable students to achieve. We note also that the bulk of past research in educational measurement has been primarily focused on the achievement of students in cognitive areas. As economists, and increasingly, the general public, have demanded systems evaluation, educators, politicians, and systems managers are beginning to seek a more common understanding and basic definition of the problem.

Existing conditions of practice in the evaluation of school system outputs

In this section of our study we propose to review the evidence of existing practice in the evaluation of educational outputs. We shall describe the following categories:

- the purposes of testing in educational systems;
- the types of evaluation instruments used;
- the frequency of testing and uses made of the data obtained,
- implications of the foregoing.

The purposes of testing in educational systems

The following conclusions may be drawn as to the purposes (explicit and implicit) of achievement testing:

(a) to relate pupil learning to teachers' expectations,

(b) to determine the ability of students to report and apply information which has been learned and to demonstrate learned attitudes and desired behaviours;

(c) to establish evidence for the elimination of students from the school system, or for the repetition of years, or for assignment to special tracks or schools;

(d) to support the admission (or refusal of admission) of students to further levels of education;

(e) to serve as a basis for the award of certificates, degrees, or licences;

(f) to provide students with an additional motivation as well as further learning opportunity by means of a test situation;

(g) to establish individual student status within a group, group status in a school, etc., as regards a prescribed curriculum. Such status may be reported as class position, test score, percentile of the tested group, or grade placement in a standardized norm.

The foregoing purposes for achievement testing of students are common to all countries, developing as well as developed. They should be read, however, as a descending order of frequency. Thus the first five items (a-e) are generally found much more frequently, whether expressed or implied, than the later purposes.

In addition, we find a set of purposes for achievement testing which are quite infrequent, being notably operative in some British and American schools:

(a) to determine whether the teaching (and the teaching staff) has been effective;

(b) to establish within a school-community relationship the concept of accountability, i.e., that the system is responsible for pupil learning and will see to it that evidence of non-learning is studied and used for change;

(c) to permit the modification of curriculum or moves toward new curriculum development in terms of evidence that poor test results have come about within an existing curriculum;

(d) to permit the re-thinking of teaching methodologies so that following poor test results, the methods which were used with low-achieving students might be improved;

(e) to establish a relationship between the results of achievement testing and the effectiveness of school systems.

As we noted earlier, the foregoing uses of achievement testing are less frequent than the earlier list (supra), and these, too, are in a descending order of frequency. In fact, the final item referring to system effectiveness is not currently found through research or systematic study. It is, rather, occasionally used as a critical concept: that the schools are doing a very poor job altogether, look at the poor achievement of pupils! In a number of countries this evaluation of local school systems finds expression in migration from rural to urban areas, and by the movement of middle-class families to suburban districts and private education.

The types of evaluation instruments used

All evaluation of pupil achievement begins with observation by teachers and parents, though on occasion employers and other interested persons offer evaluation as well.

We find, therefore, the following measures of evaluation most generally used:

(a) teacher observation: pupil achievement is reported as observed or in retrospect, summatively,

(b) teacher observation: pupil achievement is reported with specific forms, rating sheets, or other standardized materials,

(c) in-class tests of current achievement, provided as part of textbook or workbook materials,

(d) in-class tests, prepared by teachers periodically as determined by the content of the curriculum and the lessons (pupil experiences) which had been completed.

(e) external examinations, prepared by individuals or committees with special expertise to evaluate student learning, ability to express knowledge, and ability to conform to the exigencies of testing conditions.

(f) external examinations, standardized with relation to the requirements of the curriculum and to the quality of the student population of a common age, experience, and ability in the subject matter of the examination.

In the preceding list, we find no specific order of incidence of objective or subjective measures. The teacher who reports observations ad hoc is more subjective, hence less reliable, than the teachers who are required to report their observations of pupils on a form, a rating scale, a three-point choice, or other such 'guiding' device.

On the same basis, in-class tests of pupil achievement may be subjective when they are: (a) teacher-made; (b) responded to by pupils writing freely; and (c) when they are scored by the teacher who prepared them. We stress the relative objectivity (that is, there is no question as to what the student's answer is) of tests which call for decisions on one of three possible answers, a 'true/false' choice, a matching of related items, or a completion item in which a single word or phrase is to be inserted. Not withstanding the rigidity of such testing instruments and their indication of limited aspects of learning, they do permit non-biased evaluation and generalizations about groups of students. Similar comments may be made with regard to tests which are included in textbooks. In the latter case, however, the testing-and-learning potential is also evident since the book is available to the student for checking, review, and further study.

In general, there is an attitude prevalent among researchers which favours the adherence to objective procedures so that those aspects of a student which are not being tested will not intrude.

We refer to student personality, verbal fluency, behaviour under stress, relations with the teacher or with examiners, and so forth. On the other hand, the ability to satisfy an oral examination jury, or to present an extended essay or dissertation with no prior knowledge of the topic to be selected is seen by many as a most relevant evaluation of student competence and of system effectiveness as well. The problem is not easily resolved except as may be determined by conditions of explicit system goals, of adequate resources in personnel for teaching and for examining, and of system openness. Also to be taken into account are the effects of the 'informal' educational system and of the non-school variables we have mentioned earlier.

Frequency of testing and uses made of the data obtained

Traditionally, testing of student achievement has taken several very specific forms: (i) testing of students as they completed a school form so as to permit their advancement to the next; (ii) testing at the conclusion of a school period or cycle to indicate completion of requirements for a certificate and for entry to a continuing, higher cycle; (iii) testing at the beginning of a new educational period so as to confirm entry to higher levels of study.

Thus, pupils, parents, and teachers have known, in all countries, the official and regional examination times. As Atiyeh notes:

> "Quantitatively, in many countries, a good student during a career of 12 years normally sits for approximately 16 formal examinations, each covering a variety of topics."[1]

Attached to, and usually in conjunction with testing, has been the continuous assignment of marks by teachers to students, presumably on the basis of objective evidence of accomplishment and behaviour.

---

1 Naim N. Atiyeh, "Examinations: Trends and Prospects", in Joseph A. Lauwerys, and David G. Scanlon, Examinations. The World Yearbook of Education, London: Evans Brothers Limited, 1969, p. 376.

There is a developing tendency to reduce the frequency of required examinations, in fact, to eliminate them altogether. In Swedish schools[1], continuous assessment by teachers is now the norm with no compulsory written tests or final examinations in the nine-year comprehensive school. Standardized tests are available at the teacher's option and it is interesting to observe that the use of these tests has increased to the point where 90-95 per cent of all pupils in the comprehensive schools are taking tests. In the gymnasium, which follows the comprehensive school, tests are compulsory for stated forms and are spread out from December to May.

In the Federal Republic of Germany a similar condition is developing, with emphasis on teachers' marks, a minimum number of required internal examinations, and, up to the completion of secondary education, no compulsory tests. Research has shown that previous required tests for entry to secondary education were no more valid than the teachers' continuous assessment.[2] It is only at the completion of secondary education (gymnasium) that the Abitur examination is required. Standardized tests are relatively not used except for primary school admission as devices for prognosis.

The frequency and nature of testing in France is reported by Legrand, who notes the conditions which persist despite reforms:

> "... The reforming zeal, devoted first to the educational structure and now to the definition of the syllabus and teaching methods, has not yet touched the examination system. So it continues as before, rendered even more complicated by the reforms ..."[3]

---

1 Sven-Eric Henricson, "Continuous Evaluation in Swedish Schools", (duplicated), National Board of Education, Stockholm, 30.IX.1969, 8 p.

2 Reported by Karlheinz Ingenkamp in Lauwerys and Scanlon, op.cit. pp. 140-145.

3 Louis Legrand, in Lauwerys and Scanlon, op.cit., pp. 123-129.

Thus the French child goes from test to test - some optional - but mostly taken as if compulsory:

Examin d'entrée en 6ème: for entry to the first level of secondary education, at about ages 10-11;

Certificat d'études primaires: at completion of the primary period, at about age 14;

Brevet d'enseignement du premier cycle: at the end of the first level of secondary education, at about age 16;

Brevet d'enseignement professionnel: at the end of the secondary vocational (short) course, at about age 19;

Then at the end of the secondary vocational (long) course, at about age 20, the following choices:

Baccalaureat des techniciens;

Brevet industriel or commercial;

Certificat d'aptitude professionnelle;

Finally, at the end of the academic second secondary cycle, at age 19, the Baccalaureat.

In reviewing this situation, Capelle proposes that a single certificate be awarded following the completion of compulsory studies and that it be on the basis of observation and guidance. He sees the value of an examination at this point only as an arbitrating instrument where parents disagree with the proposals of a school orientation committee.

Capelle approves examinations for vocational qualifications since they constitute "attestations of capability recognized by contracts of employment ... ". Thus, a "conventional examination seems indispensible."[1]

To indicate completion of the 'long' general secondary course he suggests a generally standardized programme of teacher marks which

---

1 Jean Capelle, Tomorrow's Education. The French Experience, London: Pergamon Press, 1967, pp. 124-144.

provide a summary of the schooling completed. For entrance to university level studies, then, a national written examination would be compulsory, while university authorities could still select their students from within the group completing this test successfully. Thus the proposed improvement becomes, with a limited number of places, a highly competitive matter.

In England and Wales, the basic examination structure requires testing of primary school students for placement within secondary education and further testing at the end of secondary education with either the General Certificate of Education (G.C.E.) or the Certificate of Secondary Education (C.S.E.).[1] The G.C.E. is a pass/fail examination generally taken by candidates for higher and further education. This examination is prepared and administered by eight independent boards for students of grammar schools, i.e., those considered to be in the upper 20 per cent of the intelligence range.

The C.S.E. is prepared by fourteen regional examining boards for the next lower 40 per cent of the intelligence range. Whalley describes the three modes of C.S.E. examinations which seem to bring it closer to the needs and interests of regional conditions. He describes Mode 1 as an external examination on syllabuses produced by the regional board; Mode 2 is an examination based on a school's own syllabus; Mode 3 is an examination by individual teachers of their own candidates using their own syllabuses. All of course are with the supervision and approval of the regional boards on which are representatives of the national Schools Council for the Curriculum and Examination.[2]

---

1 Henry G. McIntosh and Daisy M. Penfold, in Lauwerys and Scanlon, op.cit., pp. 110-119.

2 G.E. Whalley, The Certificate of Secondary Education, Leeds: The University of Leeds, Institute of Education, March 1969, pp. 16-19, 41-47.

The primary examination, which is still required in some parts of England (the 11+), includes a test of intelligence, a test of proficiency in English, and a test of arithmetic. Results, which are combined with the cumulative school record, lead to decisions regarding selection and orientation of students toward various educational streams.

In the United States, teacher-made examinations as well as teachers' continuous assessment are found in all states in both primary and secondary levels. In addition, periodic standardized achievement tests are required at various points in both levels, with the results utilized for student guidance, parent information, and, occasionally, as indication of teaching effectiveness. There are no school-leaving examinations at the end of primary education, nor are there admissions examinations for the secondary schools, which are almost all comprehensive.

For admission to universities, examinations are used (prepared by the College Entrance Examination Board, American College Testing Program, National Merit Scholarship Corporation) in conjunction with secondary school records to determine admissions and scholarship awards. More recently, a number of universities have granted admission to students with certificates indicating completion of secondary studies and strong recommendations from secondary school officials or community representatives. This is especially true for members of minority ethnic groups without economic means.

Admission to primary education in the Soviet Union is purely by age and no examinations are required. During the primary period, continuous assessment by teachers is the rule. End-of-year examinations are to be found in some schools, but they are optional and are locally determined. There is no school-leaving examination for the primary period, nor are there secondary school admission examinations.

At the end of the five-year (incomplete) secondary period, an examination is required in no more than two or three subjects, given

by the teachers of the school. This examination is in both the written and oral forms. Following this secondary period students may proceed to special vocational schools where admissions examinations reflecting the schools' particular curriculum are required.

Students who continue in the general secondary school cycle for two more years are required to pass a 'maturity' examination which results in the award of a certificate of completion. This examination is also internal, though it may be administered by teachers of the school who have not actually taught the specific students. The 'maturity' examination covers in its written form, language, literature and mathematics. The other subjects of the curriculum are evaluated orally.

As can be seen, the uniform school curriculum in the Soviet Union is a basis for uniform testing, but such tests are constructed, under supervision, in the respective schools.

Students who complete the secondary technical programme must complete examinations in both the 'maturity' aspect and in the area of their school's specialization.

Admission to universities is determined by examinations which are taken in the faculty of the student's intended university specialization.

At present, there is continuing discussion as to the need for students to be subjected to dual examinations at the end of their full secondary education: the one in the spring, the 'maturity' examination and in early fall, the university admission test of the faculty. The tendency of such discussion is to eliminate the former, since secondary school completion can be attested to by school records and the university admissions examination can continue to serve its screening purpose.

In Japan, since World War II interest in the development of tests has continued to grow. Standardized tests are produced by local publishers and, though not required, are used with pupils in the

elementary and secondary schools. The results are recorded on standardized school record forms. Although these tests of subject matter achievement are used optionally in the schools, they respond to a strong motivational situation for students who wish to continue their education. We note that the tests are available, with keys and manuals, for sale in the open market.

Although Japan has a school system much like that of the United States (6-4-4), testing is conducted in conjunction with screening for the upper secondary school. The screening is based on the results of previous scholastic records and achievement tests. Third-year pupils in the lower secondary school take a prefectural achievement test which is included in the screening procedure.

Further testing is required for admission to a university, usually with tests proposed by the individual institution. These are generally objective tests and have been criticized for their tendency to restrict the patterns of student learning in the preceding educational levels.[1]

We may summarize as follows:

(a) in most countries continuous assessment is done by teachers throughout the primary years;

(b) there is a tendency to maintain school-leaving examinations where they now exist, but to join their scores with the cumulative records;

(c) the frequency of primary level testing is diminishing on a close-of-year basis, though a growing trend is shown toward standardized achievement testing at specified points;

(d) the frequency of start-of-secondary tests seems to be maintained where they have existed. In this aspect, pressure is being felt to admit students to secondary education on the basis of the completed primary cycle;

---

[1] Reported by: K. Nakayama, in Lauwerys and Scanlon, op.cit., pp. 223-226.

(e) testing in the secondary school is generally stable - internal tests seem to be the rule. Some uses of nationally standardized tests are seen with increasing frequency;

(f) end-of-cycle tests are required at the secondary level. These serve as certificates or as credentials for further education. In many countries, universities require and provide their own orientation-testing procedures;

(g) in almost all countries, the tendency is to reduce the frequency of national, required tests, except as noted above in (f);

(h) although end-of-cycle and admissions testing has been usually written (essay) and oral, the over-all tendency is toward objective, short-answer test items, with a lessening frequency of oral examinations.

It is clear that the data obtained from evaluation of students' achievement are used to assign them to an educational track, to retain them for a year's repetition or to anticipate their leaving the school system. In a number of countries more attention is paid to student vocational guidance and to orientation of parents.

Implications

The developments reported above are associated to a great extent with the slow-moving, but coninuous trend toward egalitarianism in the social structure, presumably to be achieved by egalitarianism in the schools. Reduction of fixed and required examinations and increased activities to provide educational experiences which will compensate for socio-economic and cultural inequalities are factors in the expressed desire to reduce focus upon 'elites'. The giving of exclusive attention to the development of highly selected and well-nurtured (educationally) cadres is not eliminated in any country, but is decidedly diminished and thus is more in keeping with the expressed goals for national systems.

Inferentially, at least, evaluation of teaching effectiveness is considered as a by-product of student success or failure in the evaluation process. There is no evidence that generalized curricular studies or reforms are instituted following study of evaluation data. In some countries, structural reforms such as extension of the compulsory schooling age or introduction of comprehensive lower schools, are in evidence, related in part to evaluation data, but even more to the continuing clamour for school opportunity and to the rising qualifications called for by the labour market. These latter changes are occasionally associated also with revision of over-all school system objectives.

In terms of the purposes of this exploratory study, that is, to determine the means of evaluation of educational output, we find the first and foremost attention being given for student achievement at some point within, or at the end of, the educational system.

As to the evaluation of the system itself, this is generally oblique or by inference. The student is evaluated as a learner. His achievement as a learner is seen, in some situations, as reflecting his socio-economic or cultural environment, including parental and family conditions. In some aspects, educational investments (inputs) are considered to be related to achievement: school budgets, teacher qualification, teacher salaries, size of classes, available and used educational materials, etc. There is little or no evidence of studies of base-line achievement data and of ensuing achievement of controlled populations. The value-added concept is seldom utilized in system evaluation.

Thus we find, almost in polarized contention, the preferences of economics-oriented observers and those of educational-achievement-oriented partisans. The former express their evaluation of educational system outputs in terms of manpower production, unit costs, and ultimately cost-efficiency assumptions. The latter seek to evaluate educational outputs from the quality basis of student achievement,

operational relationships between curricular processes and student behaviour, and with the contention that system goals should be made explicit so that their achievement may be determined and quantified for reliable evaluation.

It would appear that one group concedes the difficulty in quantifying quality elements and prefers to develop evaluative constructs for the already-quantified elements. The other group insists that determined quantities are not really evaluative of the system or of its processes, and maintains that there is a verifiable relationship at near-term and at long-term between inputs, processes and student quality outputs.

We would admit to being partisan in the debate since we feel that there are few remaining areas open to consideration once the techniques of cost-analysis, unit costs and manpower output are utilized. What remains is the calculation continually of ever-more refined input-output relationships.

It is our view that a more productive area for study resides in the consideration of system objectives which ought to be explicit and therefore more controlling in terms of inputs and processes. Student achievement in cognitive areas is now quantifiable and, with further research, may be more reliably correlated with inputs and curricula. Well-controlled studies on a longitudinal basis may provide further interpretations for the evaluation of system effectiveness. We look also to the heuristic potential of continuing research in assessing non-cognitive achievement goals.

We see, therefore, not a continuing polarization of hostile views, but an amelioration of differences and a more useful synthesis based upon continuing research in the various areas of school functioning.

In the following section of this report, we will describe current and developing research to achieve these purposes.

PART III

CURRENT RESEARCH ON EVALUATION OF EDUCATIONAL OUTPUTS

Research presently being carried out

In this section of our study we will review the major efforts currently being made in the evaluation of educational output. As will be noted, our emphasis is upon the qualitative aspects of school systems and those aspects which are related to student achievement of system objectives. Much of the work reported here is centred around student achievement although there is considerable interest in the determination of other indicators of system effectiveness.

Since other IIEP studies have dealt with economic indicators, our discussion will refer to these only slightly if at all.

On the international scale, the major work has been done by Husén and his colleagues in the International Project for the Evaluation of Educational Achievement.[1] This work was based upon an earlier feasibility study of over 8,000 thirteen-year-old students in 12 countries with tests in science, geography, reading comprehensiveness and non-verbal intelligence.[2]

The Husén study (known as the I.E.A.) evaluated mathematics achievement in 12 countries by developing nationally-standardized tests as established by committees of specialists in each country.

Two major groups in the secondary schools were sampled: (a) thirteen-year-olds, being the oldest students still in full-time secondary education in all 12 systems, and (b) pre-university grade students. In addition to the mathematics tests, students responded to items about their attitudes and their backgrounds. Additional data were supplied about the individual schools and the national system. A total of 133,000 students in 5,450 schools were tested and questionnaires were received from 13,500 teachers and 5,400 principals.

---

1 Torsten Husén, op.cit.

2 A.W. Foshay, Educational Achievements of Thirteen-year-olds in Twelve Countries, Hamburg: Unesco, Institute of Education, 1962.

Husén describes the operations of the secondary mathematics study and reports that:

" ... it was possible to study the 'effectiveness' of different educational practices in school or class organization in relation to both cognitive and non-cognitive outcomes. These outcomes were also related to social, economic and ecological factors."[1]

He points out that the important thing:

" ... in carrying out cross-national evaluations, is to cover the broadest possible spectrum of objectives and contents at a particular age or grade level."[2]

Husén emphasizes that the study did not seek to make comparisons between countries as if in a contest.

In the second phase of the I.E.A. Project ten-year-olds and fourteen-year-olds will be studied on achievement in science, mother tongue, civics, and a foreign language, in 19 countries. In addition to seeking to relate multiple variables to school achievement, this phase will attempt to evaluate system retention as a form of relative productivity.

While Tyler notes the potential of cross-national evaluation, he agrees that the:

"most critical theoretical problem is the definition of comparability in terms of objectives, content, and target populations ... "[3]

In an interview with Husén, we asked about the interest of I.E.A. in studies of a longitudinal nature which would test the same sample of students at least a second time. The results might provide

---

1 T. Husén, "International Impact of Evaluation", in Ralph W. Tyler, (editor), op.cit., pp. 340-341.

2 Ibid., p. 343.

3 Ralph W. Tyler, op.cit., p. 397.

further meaningful information about system effectiveness. The reply was that this proposal was not at present in the plans of I.E.A. While such cross-national studies may not be forthcoming in the near future, we look to the development of longitudinal testing within the individual countries.

On the national scale, numerous countries were not reported in the literature as conducting studies, nor was there material available to indicate that studies of output evaluation are of current interest or pre-occupation. In several countries, officials who were sent letters of inquiry about system evaluation, student achievement testing, or curriculum evaluation, failed to respond. Therefore, the absence of particular countries in the following review indicates our own inability to obtain information of consequence to this study.

For England and Wales, a major research centre is the National Foundation for Educational Research. While many of the research projects of the NFER are concerned in some way with the evaluation of educational effectiveness, the greatest emphasis is now placed upon the collaboration with I.E.A. for the coming phase of testing in England. Thus, the NFER is involved in preparation of tests and in the necessary planning for data collection in both the primary and secondary stages of this project.

Studies in England of pupil progress as related to motivation, expectation and home factors were reported by Pidgeon.[1] Additional interest is shown in the effects of separating pupils for placement in the modern schools and in the grammar schools. Considered as having effects upon achievement are such curricular and structural elements as grade placement versus age grouping and streaming, age of beginning school, teacher training, and differing social conditions.[2]

---

1 Douglas A. Pidgeon, Expectation and Pupil Performance, Stockholm: Almqvist and Wiksell, 1970.

2 Ibid., pp. 82-85.

As reported by Wall[1], the NFER has in the past conducted research studies on curriculum evaluation. Projects of the Nuffield Foundation have been studied as have earlier curricular programmes in reading and mathematics. The NFER for many years has provided test services for diagnosis of educational difficulties, day-to-day evaluation of pupils' work, and for educational guidance.

The Examinations and Tests Research Unit of the NFER studies the established standards and practices of C.S.E. examining boards to ensure comparability. This Unit is also undertaking research on secondary school pupils' activities after having taken the C.S.E.

Research reported in Wales[2] notes a critical analysis of previous research on the effects upon academic attainment of co-education. This study was especially concerned with mathematics and English. Other studies were of bilingualism (Welsh-English) and educational achievement, a problem which is relevant to both developed and developing countries.

Among studies reported which reflect sociological aspects, are several which conclude that:

> "the general improvement in academic achievement in relation to ability in the post-war period has not been evenly distributed. The improvement has been greater in the upper and middle classes and the reduction in the differential rate of achievement between the social classes has been small. This suggests that there is a continuing and ... even accelerating inadequacy in the British system of education in relation to

---

1 W.D. Wall, "The Work of the National Foundation for Educational Research in England and Wales", in H.J. Butcher (editor), Educational Research in Britain, London: University of London Press, 1968.

2 Schools Council, Welsh Committee, Educational Research in Wales, H.M.S.O., 1968.

its generally accepted selective function and that it is not an efficient agent for the sorting out of available ability in all classes.[1]

The Scottish Council for Research in Education has undertaken surveys of scholastic achievement and has participated in international comparative studies of pupil achievement.[2]

In his summary of curriculum evaluation research in Britain, Williams refers to important research strategies:

(a) the taking into account of the special outcomes expected of each curriculum as well as the over-all common outcomes and the need to formulate such expectations clearly;

(b) the use of testing instruments that will have implications for the pupil's performance later in his career;

(c) the focus on short-term studies on the components of the curriculum rather than to attempt to treat the curriculum as a whole in evaluation;

(d) evaluation of curricula should take place on a recurrent basis so as to determine patterns of effects;

(e) the distribution of testing should make it possible to assess a variety of aspects of curriculum effectiveness by representative and equivalent pupil samples.[3]

---

1 J.L. Williams, "Sociology and education in contemporary Wales", in Schools Council, op.cit., pp. 42-43. In this report, Williams refers to the contributions of English researchers in sociology and education.

2 D.A. Walker, "The Work of the Scottish Council for Research in Education", in H.J. Butcher (editor), op.cit., pp. 33-44.

3 John Williams, "The Curriculum: Some Patterns of Development and Designs for Evaluation", in H.J. Butcher (editor), op.cit., pp. 207-208.

In general, the above discussion is a relatively clear picture of the research on system evaluation now prevalent in Britain. Examinations are monitored for standardization, individual subject curricula are studied as to student achievement and with some follow-up studies, and international comparative studies continue.

However, the problems continue to be given serious consideration as witnessed by the following:

> "There is a need for a look at the whole exam system. The Schools Council has just set up a committee to look into the whole curriculum and a similar committee is needed on exams. Then the questions might be raised about where we need predictive tests, and where we need attainment tests, ... and what other forms of assessment need to be built in (teachers' reports, course work, and so on).
>
> "Now is a good time. The Council's curriculum projects are a stimulus to look at new forms of evaluation."[1]

It may be the task of the educational sociologist (or his reward) to engage in the more challenging considerations and research which were noted by Williams (supra).

With regard to research on evaluation in France, we have noted already (supra) that reforms are being undertaken in various aspects of the educational system, but the examination system remains relatively unchanged. There is, however, a growing tendency to rely upon teacher judgment for the evaluation of student achievement.

Research on marking procedures, therefore is being undertaken as reported by F. Bacher in the World Yearbook of Education.[2] She notes the need to study methods for improving objectivity of teacher evaluations by marks. She concludes that the aims of pupils' education must first be defined, and only then can an attempt be made:

---

1 Anne Corbett, "The Story of O and F", New Society, 2 July 1970, pp. 16-18.

2 op.cit., pp. 95-100.

"to apply techniques designed to show whether they have or have not been attained. This is a field of research which has been little explored as yet."[1]

Interviews with M. Reuchlin (Director of the Institut nationale d'étude du travail et d'orientation professionnelle) and with J. Drevillon at the University of Caen indicated that research activities in France are directed at improved and expanded efforts in student orientation and guidance. This is considered even more critical than before in view of increased enrolments in secondary education, the growing involvement and interest of students and parents in decisions about their continued schooling, and the greater reliance upon student dossiers for such guidance.

French researchers, including M. Reuchlin, are also participating in the I.E.A. studies reported earlier.

An instance of the growing concern for further research is the statement of J. Majault:

" ... thought precedes action. Particularly as regards education.

"Educational research is relevant to the contents of instruction, to methodology, and to the means of carrying it out. That is to say that all the problems are to be considered: /including/ studies of the reaction of students to different procedures followed in the various subjects and to the results achieved."[2]

While the above does not describe on-going or planned research, it is a call for the undertaking of research and the exposure of numerous problems to careful and objective study.

---

1 Ibid., p. 100.

2 J. Majault, La revolution de l'enseignement, Paris: Robert Laffort, 1967, pp. 181-184.

Bourdieu and Passeron[1] make a strong case against the utilization of successive tests to continually eliminate members of lower social classes. Thus testing is used to verify and support prejudgments of teachers and examiners. The authors see testing and certification as a means of 'legitimizing' existing culture and the established order.

However, their sociological outlook is that class distinctions separate the 'inheritors' from the 'disinherited' and that the educational system serves to maintain this separation.[2] They make no specific proposals for research although the picture is a pessimistic one.

Our studies lead us to conclude that educational evaluation in France continues to focus upon numbers, output of certificate holders, baccalaureats, and other graduates. It remains yet for critical evaluation of system aims, as suggested by Mlle. Bacher, to become a springboard for research in the qualitative evaluation of system outputs.

As we reported in the previous section, the Federal Republic of Germany demonstrates a tendency to relinquish reliance upon excessive testing. Much confidence is given to teachers' marks and notations in pupil dossiers. With the autonomous educational control in the Laender, curriculum development, evaluation of student achievement, and attempts to determine system effectiveness are to be found in variously-located research centres.

---

1 Pierre Bourdieu et Jean Claude Passeron, La reproduction: Elements pour une theorie du système d'enseignement, Paris: Les Editions de Minuit, 1970, pp. 194-206.

2 Ibid., p. 253.

In the Berlin Institute for Educational Research (of the Max-Planck Institute), researchers have been investigating the determinants of school achievement in three school subjects in grade 7 (gymnasium): German, English and mathematics. Difficulties arise in the attempts to establish comparability of teaching procedures and evaluative criteria in the different schools.[1]

Reports of additional studies of achievement are made by the Berlin Paedagogisches Zentrum, relating to objective procedures for admissions to higher education, comparisons of student achievements in the secondary Hauptschule and Realschule, and learning assessment in the foreign languages.[2]

At Frankfurt the University Institute for Social Studies is working on the effectiveness of political studies (civics) in elementary, intermediate and vocational schools.

Also in Frankfurt, in the German Institute for International Educational Research, researchers are participating in the I.E.A. Project, developing studies of school test construction, and seeking to establish indicators of curriculum effectiveness.

It would appear that research in the German Federal Republic is moving along the lines of achievement study in the primary and secondary schools. We find little evidence that such studies are producing indicators of system effectiveness in the Laender other than the acceptance of the teacher as a major observer of such

---

1 Edelstein, Sang, Stegelmann, "Unterrichtsstoffe und ihre Verwendung in der 7 Klasse der Gymnasien in der BRD", Studien und Berichte 12, Berlin: Institut fuer Bildungsforschung, 1968.

2 This and the following research were reported in H. Neubert (editor), Bestandserhebung von Projekten der Paedagogischen Forschung in der Bundesrepublik Deutschland fuer den Zeitraum, 1967/68-1969. This publication was produced under the sponsorship of the Volkswagen Foundation.

effectiveness in the individual case. It is assumed that the growing tendency toward curriculum study and evaluation will result in more highly focused efforts at output evaluation.

Japanese activities in research or evaluation centre around the work of the Educational Test Research Institute, which was established in 1963 in co-operation with the Ministry of Education.

The Institute is involved in test construction activities (aptitude tests, achievement tests, college admissions tests), in the gathering of data about evaluation procedures, and in establishing student guidance services. The ETRI is seen as similar to the Educational Testing Service in the United States.

With considerable student and teacher opposition the ETRI tests for university admission are not being adopted, though research efforts continue in the over-all areas of the Institute's aims.[1]

In the highly centralized school system of Sweden, researchers are developing a number of studies which may have interesting promise in the study of educational outputs.

The Swedish collaboration with the I.E.A. Project continues, with Professor Torsten Husén as chairman of the national committee.

A project which will involve student follow-up through the secondary and vocational schools is that of Henrysson, dealing with the forecast of progress in secondary school studies.[2]

This research will attempt to relate teachers' marks of individual pupils with objective tests of achievement and intelligence tests. Along with marks and tests, data were obtained by questionnaire

---

1 K. Nakayama, in Lauwerys and Scanlon, op.cit., pp. 226-232.

2 S. Henrysson, "Forecasts of Progress in Secondary School Studies", in School Research Newsletter, March 1970, National Board of Education, Stockholm.

about the students' social background, attitudes, aspirations and interests. The researchers hope to be able to determine how forecasts of student success can be improved.

Recent studies reported by Dahlloef indicate relationships between student groupings, process variables and standardized achievement test results.[1] Dahlloef and his colleagues determined that positively selected classes in secondary school mathematics learned more efficiently (in shorter time) than did unselected class groups. Achievement may be enhanced by a grouping process and outputs in a specific subject may be objectively so evaluated.

He takes an interesting view in his comments:

> " ... most educational research has concentrated on achievement level in rather general functions as the dependent variable without much concern for the educational process that has been producing these results and which - at least as regards studies on the macro-level - may be regarded as an almost neglected intervening variable."[2]

This view indicates the need for closer attention to curriculum experiences as they relate to objectives and as they may influence achievement both in terms of achievement and use of time. The latter is a major consideration which is seen as receiving too little concern in systems which stress comprehensive schooling.

---

1 Urban S. Dahlloef, Ability Grouping, Content Validity and Curriculum Process Analysis. Project Compass 13, Goeteborg: Institute of Education, University of Goeteborg, June 1969, pp. 14-20, 49-50.

2 Dahlloef, op.cit., p. 60.

In a later report[1], Dahlloef attempted to determine effectiveness indicators ('positive frame conditions') of mathematics learning in a sample of secondary students in Goeteborg.

Use was made of class observations (for interaction analysis), achievement tests, and questionnaires for students and teachers. As of the date of the report, all data were received and will be analysed so as to note the relationships which may be operative.

Although education in the United States is politically and financially decentralized, in recent years there have been trends toward common goals, similar curricula, almost universal methodologies for instruction, and nationally standardized achievement testing.

In view of the considerable mobility of teachers and of students, the 50 'independent' State systems and the several thousand relatively 'autonomous' school districts have become more alike than different. Their problems have become similar as well.

Calls have been heard across the land for improved achievement and for the provision of compensatory educational services where social and economic factors are detrimental to student progress. Nearly-uniform curricula have not produced nearly-uniform acceptable achievement. Pockets of non-achievement seem to be found in close conjunction with pockets of poverty. As a partial response to this situation, federal funds were contributed to States and school districts so that new programmes could be developed. Although various special allocations were made, we are primarily concerned with Title I of the Elementary and Secondary Education Act (1965): 'Financial Assistance to Local Education Agencies for the Education of Children of Low-Income Families'.

---

1 Urban S. Dahlloef and Ulf P. Lundgren, Macro and Micro approaches combined for curriculum process analysis: A Swedish educational field project. Project Compass 23, Goeteborg: Institute of Education, University of Goeteborg, April 1970.

Title I of the Act defined 'low-income', established guidelines and provided grants to local districts through the existing State departments. In the guidelines were explicit requirements that each project funded under Title I was to be evaluated every year, and that the plan for evaluation was to be included in the design of the project.

As examples of some of the projects undertaken in one city (New York), we list the following:[1]

After-school study centres: remedial instruction provided from 3-5 p.m. in 52 elementary schools (grades 3-6) and 47 junior high schools (grades 7-9).

Open enrolment: transportation was provided for pupils who preferred to attend schools outside their home districts; additional services 'followed' the child to receiving schools.

Improved services: additional teachers, counsellors, expanded services and increased materials were provided in 207 elementary and 24 junior high schools.

Transition schools: additional services were provided to improve the holding power of schools in communities in the process of social and economic transition; enrichment materials, special classes, additional teachers, etc., were provided.

More effective schools: limited class size, additional teachers, special materials were provided.

This listing is necessarily curtailed; there were many more projects. In 1965/66, New York City's Title I expenditure (including summer 1966) was $49,706,207. Nationally, Title I allocations in this period were $1,177,410,630.

---

1 Barbara R. Heller, A History and Description of ESEA Title I in New York City, 1965-1968. New York: Center for Urban Education, June 1968, pp. 1-50.

Since evaluations were required, evaluation research played a large part in the study and review of Title I projects. The limitations of this study do not permit further amplification. Yet, a few comments about the specific situation in which this writer participated are in order.

In New York City, evaluation research was conducted by the Center for Urban Education, for the most part. The Center is a 'Regional Education Laboratory', also funded under the Act (Title IV), and designated to undertake research in broad areas for its geographic region.

The goals of the Title I projects were generally the following:

(a) to improve the school achievement of children,

(b) to improve the attitudes of students toward schooling,

(c) to maintain community stability and sustain ethnic integration when possible,

(d) to satisfy parents and other community representations with regard to educational services.

Since evaluation of the projects was to be conducted and concluded each year, and since it was intended that feedback from evaluation would result in revision of the projects themselves, there were pressures of time upon researchers as well as the demand that research results provide for useful feedback. While the time element was usually met, there is no way to determine the effectiveness of the research as a mechanism for improved project planning. The time factor was in itself an impediment to such results.

For the evaluation of most projects, evaluation teams (or individuals - this writer was involved in both aspects) attempted to do the following:

(a) to establish the prior level of operation of the schools and their target population. This often consisted of determining achievement status, attendance, student and family mobility, attitudes, and expectations. It was

not possible to undertake all such procedures, though baseline data were sought and established in terms of the specific project,

(b) to establish control groups where this was feasible. The element of time pressure mitigated against this procedure in most cases, though often the median city-wide population or the population which established test norms could be utilized for comparison with target groups,

(c) achievement tests, observations by trained observers, attitude tests, interviews, and other data-gathering measures were utilized to determine outcomes following operation of the project process variables. Thus, intermediate and post-test procedures were utilized over a short-term longitudinal study to the extent that the target population remained stable and available. It was assumed, of course, that certain projects could be evaluated not by tests, but by the indication of whether families remained in the community or sought to 'emigrate' to more promising school situations. Other projects were considered as not appropriately subjected to evaluation over a one-year period. In fact, continuing projects were evaluated each year, usually by similar means, to determine whether outcomes were more promising over the long-term;

(d) as noted above, changes in behaviour could indicate the effectiveness of projects if all other variables remained relatively constant. Thus, the post-test differences were assumed to indicate project effectiveness. What was found (or sought) was evidence of educational value added in terms of achievement, of improved student attendance at school, of statements of satisfaction by school clients, and of maintained stability in the community 'mix'.

While New York City researchers performed their tasks as described above, over the country similar activities were under way. We have not found available to us a summary of the over-all Title I evaluations, or an evaluation of the evaluations themselves. Title 1 projects are still in operation, though perhaps with reduced funds, and evaluation research continues to be undertaken.

Federal funding provided an incentive to a variety of undertakings in educational evaluation. Another such was the establishment at the University of California, Los Angeles, of the Center for the Study of Evaluation of Instructional Programs (CSEIP). The Center was:

> "devoted to studying and improving the evaluation of instructional programs; ... to (clarify) the process of evaluating instructional programs by formulating appropriate theory; to identify, measure and study variables relevant to evaluation, ... and to develop and field test systems for evaluating educational programs and institutions."[1]

Major activities of the Center (now called Center for the Study of Evaluation - CSE) have included publication of the newsletter, Evaluation Comment, the holding of a symposium on 'Theory of Evaluation of Instruction' (December 1968), the development and field testing of preliminary work on an 'Elementary School Evaluation System', and the establishment of an exchange for the collection of instructional objectives. The latter activity is part of a Center project for research on objective-based evaluation.

Since the CSE is in an interesting development stage and is moving toward theory elaboration and field operations, we may anticipate progress in this area of research.

---

1 From "Statement of Intent", Evaluation Comment, CSEIP, Los Angeles: UCLA, January 1968, Vol. I, No. 1.

An attempt to develop a programme of educational evaluation on a large scale in the United States is the National Assessment of Educational Progress (NAEP). This project has been supported by foundation and U.S. Office of Education grants since 1964. Its basic purposes are to:

> "provide information that can be used to improve the educational process, to improve education at any and all levels where knowledge will be useful about what students know, what skills they have developed, or what their attitudes are."[1]

Plans include the testing of four different age groups, 9, 13, 17, 26-35 in 10 subject areas: art, career and occupational development, citizenship, literature, mathematics, music, reading, sciences, social studies, and writing.

For each testing area, the objectives were to meet the following criteria:

> "(1) The objectives must be the satisfactory goals for each subject area as seen by subject matter specialists.
>
> (2) The objectives must be ones which currently are accepted as goals of American education by most schools.
>
> (3) The objectives must be ones which are acceptable to thoughtful lay adults as reasonable goals of American education."[2]

The 'exercises' (rather than test items) were expected to be appropriate to the assessment of objectives, were to be samples of important skills, knowledge, or attitudes, and were to test the abilities of the greatest number, the average number and the most able 'assessees'.

---

1 Frank B. Womer, What is National Assessment? Ann Arbor, Michigan: Education Commission of the States, 1970, p. 1.

2 Ibid., p. 5.

The goal of the programme is:

"to be able to report, to summarize, in as meaningful a fashion as possible, the behaviour exhibited by groups of representative individuals."[1]

It is not a comparing procedure for individuals or groups.

Reporting categories are to be as follows:

(a) Age groups
  (i) 9-year-olds;
  (ii) 13-year-olds,
  (iii) 17-year-olds;
  (iv) adults between 26 and 35 years of age.

(b) Geographic regions of the U.S.
  (i) northeast;
  (ii) southeast;
  (iii) central;
  (iv) west.

(c) Size and type of community
  (i) large cities (above 200,000 population);
  (ii) urban fringes (cities adjacent to the large cities);
  (iii) middle-size cities (25,000 to 200,000);
  (iv) small town-rural areas (below 25,000).

(d) Sex

(e) S.E.S. 'socio-educational status'

"The intent is to be able to report results separately for assessees from disadvantaged homes."[2]

(f) Colour (black and non-black).

---

1 Ibid., p. 35.

2 Ibid., p. 40. This category was somewhat more clearly described as 'educational level of the parents of those assessed' in NAEP (newsletter), Vol. III, No. 3, September-October 1970, p. 5.

Material from the publications reporting on NAEP reiterate the information-seeking purpose of the project. It is not a model of standardized testing.

> "Appropriate standards of achievement should be and must be determined by persons knowledgeable about the abilities that youngsters of a given age bring to the learning process. A very important ingredient in determining such standards is a knowledge of the levels of achievement at which students are functioning. But present levels of achievement are not necessarily appropriate standards themselves."[1]

The plan for NAEP is for a series of cycles through 1981 to test reading, mathematics and science every three years and all other subject areas every six years.

> "The ultimate goal" "is the measurement of change (progress in knowledges, skills, understandings, and attitudes as they relate to meaningful educational objectives."[2]

Various criticisms and misgivings have been expressed with regard to this national assessment, particularly from groups of teachers and other school personnel. These were concerned about the possibility of invidious comparisons which might be made, or of moves toward such comparisons. There were fears that national assessment would bring about more rigid curricular requirements as to contents, methods, and uses of materials.

When the Exploratory Committee on Assessing the Progress of Education (ECAPE) was founded in 1964, its membership did not include representatives of such groups as: the American Association of School Administrators, the Chief State School Officers, the National Association of Secondary School Principals, the Department of

---

1 Frank B. Womer, op.cit., p. 45.

2 Ibid., p. 46.

Elementary School Principals, the National Education Association, the American Federation of Teachers, the National Congress of Parents and Teachers, the National Association of State Boards of Education, and the National School Boards Association. In July 1968, when the project became CAPE, representatives of these groups were included. Since then less concern has been voiced about the negative effects of the project upon schools, curricula, and school personnel.

Although the project documents repeat that its data will help to answer questions about educational effectiveness, we must wait for more explicit information as to how its sampling, testing, and measurement of growth in attainment will be related to such effectiveness.

Other recent work on evaluation of outcomes is being done by the RAND Corporation of Santa Monica, California. Studies include the analysis of programme effectiveness in elementary and secondary education[1], an evaluation design for a school district's compensatory education programme[2], the specification of objectives for system evaluation[3], the study of relationship of school inputs to school performance[4], study of experimental design and evaluation of educational innovations.[5]

Continual work on assessment of educational achievement is undertaken by the Educational Testing Service of Princeton, New Jersey. This organization not only produces standardized test materials, but conducts an annual Conference on Testing Problems whose proceedings are important explorations of current and developing topics in the field.

---

1 Report of Rand Corporation: P-4035, February 1969.

2 op.cit., RM-5903-S.J.S., May 1969.

3 op cit., P-4099, May 1969.

4 op.cit., P-4211, October 1969.

5 op.cit., P-4360, April 1970.

Other individuals and teams have been active in developing research designs and proposals for evaluating educational outcomes. We note the continuing work of Bloom, Guba and Stufflebeam, Tyler, Stake, Popham, and others. Among the most useful of recent compilations describing American (and other) activities is the recent issue of the Review of Educational Research[1] of April 1970, dedicated to the topic, 'Educational Evaluation'.

We would be remiss if we did not refer also to the pioneer work of Flanagan and his associates in Project TALENT[2] and of James S. Coleman who, with his colleagues, investigated the quality of educational achievement as related to opportunity.[3]

It has been our purpose, in the foregoing paragraphs to describe the extensive interest and activity of American researchers in educational evaluation. We have gone into some detail as regards several projects and noted the general outlines of others. This does not signify value judgment, but rather the interest of this researcher in current work which appears to be widely applicable (or having such potential) and more readily focused on the broadest possible populations.

In the following paragraphs we review briefly the indications of evaluation activity and research in a number of other countries and organizations.

---

1 Published by the American Educational Research Association, Washington, D.C. Also important is the December 1969 issue of RER on "Methodology of Educational Research".

2 John C. Flanagan et al., The American High School Student, Pittsburgh: Project TALENT Office, University of Pittsburgh, 1964, Final Report.

3 Equality of Educational Opportunity, Washington: Government Printing Office, 1966.

A recent document from Cuba[1] reports that teachers in the primary grades (first, second and third) are responsible for determining attainment of educational goals by quizzes and continuous judgment. In the second primary cycle (grades four through six), some differentiation is indicated as between 'periodic tests' and quizzes for diagnostic purposes.

> The following statement indicates some research development:
> " ... The changes made in the evaluation of this (second) cycle have not yet reached a satisfactory stage of organization and, therefore, this situation continues to be the object of studies and analyses with the aim of changing quantity into quality, not only in regard to points given, but also in that the evaluation control the development of skills and abilities which are much more important in the elementary school cycles than the information contents."[2]

In Peru, a recently-established Centre for Educational Research (1967) has made some progress. Organized with the aid of a university of Pittsburgh team and US AID, the Centre's main task is to undertake research studies of problems in the qualitative aspects of school development.

Major studies of the Centre include projects on the improvement of operational efficiency, a follow-up study of secondary technical school graduates, the problems of drop-outs in barriadas, and regional variations in child growth, school performance, and over-all development. Further research projects using data from on-going studies will deal with testing, achievement, and evaluation.[3]

---

1 The Educational Movement: Cuba, 1969/1970, Conference on Public Instruction, Geneva, July 1970.

2 Ibid., p. 43.

3 R.G. Paulston, "Peru: Developing an Educational Research and Evaluation Centre", International Newsletter, Princeton, New Jersey: Educational Testing Service, January 1969.

In a publication of the Ceylon Ministry of Education[1], J. Alles discusses the need for evaluation and assessment. He urges that it be process-oriented rather than centred on individual achievement.

We are not aware of research on evaluation of educational outcomes in the USSR. Assessment procedures have been described (supra) and the criteria for assessment of pupil progress are delivered by the ministries of education of the various republics.

School evaluation (though not research) is described in the Ukraine as being conducted by teams of Ministry of Education observers. Studies were undertaken in sample schools by sitting in classes, verifying curriculum conformity, and conducting tests in language and mathematics to validate teachers' marks.[2]

A report from Korea indicates that research studies are being undertaken on such topics as: the entrance examination system (1967), the uses of technology to improve educational effectiveness, and a "national assessment of scholastic achievements for the qualitative improvement of education in elementary and middle schools", (both 1970).[3]

An interesting development is reported in a document on the promotion of educational research in Asia.[4] In an extended discussion,

---

1 J. Alles, Notes on structural and functional aspects of an educational system relevant to educational administration, Colombo, Ceylon: Ministry of Education, 1967, p. 14.

2 K. Kovalevskii, "The Ministry Verifies, Studies, and Recommends", reported and translated in Soviet Education, September 1969, Vol. IX, No. 11, pp. 26-29.

3 Hyun Ki Paik, Introduction to Central Education Research Institute, Seoul, Korea, 1970, pp. 5-7.

4 WORKSHOP REPORT, Unesco Regional Programme for Promoting Educational Research in Asia, National Institute for Educational Research of Japan 1968, pp. 16-19.

the question of curriculum evaluation is reviewed as necessary to educational development in the Asian region. This evaluation aspect is linked to objectives, content, time, teaching materials and methodologies. It is also proposed to assess pupil growth in relation to curricular objectives.

Recommended procedures include tests of achievement, observations, interviews, anecdotal records, rating scales, self-inventories and projection techniques. The potential of both external and internal evaluation is discussed.

Interesting discussions about research needs in Norway are presented by Frøyland[1] of the National Council of Innovation in Education and Dalin in a paper for the OECD.[2]

In a recent OECD conference, prepared materials on 'Differences in School Achievement and Occupational Opportunities'[3] reviewed the factors involved in student success or failure in school.

The studies included data from several European countries who participated in I.E.A. testing. The evidence of the data was that children from deprived backgrounds were more likely to drop out of school; social differences are cumulative through the university level, but begin to have effect at the primary level; and genetic factors are less important than home environment. There was agreement that more research is needed as to the relative importance of school factors for pupil achievement.

Boudon writes in an OECD discussion paper[4] on longitudinal studies regarding the complementary nature of such studies with

---

1 "How to Change? Curriculum Development for the Eighties and Onward", duplicated, Oslo 1970.

2 Per Dalin, The Process of Innovation in Education, duplicated, Paris, OECD, 1969.

3 Background Study No. 10, Conference on Policies for Educational Growth, Paris: OECD, May 1970, pp. 1-22.

4 Educational Growth and Educational Opportunity, Paris: OECD, September 1970.

cross-sectional studies (I.E.A.), and that both must be linked in further research on educational achievement. He stresses that pre-school and school variables cannot be overlooked in an analysis of post-school mobility.

Indications of developing research

In this section of our report, we wish to give special emphasis to a number of research developments which seem to offer exceptional promise. None are, to our knowledge, in operational stages although they may include elements which have been applied in field research.

Hemphill has proposed that evaluation studies be undertaken within a framework of decision-making.[1]

> "An evaluation study becomes a process of acquiring further information ... that can be used by the decision-maker as a conditional modifier of his present information. His probability estimates of the consequences of his contemplated decisions can be changed ... as a direct result of expected outcomes of an evaluation study. The outcomes also have ... an estimable probability ... "[2]

Decisions may then be made as to the evaluation itself, and finally, as to instituting new or modified programmes.

Hemphill stresses the need for more effective evaluations so that decisions which may be costly can be taken or not under reliable circumstances.[3]

---

1 John K. Hemphill, "The Relationships Between Research and Evaluation Studies", in Ralph W. Tyler, op.cit., pp. 189-220.

2 Ibid., p. 219.

3 A somewhat generalized discussion along similar lines is offered in Lee J. Cronbach and Patrick Suppes (editors), Research for Tomorrow's Schools: Disciplined Inquiry for Education, Toronto: MacMillan Co., 1969, pp. 170-200.

A discussion by Thomas[1] on benefit and cost implications for educational system outputs incorporates suggestions for improved school system analysis and evaluation. Analysis of data on inputs and output relations by operations research techniques can lead to some conclusions about school system productivity.

The Conclusions of the June 1970 OECD Conference on Policies for Educational Growth include as guidelines for policy:

> "Goals for educational growth and change in the 1970's should be made more explicit, and where possible indicators which would measure the performance of the educational system, both in relation to educational goals as such and the contribution of education to the wider social and economic objectives, should be established."[2]

Reporting on a project on the Elaboration of a System of Human Resources Indicators[3], Solomon notes that the meeting of experts proposed a list of "indicators of educational flow and efficiency of the educational system".

These indicators are purely quantitative, referring to enrolments, completions of levels and degrees, pupil/teacher ratios, teacher qualifications and infant (9 months to 26 months) protein intake.

---

1 J. Alan Thomas, "Cost-benefit Analysis and the Evaluation of Educational Systems" in Proceedings of the 1968 Invitational Conference on Testing Problems, Princeton; Educational Testing Service, 1969, pp. 89-99.

2 Conclusions, Paris: OECD, June 1970, p. 2.

3 Notes reported by E.S. Solomon in Annex C of the Report of the First Meeting of the Panel of Experts on Methodology of Human Resources Indicators, Unesco, Paris, 15-20 December, 1969.

Other indicators are concerned with employment of out-of-school youth (age 15-24), stock of high-level manpower, educational attainment of the labour force, labour force utilization, health-nutrition status, labour force modernization (literacy, birth rate, etc.), and magnitude of educational effort in monetary terms.

While we consider these 'indicators' as limited and presently unrelated to determination of educational quality, we believe that there is justification for increased collaboration between product-oriented and process-oriented researchers. We note this particularly in regard to the need to relate educational system 'efficiency' and monetary effort to the problems of curriculum development and realistic output evaluation.

PART IV

PROPOSALS FOR FURTHER STUDY

In the light of this exploratory report, and the reviewing of broad areas which seem to be most appropriate for continued research on evaluation of educational system outputs, we would make the following suggestions:

(1) A case study of several educational systems in order to:

(a) determine the setting of system goals, their definition, relevance, and interpretation as process elements in curricula;

(b) study the development of curricular activities, methods, uses of materials, and administrative arrangements which indicate the implementation of explicit system goals;

(c) follow the sequence from goal-setting to curricular processes to evaluation of outcomes in terms of goal criteria;

(d) analyse evaluation instruments and techniques to determine their use in measurement of school-determined outcomes;

(e) study system evaluation procedures to relate how knowledge of output quality affects reforms of curriculum, administration, and teaching practices;

(f) provide educational planners with quality data which will enter significantly into plan proposals.

The study proposed above would be a systematic global approach to the problems of educational growth and to the kinds of decisions which economists and educators will be required to propose in their collaboration in planning. Coming from this research would be more useful quality indicators of system effectiveness which could be related to social, economic and political indicators now being empirically studied.

(2) A case study of the ways in which such frequently-mentioned aims as 'character development', 'emotional growth' and 'rounded personality' are being implemented in school system processes and evaluated as explicit output. Such a study would be related also to the problems of non-cognitive educational skills (which we have discussed) and to the increased reliance upon continuous teacher judgment to determine progress.

(3) A case study of the ways in which evaluation of educational value added is undertaken within school systems. A major aspect of this research would be to study how external, non-school variables may be held constant and how compensatory education decisions are made.

In each of the above proposals we would hope to study the effects of economically-oriented decisions upon evaluative criteria and procedures. It seems reasonable to note that budgeting decisions and their implementation on a national level require the undertaking of concomitant evaluation decisions and implementation. Thus far, we have noted few examples which recognize this set of circumstances.

Finally, it is our belief that case studies of the nature proposed could be feasible in countries for which IIEP already has data and economic case study experience. However, we would prefer to limit output evaluation studies to countries where the primary education enrolment has already reached at least 50 per cent of the age cohort.

PART V

BIBLIOGRAPHY

ABLIN, Fred, (ed.), Education in the USSR. A Collection of Readings from Soviet Journals, New York, N.Y., International Arts and Sciences Press, 1963, Vol. I.

AGAZZI, Aldo, Les aspects pédagogiques des examens, Strasbourg, Conseil de la Cooperation Culturelle du Conseil de l'Europe, 1967, (L'éducation en Europe. Série II - Enseignement général et technique No. 10).

ALKIN, M.C., "The Development of Evaluation Theory", in: Evaluation Comment, Los Angeles, California, Center for the Study of Evaluation, October 1969, Vol. 2, No. 1, pp. 1-8.

ALLES, Jinapala, et al., "An Attempt at Restructuring Some Conceptual Frameworks used in Curriculum Development and Evaluation", in: Theoretical Constructs in Curriculum Development and Evaluation, Ceylon, Division of Secondary Eduation, Ministry of Education, 1967, (Working Paper No. 1).

ALLES, Jinapala, Notes on Structural and Functional Aspects of an Educational System Relevant to Educational Administration, Colombo, Ceylon, Ministry of Education, 1967.

ALTSZULER, Ida, Research Concerning the Functions of School Tests, Warsaw, 1960, 231 pp.

ANDERSON, C.A., The Social Context of Educational Planning, Paris, Unesco/IIEP, 1967.

ANDERSON, Richard, Comments on Professor Gagné's Paper entitled "Instructional Variables and Learning Outcomes", Los Angeles, University of California, September 1968 (Occasional Report No. 17).

ATIYEH, Naim N., "Examinations : Trends and Prospects", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969, p. 376 ff.

BACHER, Françoise, "Some Inquiries into the Problems of Real Assessment (Docimology)", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969, p. 95 ff.

BECKER, Hellmut, "Research and Planning in Education", in: Western European Education, White Plains, N.Y., International Arts and Sciences Press, Winter 1969/70, Vol. I, No. 4, p. 23 ff.

BEEBY, Clarence E., (ed.), Qualitative Aspects of Educational Planning, Paris, Unesco/IIEP, 1969.

BEEBY, Clarence E., The Quality of Education in Developing Countries, Cambridge, Harvard University Press, 1966.

BEN-PORATH, Yoram, "Aggregate Costs, Output and School Achievement", in: Donald E. Super (ed.), Toward a Cross-National Model of Educational Achievement in a National Economy, New York, N.Y., Columbia University, Teachers College, 1969, p. VII-1.

BEREDAY, George Z., (ed.), Essays on World Education. The Crisis of Supply and Demand, New York, N.Y., Oxford University Press, 1969.

BERGER, Wolfgang, Zur Theorie der Bildungsnachfrage. Ein Beitrag zur Identifizierung der Determinanten privater Nachfrage nach formaler Bildung, Berlin, / On the theory of assessing demand for education. A contribution to identify what determines private damand for formal education /, Institut fuer Bildungsforschung, 1969 (Studien und Berichte 19).

BERLAK, Harold, "Values, Goals, Public Policy and Educational Evaluation", in: Review of Educational Research, Washington, D.C., American Educational Research Association, 1970, Vol. 40, No. 2, p. 261 ff.

BLOOM, Benjamin S., "Learning for Mastery", in: Evaluation Comment, Los Angeles, California, Center for the Study of Evaluation of Instructional Programs, May 1968, Vol. 1, No. 2.

BLOOM, Benjamin S., Stability and Change in Human Characteristics, New York, N.Y., London, Sydney, John Wiley and Sons, 1964.

BLOOM, Benjamin S., (ed.), Taxonomy of Educational Objectives, Handbook I: Cognitive Domain, New York, N.Y., McKay Comp, 1956.

BLOOM, Benjamin S., Toward a Theory of Testing which includes Measurement - Evaluation - Assessment, Los Angeles, Center for the Study of Evaluation of Instructional Programs, University of California, 1968 (Occasional Report No. 9).

BOUDON, Raymond, Equal Educational Opportunity. Meeting on Longitudinal Studies Concerning Educational and Occupational Opportunities (30 September - 2 October, 1970), Paris, OECD., Centre for Educational Research and Innovation, 1970, (Educational Growth and Educational Opportunity: Common Project CERI XXII: CERI/EG/29/70.01). (Restricted).

BOURDIEU, Pierre., PASSERON, Jean-Claude, La réproduction. Elements pour une théorie du système d'enseignement, Paris, Les Editions de Minuit, 1970.

BRAYBROOKE, David., LINDBLOM, Ch., A Strategy of Decisions, New York, N.Y., The Free Press, 1963.

BRUGNOT, Jack., PLENT, Albert, Information sur l'enseignement en France et en Europe. Premier et second degré, Paris, Les Editions Sociales Françaises, 1968.

BURNS, Hobart W., (ed.), Education and the Development of Nations, Syracuse, Syracuse University Press, 1963.

BUSHNELL, David S., "An Educational System for the 70's", in: Phi Delta Kappan, Bloomington, Ind., December 1969.

BUTCHER, H.J., (ed.), Educational Research in Britain, London, University of London Press Les., 1968.

CALDWELL, Michael S., "An Approach to the Assessment of Educational Planning", in: Educational Technology, Englewood Cliffs, N.J., October 1968, p. 5 ff.

CAPELLE, Jean, Tomorrow's Education. The French Experience, London, Pergamon Press, 1967.

CARPENTER, M.B., RAPP, Majorie L., The Analysis of Effectiveness of Programs in Elementary and Secondary Education, Santa Monica, California, Rand Corporation, February 1969 (Selected Rand Abstracts 1969, No. 4035).

Central Advisory Council for Education (United Kingdom), Children and their Primary Schools, London, Her Majesty's Stationery Office, 1967, Vol. I : Report, Vol. II : Research and Surveys.

Central Education Research Institute, Korea, Introduction to Central Education Research Institute, Seoul, Korea, 1970.

CERVALL, Lennart, New System of Curriculum Development, Stockholm, National Board of Education, 1969 (duplicated).

CHOPPIN, Bruce H., "Can literary appreciation be measured objectively?", in: International Review of Education, Hamburg, Unesco Institute for Education, 1969, Vol. XV, No. 2, pp. 241-247.

COLEMAN, James S., Quality of Educational Opportunity, Washington, D.C., U.S. Government Printing Office, 1966.

Committee of Educational Development, (CED), "Innovation in Education: New Directions for the American School", Book Review by James Welsh in: Educational Researcher, 1969, No. 2.

COOMBS, Philip H., The World Educational Crisis. A Systems Analysis, New York, N.Y., Oxford University Press, 1968.

CORBETT, Anne, "The Story of Q and F. Society at Work", in: New Society, London, July 1970.

COX, Roy, "Reliability and Validity of Examinations", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969, Chapter 7, p. 70 ff.

CRONBACH, Leo J., SUPPES, Patrick, (eds.), "Research for Tomorrow's Schools: Disciplined Inquiry for Education", Report of the Committee on Educational Research of the National Academy of Education. National Academy of Education, Toronto, Ontario, McMillan Co., 1969.

/ CUBA /, The Educational Movement: Cuba 1969/1970. Conference on Public Instruction convoked / sic / by OIE / sic /, and the Unesco, Geneva, July 1970.

DAHLLOEF, Urban S., Ability Grouping, Content Validity and Curriculum Process Analysis, Goeteborg, University of Goeteborg, 1969, (Reports from the Institute of Education, No. 7) (Project Compass 13).

DAHLLOEF, Urban S., LUNDGREN, Ulf P., Macro and Micro Approaches Combined for Curriculum Process Analysis: A Swedish Educational Field Project, Goeteborg, University of Goeteborg, 1970 (Reports from the Institute of Education, No. 10) (Project Compass 23).

DAHLLOEF, Urban S., The Materials and Methods of Implementation in the Development of the Curriculum. Outline of a Model and some Illustrations from Sweden, Goeteborg, University of Goeteborg, 1970 (Reports from the Institute of Education, No. 9) (Project Compass 21).

DALIN, Per, The Process of Innovation in Education, Paris, OECD., Centre for Educational Research and Innovation, 1969, (CERI/EI/69.10) (Restricted), (Duplicated).

Deutscher Ausschuss fuer das Erziehungs- und Bildungswesens, (ed.), Rahmenplan zur Umgestaltung und Vereinheitlichung des allgemeinbildenden oeffentlichen Schulwesens, Klett, Stuttgart, 1959, pp. 1-55, (3rd issue).

Deutscher Bildungsrat, Empfehlungen der Bildungskommission 1967-1969, Klett, Stuttgart, 1970.

Deutscher Bildungsrat, "Strukturplan fuer das Bildungswesen", Empfehlungen der Bildungskommission, Bonn, Bundesdruckerei, 1970.

DOUGLAS, J.W.B., The Home and the School, London, 1964.

DUKE, Benjamin, C., "The Karachi Plan - Master Design for Compulsory Education in Asia", in: International Review of Education, Hamburg, Unesco Institute for Education, 1966, Vol. XII, No. 1, p. 73 ff.

DYER, Henry S., "Statewide Evaluation - What are the Priorities?", in: Phi Delta Kappan, Bloomington, Ind., June 1970, pp. 558-559.

EDDING, Friedrich, "Educational Resources and Productivity", in: George Z.F. Bereday (ed.), Essays on World Education: The Crisis of Supply and Demand, New York, N.Y., Oxford University Press, 1969, pp. 22-23.

EDDING, Friedrich, Methods of Analyzing Educational Outlay, Paris, Unesco, 1966, (Statistical reports and studies).

EDELSTEIN, Wolfgang., SANG, Fritz., STEGELMANN, W., Unterrichtsstoffe und ihre Verwendung in der 7 Klasse der Gymnasien in der Bundesrepublik Deutschland (Teil I), Berlin, / Curriculum content and its application to the 7th grade of Gymnasia in the Federal Republic of Germany (Part I) /, Institut fuer Bildungsforschung, 1968, (Studien und Berichte 12).

Education in France, Cultural Services of the French Embassy, New York, N.Y., 1967, No. 36.

Educational Testing Service, Proceedings of the 1968 Invitational Conference on Testing Problems, Princeton, N.J., 1969.

Educational Testing Service, Toward a Theory of Achievement Measurement, Proceedings of the 1969 Invitational Conference on Testing Problems. Princeton, N.J., 1969.

EIDELL, Jerry., KITCHEL, Joanne, (eds.), Knowledge, Production and Utilization in Educational Administration, Columbus, published jointly by the University Council for Educational Administration and The Ohio State University, 1968.

ELAM, Stanley., SWANSON, Gordon J., (eds.), Educational Planning in the United States, Itasca, Ill., Peacock, 1969.

FINDLEY, Warren G., (ed.), "The Impact and Improvement of School Testing Programs", The 62nd Yearbook of the National Society for the Study of Education, Chicago, Ill., University of Chicago Press, 1963.

FINDLEY, Warren G., "Overview of the Research on Educational and Psychological Testing", in: Review of Educational Research. Educational and Psychological Testing, Washington, D.C., American Educational Research Association, (ed.), February 1963, Vol. XXXVIII, No. 1.

FLANAGAN, John C., et al., The American High School Student, Pittsburgh, University of Pittsburgh, Project TALENT Office, 1964, (Final Report).

FLECKENSTEIN, Helmut, Hoehere Schulen auf neuen Wegen, Frankfurt, / New roads for secondary academic schools /, Hirschgraben, 1958, 128 pp.

FOLEY, Walter J., Educational Information Project, Washington, D.C., Department of Health, Education and Welfare, 1968.

FOSHAY, A.W., Educational Achievements of Thirteen-Year-Olds in Twelve Countries, Hamburg, Unesco Institute of Education, 1962.

FRANCE, Norman., WISEMAN, Stephen, "An Educational Guidance Programme for the Primary School", in: The British Journal of Educational Psychology, London, June 1966, Vol. XXXVI, Part 2, p. 210 ff.

FRANZEN, Erich, Testpsychologie. Persoenlichkeits- und Charaktertests, Frankfurt/M., / Test psychology. Personality and character tests /, Ullstein, 1958, 173 pp.

FRØYLAND, Egil, How to Change? Curriculum Development for the Eighties and Onwards, Washington, D.C., The National Council of Innovation in Education, 1970, (duplicated).

GAGE, N.L., Comments on Professor Lortie's Paper entitled "The Cracked Cake of Educational Custom and Emerging Issues in Evaluation", Los Angeles, University of California, Center for the Study of Evaluation of Instructional Programs, September 1968, (Occasional Report No. 21).

GAGE, N.L., (ed.), Handbook of Research on Teaching, Chicago, Rand McNally and Co., 1963.

GAGNE, Robert M., "The Implications of Instructional Objectives for Learning", in: Report of the Regional Commission on Education, Co-ordination and Learning Research, Pittsburgh, University of Pittsburgh Press, 1964, Chapter 23.

GAGNE, Robert M., Instructional Variables and Learning Outcomes, Los Angeles, University of California, Center for the Study of Evaluation of Instructional Programs, September 1968, (Occasional Report No. 16).

GLASS, Gene V., Comments on Professor Bloom's Paper entitled: "Toward a Theory of Testing which Includes Measurement - Evaluation - Assessment", Los Angeles, University of California, Center for the Study of Evaluation of Instructional Programs, September 1968 (CSE Report No.11).

GOODLAD, John J., School Curriculum Reform in the United States, New York, N.Y., The Fund for the Advancement of Education, March 1964.

GRZYWAK-KACZYNSKA, Maria, Tests in Schools, Warsaw, 1960, 267 pp. (Bibliography).

GUBA, Egon G., "The Failure of Educational Evaluation", in: Educational Technology, Englewood Cliffs, N.J., May 1969, pp. 29-39.

GUILFORD, J.P., Comments on Professor Bloom's Paper entitled: "Toward a Theory of Testing which Includes Measurement - Evaluation - Assessment, Los Angeles, University of California, Center for the Study of Evaluation of Instructional Programs, September 1968, (Occasional Report No. 12).

HALLS, W.D., "Analysis of Aims and Content as a Basis for Assessment of School Courses", in: Comparative Education, Oxford, Pergamom Press, December 1969, Vol. 5, No. 3, pp. 213-220.

HAND, Harod C., "National Assessment Viewed as the Camel's Nose", in: Phi Delta Kappan, Bloomington, Ind., September 1965 (duplicated).

HARTLEY, Harry J., "Limitations of Systems Analysis", in: Phi Delta Kappan, Bloomington, Ind., May 1969, p. 515 ff.

HELLER, Barbara R., A History and Description of ESEA Title I in New York City, 1965-1968, New York, N.Y., Center for Urban Education, June 1968.

HEMPHILL, John K., "The Relationships Between Research and Evaluation Studies", in: Ralph W. Tyler (ed.), Educational Evaluation: New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education, Chicago, Ill., University of Chicago Press, 1969, (Part II).

HENRICSON, Sven-Eric, Continuous Evaluation in Swedish Schools, Stockholm, The National Board of Education, Informationssektionen, September 1969, 8 p. (duplicated).

HENRYSSON, Sten, "Forecast of Progress in Secondary School Studies", in: School Research Newsletter, Stockholm, National Board of Education, March 1970 (No. 1970.1 (7)).

HIRSCH, Werner Z., Population and Education : Meeting Change with Change, Los Angeles, California, UCLA Institute of Government and Public Affairs, 1969.

HUSEK, T.R., "The Implications of different kinds of Evaluation for Test Development", in: Evaluation Comment, Los Angeles, California, Center for the Study of Evaluation, October 1969, Vol. 2, No. 1.

HUSEK, T.R., SIVOTNIK, Ken., "Matrix Sampling", in: Evaluation Comment, Los Angeles, California, Center for the Study of Evaluation, December 1968, Vol. 1, No. 3.

HUSEN, Torsten., and BOALT, G., Educational Research and Educational Change: The Case of Sweden, New York, N.Y., Stockholm: Wiley: Almqvist and Wiksell, 1968.

HUSEN, Torsten, "International Impact of Evaluation", in: 68th Yearbook of Education. Part II, "Educational Evaluation : New Roles, New Means", Chicago, National Society for the Study of Education, (ed.), 1968.

HUSEN, Torsten, et al., International Study of Achievement in Mathematics. Phase I. A Comparison of 12 Countries, Stockholm: New York, N.Y., Almqvist and Wiksel: J. Wiley, 1967.

HUSEN, Torsten, "Responsiveness and Resistance in the Educational System to Changing Needs of Society", in: School Research Newsletter, Stockholm, Sweden, 1968, No. 6.

HUSEN, Torsten, "Talent, Opportunity and Career: A 26-Year Follow-up", reprinted for private circulation from The School Review, Stockholm, June 1968, Vol. 76, No. 2.

IIEP, Manpower Aspects of Educational Planning, Paris, Unesco/IIEP, 1968.

IMMEGART, Glenn., PILECKI, Francis, J., "Assessing Organizational Output: A Framework and some Implications", in: Educational Administration Quarterly, Danville, Ill., Winter 1970.

INATOMI, Eijiro, Practical Moral Education Series: Teaching of Moral Education in Grades 1-6, Tokyo, Maki Shoten, 1959, 3 Vols.

INATOMI, Eijiro, Practical Moral Education Series: Teaching of Moral Education in Lower Secondary Schools, Tokyo, 1958, 198 pp.

INGENKAMP, Karlheinz., MARSOLEK, Th., (eds.), Moeglichkeiten und Grenzen der Testanwendung in der Schule, Weinheim, Berlin, / Possibilities and limitations of the application of tests in the school /, 1968.

INGENKAMP, Karlheinz, Untersuchungen zur Uebergangsauslese, Weinheim, Berlin, / Investigations of selection procedures in transitional stages of schooling /, 1968.

INGENKAMP, Karlheinz, "West Germany", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969, London, Evans Brothers Ltd., 1969, p. 140 ff.

Institut Pédagogique National, Direction de l'enseignement du second degré, Paris, 1954, 82 pp. (Mémoires et documents scolaires, nouv. sér. 4).

Institut Pédagogique National, La documentation Française, Paris, L'Organisation de l'enseignement en France, 1957, 114 pp.

Institut Pédagogique National, La promotion sociale, Paris, 1960, 32 pp.

ISAMBERT-JAMATI, Viviane, "Extension du public et 'baisse de niveau' dans l'enseignement du second degré", in: Revue Française de Sociologie, Paris, avril-juin 1970, XI, 2, p. 151 ff.

JACOBSSON, Bengt, School Reforms in Sweden. International Working Party on Educational Technology and the Learning Process, Geneva, Unesco, 1970 (ED/Conf.15/5) (duplicated).

JAEGER, Richard M., Evaluation of National Educational Programs: The Goals and the Instruments, Minneapolis, Minn., 1970 (duplicated) (Paper presented at the 1970 Annual Meeting: American Educational Research Association).

JENSEN, A., "How much can we boast IQ and Scholastic Achievement?", in: Harvard Educational Review, Cambridge, Mass., Winter 1969.

KATZENMEYER, William G., "School Organisation and Student Achievement. A Study based on Achievement in Mathematics in 12 Countries", in: Educational and Psychological Measurement, Richmond, Va., September 1968, pp. 647-648.

KATZMAN, Martin T., ROSEN, Ronald S., "The Science and Politics of National Educational Assessment", in: Teachers College Record, New York, N.Y., May 1970, Vol. 71, No. 4, pp. 541-586.

KIESLING, H.J., The Relationship of School Inputs to Public School Performance in New York State, Santa Monica, California, Rand Corporation, October 1969, (Selected Rand Abstracts, 1969, P-4211).

KOVALERSKII, K., "The Ministry Verifies, Studies and Recommends", reported and translated in: Soviet Education, September 1969, Vol. IX, No. 11.

KRATHWOL, David R., BLOOM, Benjamin S., MASIA, B., Taxonomy of Educational Objectives. Handbook II: Affective Domain, New York, N.Y., McKay Comp, 1964.

KRAVETZ, Nathan, "The Diagnosis of Educational Systems and Operations as a Prelude to Development Planning", in: The Fundamentals of Educational Planning: Lecture-Discussion Series, Paris, Unesco/IIEP, 1970, (IIEP/TM/47/70).

KRULEE, Gilbert., NADLER, Eugene B., "Career Choice and Curriculum Evaluation", in: Journal of Engineering Education, December 1961, Vol. 52, No. 3, p. 143 ff.

KUNKEL, Richard C., McELHINNEY, James H., A Rationale for the Evaluation of Curriculum, Muncie, Indiana, Ball State University, 1970, (photocopy, individual distribution).

KUNKEL, Richard C., McELHINNEY, James H., LUCAS, Lawrence A., An Evaluation of Curriculum Projects as Cues to Curricular Change, Minneapolis, Minn., March 1970. (A paper presented at the Annual Convention of the American Education Research Association).

LANDSHEERE, Gilbert, de., Equal Educational Opportunity. Meeting on Longitudinal Studies Concerning Educational and Occupational Opportunities. (30th September - 2nd October, 1970), Paris, OECD., Centre for Educational Research and Innovation, 1970 (Educational Growth and Educational Opportunity: Common Project CERI XXII: CERI/EG/EO/70.02), (restricted), (duplicated).

LAUWERYS, Joseph., SCANLON, David G., (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969.

LEGRAND, Louis, "France", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969.

LINDVALL, C.M., "Introduction", Defining Educational Objectives, Pittsburgh, University of Pittsburgh Press, 1964.

LORTIE, Dan, The Cracked Cake of Educational Customs and Emerging Issues in Evaluation, Los Angeles, California, University of California, Center for the Study of Evaluation of Instructional Programs, September 1968 (CSE Report No. 19).

MACINTOSH, Henry G., PENFOLD, Daisy M., "England and Wales", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969, p. 110 ff.

MAJAULT, J., La revolution de l'enseignement, Paris, Robert Laffort, 1967.

MARKLUND, Sixten, "Qualitative Evaluation of Teacher Training", in: School Research Newsletters, Stockholm, July 1969.

MAYESKE, George W., et al., A Study of Our Nation's Schools, Washington, D.C., U.S. Department of Health, Education and Welfare, Office of Education, 1969.

McLURE, D., Curriculum Innovation in Practice, London, Schools Council, 1968, (A Report on the 3rd International Curriculum Conference, 1967).

MENSCHINSKAYA, Natalie, "Learning Research Carried Out at the Institute of Psychology at the Academy of Pedagogical Science of the USSR", in: International Review of Education, Hamburg, Unesco Institute for Education, 1966, Vol. XII, No. 1, p. 16 ff.

Ministry of Education, Nepal, Final Report. First Educational Research Seminar, Kathmandu, Nepal, 1968.

Ministry of Education (United Kingdom), 15-18 Report of the Central Advisory Council for Education, England, London, Her Majesty's Stationery Office, 1959/60, Vol. I : Report, Vol. II : Surveys.

NAKAYAMA, Kazuliko, "Japan", in: Joseph Lauwerys and David Scanlon (eds.), The World Yearbook of Education 1969 - Examinations, London, Evans Brothers Ltd., 1969, p. 223 ff.

National Assessment of Educational Progress, Summary of Report 2. Citizenship: National Results - Partial, Denver, Colo., Education Commission of the States, 1970.

National Assessment of Educational Progress, Summary of Report 1. Science: National Results, Denver, Colo., Education Commission of the States, 1970.

National Assessment of Educational Progress, First Assessment Results Reported to Public, Denver, Colo., Education Commission of the States, 1970, Vol. III, No. 3.

NEUBERT, Hansjoerg, Bestandserhebung von Projekten der Paedagogischen Forschung in der Bundesrepublik Deutschland fuer den Zeitraam 1967/68-1969 / s.l.n.d. /, / Survey of Educational Research Projects in the Federal Republic of Germany during the period 1967/68-1969 /, (Teil II: Klassifikation nach Sachkategorien).

Norwegian Council for Science and the Humanities Commission for Educational Research, The Development of Educational Research in Norway. Summary Report, Oslo, Narset Trykk-Larvik, 1969.

OECD., Conclusions. Conference on Policies for Educational Growth (Paris 3rd - 5th June, 1970), Paris, 1970, (Directorate for Scientific Affairs: duplicated).

OECD., "Differences in School Achievement and Occupational Opportunities: Exploratory Factors. A Survey Based on European Experience", Conference on Policies for Educational Growth, Paris, Committee for Scientific and Technical Personnel, 1970, (Background Study No. 10), (restricted).

OECD., Final Report on the Workshop on the Management of Innovation in Education, June 29th - July 5th, 1969. St. John's College, Cambridge, Paris, Centre for Educational Research and Innovation, 1969. (CERI/EI/69.22: 2nd Rev.), (restricted).

OECD., Special Problems of the Developing Member Countries. Conference on Policies for Educational Growth, Paris, Committee for Scientific and Technical Personnel, 1970, (Background Study No. 13: STP (70) 23, Scale 2: restricted).

OTT, Jack M., The Taxonomy of Administrative Information Needs: An Aid to Educational Planning and Evaluation, Minneapolis, Minn., 1970, (Paper presented at the 1970 meeting of the American Educational Research Association).

OTT, Jack M., A Decision Process and Classification System for Use in Planning Educational Change, Columbus, Ohio State University, Evaluation Center, 1967.

PAIK, Hyun Ki, Introduction to Central Education Research Institute, Seoul, Korea, 1970.

PAIK, Hyun Ki, "Reform of Entrance Examination System (Korea)", in: International Newsletter. Educational Evaluation and Research, Princeton, N.J., Educational Testing Service, 1969, Issue VI, p. 7.

PASSOW, Harry A., "School Organisation and Student Achievement", in: International Review of Education, Hamburg, Unesco Institute for Education, 1968, XIV, 3, pp. 362-364.

PAULSTON, Rolland G., "Peru: Developing an Educational Research and Evaluation Center", in: International Newsletter. Educational Evaluation and Research, Princeton, N.J., Educational Testing Service, 1969, Issue VI.

PHILLIPS, Charles E., The Development of Education in Canada, Toronto, W.J. Gage and Co., 1957.

PIDGEON, Douglas A., Expectation and Pupil Performance, Stockholm, Almqvist and Wiksell, 1970, (Stockholm Studies in Educational Psychology 18).

POPHAM, James W., The Instructional Objectives Exchange: Progress and Prospects, Minneapolis, Minn., 1970 (A Symposium Presentation at the American Research Association Annual Meeting), (duplicated).

POPHAM, James W., Instruction Objectives Exchange Catalog, Los Angeles, UCLA Graduate School of Education, 1966 (A Project of the Center for the Study of Evaluation).

POSTLETHWAITE, T.N., "The International Association for the Evaluation of Educational Achievement", in: International Newsletter. Educational Evaluation and Research, Princeton, N.J., Educational Testing Service, 1968, Issue IV.

POSTLETHWAITE, T.N., "International Project for the Evaluation of Educational Achievement (IEA)", in: International Review of Education, Hamburg, Unesco Institute for Education, 1969, Vol. XV, No. 2.

POSTMAN, Leo, Comments on Professor Gagné's Paper entitled "Instructional Variables and Learning Outcomes", Los Angeles, University of California, Center for the Study of Evaluation of Instructional Programs, September 1968, (Occasional Report No. 18).

PROVUS, Malcolm, Evaluation of Ongoing Programs in the Public School System. Educational Evaluation: New Poles, New Means. The 68th Yearbook of the National Society for the Study of Education, Chicago, University of Chicago Press, 1969, Part II.

RANDALL, Robert S., "An Operational Application of the CIPP Model for Evaluation", in: Educational Technology, Englewood Cliffs, N.J., July 1969.

RAPP, Majorie L., BRUNNER, G.L., An Evaluation Design for San Jose Unified School District's Compensatory Education Program, Santa Monica, California, Rand Corporation, May 1969, (Selected Rand Abstracts, RM-5903-SJS).

RAPP, Majorie L., ROOT, J.G., Some Considerations in the Experimental Design and Evaluation of Educational Innovations, Santa Monica, California, Rand Corporation, April 1970, (Selected Rand Abstracts, P-4360).

REUCHLIN, Maurice, Pupil Guidance - Facts and Problems, Strasbourg, Council for Cultural Co-operation of the Council of Europe, 1964 (Education in Europe. Section II - General and Technical Education No. 3).

Review of Educational Research, Special Issue on: "Curriculum", June 1969, Vol. 39, No. 3, Washington, D.C., American Educational Research Association, 1969.

Review of Educational Research, Special Issue on: "Educational Evaluation", April 1970, Vol. 40, No. 2, Washington, D.C., American Educational Research Association, 1970.

Review of Educational Research, Special Issue on: "Methodology of Educational Research", December 1969, Vol. 39, No. 5, Washington, D.C., American Educational Research Association.

RIST, Ray C., "Student Social Class and Teacher Expectations", in: Harvard Educational Review, Cambridge, Mass., Longfellow Hall, August 1970, Vol. 40, No. 3.

ROSENSHINE, Barak, "Evaluation of Classroom Instruction", in: Review of Educational Research, Washington, D.C., American Educational Research Association, April 1970, Vol. 40, No. 2, p. 279 ff.

SAYLOR, Galen, "Pro and Con National Assessment", in: Teachers College Record, New York, N.Y., May 1970, Vol. 71, No. 4, (duplicated).