

**Evaluation of education:
Examples from an international
and comparative perspective**

Ingemar FÄGERLIND
Institute of International Education
Stockholm University

Plenary Session

INTRODUCTION

Politicians, educators and parents alike, share the same interests in knowing how to improve the standards of achievement in educational institutions. More and more crucial is then the ability to evaluate or determine the extent to which such institutions have accomplished what they were aimed to accomplish. Evaluation of education is becoming very important when measures of Relevance, Goal Achievement and Cost Effectiveness are asked for. According to Cooley and Lohnes (1976:3) evaluation is a process by which relevant data are collected and transformed into information for decision-making, while House (1980) defines evaluation as comparative, concerned with standards, involving value judgements, and is directly oriented to the decision-making process. Evaluation produces knowledge specific to a particular context and the success of evaluation research is contingent on the extent to which it has been relevant for policy-making. More important for evaluation success is the extent to which evaluation research rigorously conforms to accepted standards and norms of research itself. Policy decisions, programmes and reforms are only as good as the quality of the research upon which they are based.

Evaluation can be directed to the analysis of specific programmes as well as to the performance of entire systems. Evaluations can and should cast their nets wide. Not only are the effects of specific educational inputs, organizational contexts and outputs legitimate evaluation targets, but so too are the assumptions, ideologies and theories underlying educational practices. Evaluation research is just as likely to be critical of current practices as it is confirmatory.

Evaluation research cuts across theoretical paradigms and research methodologies. Evaluations quite frequently focus on the evaluation of education policies within one country, but evaluation research can also be comparative, as for example the studies of the International Association for the Evaluation of Educational Achievement (IEA) which, since the end of the 1950s, have investigated student performance in subjects such as science and mathematics Hayes (1991).

An often neglected dimension of planning decisions, and of the evaluation of planning, has been the realization that all such activities are ideologically grounded Paulston (1992). The underlying assumptions upon

which plans are based are as much political as they are scientific. A clear specification of development goals, and the role that formal education should play in achieving such goals, is important for evaluation research.

NATIONAL EVALUATION

In many countries the educational systems are undergoing massive deregulation of decision-making. Authority and decision-making powers have been increasingly transferred from central to local government bodies. At the same time a lower rate of economic growth and the necessity for avoiding further tax rises mean that there is a larger demand for return from resources invested. The requirement of more efficient resource utilization has brought to the fore follow-up and evaluation.

Such activities are seen as a natural part of educational activities. In my own country, Sweden, a National Evaluation was started in 1987. The Government instructed the National Board of Education to develop the content and forms of a National Programme of Evaluation. Special attention was to be given to those parts of education system which did not consist of examinations, as well as to the way in which the different parts of the programme were related to each other and how they interacted. The purpose of the National Evaluation is to enhance the knowledge of schools and adult education and to compile input documentation for decision-making at different levels (SNBE, 1987). The evaluation is performed in order to supply information concerning schools, their activities and results. The effects of the school system as well as the conditions governing school work will be elucidated. Assessment of productivity and efficiency will also be an essential part of the National Evaluation. Thus, systematic and continuous collection, processing and dissemination of information about the state and goal achievement of the school system with reference to the Education Act and the National Curricula is carried out.

To achieve a holistic picture of school activities, the National Evaluation will include both external conditioning circumstances and the internal activities of schools, the development of students' knowledge and skills and their social and emotional development and attitudes. The collection of this information is being done through various channels. County education committees, the pupils themselves, their teachers and school management are all providing information. A wide variety of data collection methods are being used. Written questionnaires and tests are supplemented with direct questionnaires, both of the school as a whole and of the teaching

of individual subjects. Other information will also be used, e.g. official statistics from Statistics Sweden as well as data derived from local evaluations. The government and the Parliament have recommended the National Evaluation to concentrate on the compulsory school during an initial phase. A similar evaluation and assessment of the upper secondary school and municipal adult education will be carried out later. The first round, started in the academic year 1988/89, was experimental in nature and was conducted in grades 2 and 5. Reading, writing, arithmetic, art education and music were fields included in the assessment of grade 2 students. In grade 5, the subjects were Swedish, Mathematics, English, Civics, Art and Music. Assessment of the students' knowledge and skills were to be recurrent, and one round of assessment will be carried out every three years. The assessment in each round will relate to a small representative sample of classes. Data concerning the students' health status, current statistics, special student follow-up and demographic data, all from Statistics Sweden, will supplement the National Evaluation.

The evaluation findings will be summed up every three years. They will be used by the National Agency for Education in its long-range planning, in work on curricula and for ongoing development work. The result will also be disseminated in the basic and in-service training of teachers. It is hoped that the results will be used by teachers in their classroom planning and for the production of teaching materials. The overriding purpose of the National Evaluation is to give the Government and the members of the Parliament a picture of the situation and development in the school system as well as input data for political decision to be taken.

Follow-up and evaluation are to be implemented in the individual schools as well as locally and nationally. The aim is for the various levels to bear responsibility for following up and evaluating the targets adopted at the respective level.

The targets and priorities adopted in the school plan are to be assessed at the local level, and in each school the targets laid down for the individual school's work are to be followed up and evaluated.

The aim of the *Riskdag* and government is to be able to monitor and assess adherence to national targets and guidelines. They must be capable of obtaining and understanding of the results attained in education. The municipalities will be obliged to report facts and circumstances of relevance for assessing educational activities. Municipalities that fail to fulfil their obligations may incur sanctions.

The national follow-up should take such a form as to provide a comprehensive view of educational activities with respect to the targets, implementation and results of tuition as well as parameters of organisation, staff and economic situation affecting schooling. The intention is that the *Riksdag* and the Government should receive information on a regular basis in order to obtain clear picture of the state of the school system. This is to be achieved through annual reports on outcomes and costs and also through more detailed reports about the situation in different schools issued at three-year intervals that will form the basis for longer-term decisions.

Two new government authorities have been created, in Sweden partly for the purpose of national follow-up and evaluation. The National Agency for Education thus has, as one of its main functions, the task of developing and being responsible for nationally coordinated and coherent follow-up and evaluation of the state educational system. A similar authority is responsible for evaluation within the field of Higher Education.

INTERNATIONAL COMPARISONS

National results are getting more interesting when they could be compared between nations. The most well-known comparative studies are the ones run by the International Association for the Evaluation of Educational Achievement (IEA) even if the International Association for Educational Assessment in Princeton has more recently performed comparative studies of similar nature, Lapointe, Askew, Mead (1992); Lapoint, Mead, Askew (1992). IEA was founded in 1959 and a mathematics study in 12 countries was performed in 1964. The results were first reported in a volume edited by Husén (1967). In this report hypotheses and finding were presented under three major headings: (1) problems of school organization; (2) problems related to curriculum and methods of instruction; and (3) effects of social factors. During the early 1970s, achievement in six subject areas (Science education, Reading, Literature, French as a foreign language, English as a foreign language, and Civic education) were investigated in 22 countries. Several countries looked upon their results as some kind of evaluation of the reforms that had taken place since the second world war. Nine volumes reporting results from these studies were published between 1973 and 1976. During the 1980s a second mathematics study and a second science study were performed as well as a classroom environment study and a written composition study. In the beginning of the 1990s a reading/literacy study is being performed and a third mathematics-science study is on its way (Hayes, 1991).

The IEA studies have managed to develop a large set of methods for measuring different aspects of attainment. In the cognitive field the main problem has been to construct instruments that can be used for international comparisons. In addition to the more conventional tests of knowledge, new domains of achievement have been explored, for example, patterns of response to literature and written communication. The IEA has generated an enormous increase in our knowledge of educational systems in the participating countries -their organization, enrolment, resources, and practices. The influence on policy has gone through a multistage process. The interest of the mass media in overall national results -the "Olympic Games" aspect- has started a public debate. Ministers of education and other persons in charge of the system have been forced to react to these results. Negative results always seem to get more publicity than do positive ones. More close analyses of the national data have been performed as a result of preliminary results, and in some cases concrete explanations have been found and reforms with effects on classroom practices are undertaken. There is now considerable interest in comparisons of quality over time. Did the systems improve from the 1970s to the 1980s and 1990s? IEA has built an organization for international cooperation in educational research which has worked for more than 30 years. Some 40 educational systems have been involved which is in itself a major achievement (Härnqvist, 1987). The IEA studies could be also criticized from many points of view. They are extremely expensive, take too long for publishing the results, and have difficulties to adjust to systems in developing countries. However, more and more developing countries are joining IEA and their Ministers of Education and researchers seem to think that such studies are important. One obvious thing is that the centres taking part have a possibility to build up a national competence in survey research, with a bias in the direction of large-scale quantitative surveys.

EVALUATION OF HIGHER EDUCATION

The immense expansion of higher education systems, increasing financial constraints in the public sector, demands for accountability and also a general tendency in higher education towards decentralization are factors behind the demands for evaluation field. In rapidly changing Europe, evaluation in higher education is becoming a very important issue for academic, administrators and politicians. In some countries evaluation of higher education is part of a national evaluation scheme in others each university could have its own agenda. Evaluations could be performed (1) in terms of individual students in relation to stated course demands (2) in

terms of an assessment of study programs and departments in a number of respects; (3) in terms of the whole system, its main components and internal relationships as well as comparisons with other countries with systems operating under similar economic and cultural conditions (Dahllöf, 1992).

For the intensive evaluation work carried out it is becoming more and more common with special units at the university level but also at the government level. Ministries and administrators have given expression to their lack of confidence in colleges and universities regarding quality and cost effectiveness. The fund-giving authorities want to have evidence for both the legislature and the public on various indicators of effectiveness for the state system of higher education. Effectiveness, representing degree of goal fulfilment and efficiency, indicating the output level produced at the lowest cost are used as measures. However, comparative studies between university institutions can only be made on the basis of a few variables. As soon as evaluation of quality is based on other interests than control and cost efficiency broader information is obtained.

Patterns of institutional decision-making, degree of innovation and development, planning in relation to resources, importance attached to students' learning quality, professional development of teachers, access to various kinds of service and formulation of institutional policy on important issues are some of the measures used. According to the aim of the evaluation, control or improvement, different measures are used (Franke - Wikberg, 1990).

Educational exchange and research cooperation programs of the European Community like ERASMUS, LINGUA and COMETT create a need for comparative information on quality of institutions and contents of study programs and courses at various universities and other institutions of higher learning all over Europe. The Council of Europe, the OECD and other international organizations have experiences of such comparative evaluations and ratings.

THE METHODOLOGY OF COMPARATIVE EVALUATION RESEARCH

Even the best data can fail to provide an evaluation and understanding of the performance of educational systems if inappropriate methods are used, or if the results produced by particular methodologies are misunderstood or misinterpreted. There is no single appropriate evaluation methodology,

rather the methodologies are appropriate to any given set of evaluation goals and data. Sherman and Webb (1988) have collected some papers that give a careful and systematic review of the current possibilities of qualitative methods in educational research.

SCHOOL EFFECTIVENESS EVALUATIONS

Survey studies of school effectiveness and other input-output studies have used linear regression strategies (Creemers and Sheerens, 1989). Educational production functions were developed to analyses of schools and educational systems.

Many evaluation studies which have an impact on educational policy have used linear regression and the educational production function in one form or the other. In industrialized societies the works of Coleman et al. (1966) and Jencks et al. (1972, 1979) are well-known examples, while reviews of educational production function studies by Simmons and Alexander (1980) and Fuller (1987) include examples from non-industrialized countries. Jencks et al. (1972), Keeves (1972), and Fägerlind (1975) used path analysis in their studies. The statistical procedure of partial least squares (PLS), a so-called soft-modeling technique, has been used in several of the IEA studies while a hard-modeling procedure developed by Jöreskog and Sörbom (1977, 1988) have been used by Munck (1979), and Tiujnman (1989). In spite of criticism, (Freedman, 1987), regression analysis and path-modeling will continue to be utilized in many future comparative evaluation studies.

These limitations of linear regression and educational production function approaches should not mean that they represent useless strategies for evaluation research. If the assumptions underlying linear equation models are carefully examined and research designs are constructed to overcome these limitations, then much can be gained in knowledge about educational processes which may be useful for policy, and for the improvement of the educational system with regard to the desired goals of development. However alternative strategies for evaluation research also exist.

ALTERNATIVE EVALUATION STRATEGIES

Many attempts to overcome the limitations of input-output studies for evaluation research have focused on refinements to previous methodological

procedures. The need to improve measurement of variables and to expand the number of variables in evaluation models have been emphasized by many researchers. Greater sensitivity to variations in socio-economic status between societies, as well as the many school outputs other than academic achievements, are important examples.

The use of longitudinal data in comparative evaluation research have been rare. The Malmö longitudinal study, which began with 1,542 grade 3 school-children in 1938, continues to provide important information on schooling and careers in Sweden, not only for the original sample but now for their children. This study is trying to investigate the relationship between socioeconomic data, intelligence, education and life career at different age levels (Husén, 1969; Fägerlind, 1975; Tuijnman, 1989). Comparisons with longitudinal studies with younger participants make it possible to evaluate the impact of educational reforms of the 1960s and 1970s and comparisons of results from longitudinal studies in different countries have been variable.

QUALITATIVE EVALUATION

In sharp contrast to the quantitative approaches thus far discussed, qualitative evaluation methods offer an alternative which directs more attention to in-depth analyses of process rather than structure. There has been considerable recent development in the use of qualitative methods, ranging from case studies to ethnographies, and including observational, documentary, and qualitative interview data. In naturalistic evaluations there is no search for generalizable laws, but rather for insights that can transferred from one context to another (Bhola, 1990). Patton (1980), who has described at length the use of qualitative evaluation methods, is correct when he states that the challenge to the researcher is to select the methods most appropriate to the concerns of the policy-maker and to the specific evaluation situation.

CHALLENGES FOR EDUCATION AND COMPARATIVE STUDIES

Education today, and to an even greater extent in the future, faces challenges of three kinds:

First, challenges connected with the *framework* created by societal changes, such as the growth of the knowledge industry, urbanization,

migration, shared ecology, integrated trade and labour markets, and increased communications.

Second, challenges dealing with the *goals* of education, such as quality versus quantity, elite versus comprehensive education, equality of opportunity, the establishment of a bigger common core of general education.

Lastly, challenges connected with the *means* of achieving the goals emerging from the new framework (Husén, Tuijnman a Walls, 1992).

In the new European society, facing the challenges of rapprochements as well as political integration and disintegration, and in many other parts of the world, comparative evaluation will play a more important role as a wider and more comparable information system is necessary for a true understanding of your own situation relative to that of others. In addition, it helps both individuals and communities in determining the directions of future developments (Commission of the European Communities, 1991).

REFERENCES

Bhola, H.S. (1990). **Evaluating "Literacy for Development" Projects, Programs and Campaigns**. Hamburg: UNESCO Institute for Education.

Commission of the European Communities (1991). **Memorandum on Higher Education in the European Community**. Brussels.

Cooley, W.W. & Johnes, P.R. (1976). **Evaluation Research in Education**. New York: Irvington Publishers.

Creemers, B.P.M. & Scheerens, J. (1989). "Developments in School Effectiveness Research", **International Journal of Educational Research** 13, (7).

Coleman, J. *et al.* (1966). **Equality of Educational Opportunity**. Washington D.C.: U.S. Government printing Office.

Dahllöf, U. (1992). "Evaluation of Higher Education in Sweden", In V. Chinapah (Ed.) **Evaluation of Higher Education in a Changing Europe**. Stockholm: Swedish UNESCO Foundation.

- Fägerlind, I. (1975). **Formal Education and Adult Earnings**. Almqvist & Wiksell International.
- Franke-Wikberg, S. (1990). "Evaluating Educational Quality on the Institutional Level", **Higher Education Management**, 2 (3): 271-292.
- Freedman, D.A. (1987). "As others see us: A case study in path analysis", **Journal of Educational Statistics**, 12, (2): 101-128.
- Fuller, B. (1987). "What school factors raise achievement in the third world ?" **Review of Educational Research** 57 (3): 255-292.
- Jencks, Ch. *et al.* (1972). **Inequality: A Reassessment of the Effects of Family and Schooling In America**, New York: Basic Books.
- Jencks, Ch. *et al.* (1979). **Who gets ahead ? The Determinants of Economic Success in America**, New York: Basic Books.
- Jöreskog, K.G. & Sörbom, D. (1977). "Statistical models and methods for the analysis of longitudinal data". In D.J. Aigner & A.S. Goldberger (Eds.), **Latent Variables in Socio-Economic Models**, Amsterdam: North-Holland Publishing Company.
- Jöreskog, K.G. & Sörbom, D. (1988). **LISREL 7. Analysis of Linear Structural Equations**. Uppsala: Department of Statistics: University of Uppsala.
- Keeves, J. (1972). **Educational Environment and Student Achievement**. Stockholm: Almqvist & Wiksell International.
- Härnqvist, K. (1987). "The IEA Revisited", **Comparative Education Review**, 31 (1): 129-126.
- Hayes, W.A. (Ed.) (1991). **Activities Institutions and People**. IEA. Oxford: Pergamon Press.
- House, E.R. (1980). **Evaluating with Validity**. London: Sage.
- Husén, T. (Ed.) (1967). **International Study of Achievement in Mathematics: A Comparison of Twelve Countries**, New-York: Wiley.

Husén, T. (1969). **Talent, Opportunity and Career**. Stockholm: Almqvist & Wiksell.

Husén, T., Tuijnman, A. & Halls, W.D. (1992). **Schooling in Modern European Society**. Oxford: Pergamon Press.

Lapointe, A.E., Askew, J.M. & Mead, N.A. (1992). **Learning Science**. Princeton: ETS.

Lapointe, A.E., Mead, N.A. & Askew, J.M. (1992). **Learning Mathematics**. Princeton: ETS.

Munck, I. (1979). **Model Building in Comparative Education**. Stockholm: Almqvist & Wiksell International

Patton, M. (1990). **Qualitative Evaluation Methods**. Beverly Hills: Sage Publications.

Paulston, R.G. (1992). **Mapping Paradigms and Theories in Comparative Education**. Pittsburgh: department of Administrative and Policy Studies.

Simmons, J. & Alexander, L. (1980). "Factors which promote school achievement in developing countries: A review of the research". In Simmons, J. (Ed.) **The Education Dilemma**, New York: Pergamon Press.

SNBE (1987). **National Evaluation to Give a Holistic Picture of School Activities**. Stockholm: National Board of Education.

Tuijnman, A. (1988). **Recurrent Education, Earnings and Well-being. A Fifty year longitudinal Study of a Cohort of Swedich Men**. Stockholm: Almqvist & Wiksell International.