

Norberto Bottani and Bernard Favre

School evaluation

Prospects, vol. XXXI, no. 4, December 2001

N° 120

Original language: French

Norberto Bottani (Switzerland)

Director, Unit for Education Research (SRED), Geneva. Taught educational science in Italian-speaking Switzerland before contributing to a national education research policy and the development of Swiss involvement in international activities. Between 1976 and 1997 he was in charge of research programmes at OECD, where he ran the international education indicators programme, edited the series 'Education at a glance: OECD indicators' and produced the first edition of *Education policy analysis*. Education policies are his main centre of interest. Recently co-edited *In pursuit of equity in education: using international indicators to compare equity policies* (2001). E-mail: norberto.bottani@etat.ge.ch

Bernard Favre (Switzerland)

Researcher in the sociology of education, SRED, Geneva. He has focused on the evaluation of innovations in schools, relations between families and schools, and the analysis of school communities. He is currently working on an analysis of primary education reform in Geneva and a research project on the socialization of children and young adults in and out of school in a working-class district. His recent publications include *Les écoles primaires genevoises* [Geneva's primary schools] (to be published), with Françoise Osiek. Electronic mail: bernard.favre@etat.ge.ch

SCHOOL AUTONOMY AND EVALUATION

INTRODUCTION TO

THE OPEN FILE

Norberto Bottani and Bernard Favre

Evaluation is a central concern in education policy and education research today.¹ The interest in it is not new. Education systems are very much concerned by the 'demand' for assessment, which for many years focused on the activity of pupils (their learning) and, to a lesser extent, that of teachers (their teaching skills). However, for some twenty years now the field has been evolving in two ways: on the one hand, evaluation has moved beyond school circles to become a front-line political issue; and on the other, it has been extended from pupils and teachers—let us say, people taken individually—to cover schools, systems and training policies.

The institutional effect

With regard to schools, the actual concept of an 'institutional effect' and its existence were not at first recognized. At the end of the 1960s, James Coleman² and Christopher Jencks³ questioned whether it was possible to alter the distribution of academic performance by acting on school-specific variables, such as class size, teaching hours, teachers' qualifications and

the organization of the school. According to Jencks: 'Differences between schools have rather trivial long-term effects, and [...] eliminating differences between schools would do almost nothing to make adults more equal. Even eliminating differences in the amount of schooling people get would do relatively little to make adults more equal'.⁴ These conclusions, which minimized the impact of differences between schools on performance, were not accepted unanimously and were the starting point for a long series of studies which opened up the 'black box' of the school. Researchers went into schools not to assess pupils, but to see what really went on inside, to observe the way in which the school functioned and to discover factors which did influence pupil attainment. This entry into the 'black box' enabled a number of variables to be pinpointed which can be said to have a definite and positive effect on learning.

The study by Michèle Rutter and her fellow researchers, B. Morgan, P. Mortimer and G. Houston, on the effects of secondary schools in England,⁵ which was among the critical works produced in reaction to the work of Coleman and Jencks, was an important step forward and marked a necessary shift in the history of school evaluation. The British researchers demonstrated that the effects of schools on pupil attainment are far from negligible. In so doing they opened up a new field of study, generating a whole flurry of research throughout the 1980s, notably on 'school quality and development'⁶ and effective schools. OECD's International School Improvement Project (ISIP)⁷ was part of the new trend and was to contribute significantly to its development.

Editorial series on best school practice are one of the unexpected by-products of these studies on the quality and effectiveness of schools. They were produced in order to 'spread the good news', that is, to inform practitioners of solutions that, on the basis of research findings, could be said to improve education. The explanatory brochure *What works*⁸ was one of the very first examples. The idea was taken up again a few years later by CERI, the OECD Centre for Educational Research and Innovation, which in the 1990s started to publish reports in a series entitled 'Innovations that work'. At the end of the 1990s, the International Bureau of Education (IBE) carried on this work by publishing its series 'Educational practices' in collaboration with the International Academy of Education.⁹

These thirty years of research into schools have resulted in a considerable mass of data and analyses which have focused attention on the functioning of schools by providing information that has itself led to changes in how the results of teaching and the ways in which pupils learn are interpreted. As Ballion summed it up:

research work on the 'effective school' reveals that at this level [that of the school] it is possible not only to identify differential factors of effectiveness, explaining for instance that schools with identical contextual and compositional variables can have unequal results, and that in any one school, one primary school, there can be, as a recent French study has shown, teachers who are effective and others who are not; but also that the school, even if it is not the only one, is an appropriate level of action for improvement.¹⁰

Schools have therefore ceased to be a mysterious world, shielded from indiscreet and prying eyes, in which teachers tend to shut themselves off in order to defend their mission. The research has also transformed the way in which schools are presented, not as 'production units', but as educational communities, learning organizations where professional teams are at work.

From knowing about schools to evaluating them

The debate on evaluating schools has a relatively short history, as becomes clear, for instance, on reading the report of the working group on the evaluation of the pre-service training system written by Robert Ballion as part of the work of the education/training/research commission of the tenth policy guidance plan in France, published in 1991.¹¹ Reference to schools evaluation in the report is confined to a single paragraph entitled 'The evaluation of production units'. The school is presented as a production unit, that is, as an 'entity generating its own functionality'; it can therefore no longer be conceived of 'as simply an aggregate of individuals',¹² it is effective in its own right. School evaluation is not addressed in *The international encyclopaedia of educational evaluation*, published in 1990.¹³

Nevertheless, the new way of looking at the school has introduced the idea that it is just as important to evaluate education systems themselves, by evaluating the performance of the pupils taken as a whole and the quality of schools, as it is to evaluate individual performance (pupils and teachers). Moreover, this evaluation, in the 'quality and development of schools' research context, aims above all to be 'formative'; it is not a question of ranking or sanctioning schools, but of providing professionals with the tools to reflect on their practices. As J.-L. Derouet has pointed out, school evaluations are useful as a potential means of improving life at school, for instance 'as a way of bringing about a measure of civil peace within an institution riven by antagonism between clans', or an instrument 'of cultural integration for a community'.¹⁴ Evaluation can alert the teaching staff to where its knowledge is lacking, inform it of what is basically needed for the smooth running of the school, and so on.

School evaluation: fears and risks

From the early 1990s the success of educational indicators aroused even further interest in these challenges, but at the same time introduced into the field of research and policy a new dimension which had previously been approached only with great caution, namely comparison between schools, variations within schools and between schools, and the difference in results between education systems. To that was added the fact that in many countries school evaluation was imposed as part of a reform package giving schools more independence. The implication of these reforms¹⁵—freedom of choice of schools, training

vouchers, liberalization of methods of administration of school systems—increased the need for knowledge about the nature, performance and functioning of schools, but in so doing gave evaluation, at least initially, a punitive function as much as an improving one.

This development has raised both legitimate and exaggerated expectations and fears about evaluation. While, on the whole, there is agreement on the need to evaluate education and not just pupils taken individually, there is significant disagreement as to the nature and modalities of the procedures to be adopted. The proliferation of experimental and formal evaluation programmes has clarified the debate only in part. Scientists are divided on the subject of the methods and effect of evaluation. As to teachers, they remain suspicious and are unconvinced by new schemes or proposals for participatory evaluation, arguments extolling the formative nature of evaluation and demonstrations of the advantages and merits of self-evaluation as compared with any form of external evaluation. Evaluation smacks of heresy. In countries where there is a wide variety of pupil assessment practices, such as in the United States of America, most teachers have to submit routinely to the chore of tests and questionnaires which, it may be supposed, a substantial majority would willingly do without.

School evaluation is therefore being introduced in a climate marked by distrust, doubt and scepticism. The purpose of this open file is to take stock of developments in this field. To some extent it is complementary to the open file on the evaluation of education systems published in vol. XXVIII, no. 1, March 1998, of *Prospects*, which in a way foreshadowed it.¹⁶

Interesting avenues for research are opening up and some of them are explored in the articles in this open file. Having said that, we must be under no illusion: as pointed out above, the evaluation of schools can also be used to regulate the education market and stimulate competition between schools. The texts presented here do not lend themselves to this purpose and even draw attention to this kind of misuse which jeopardizes the transformation of schools into privileged places of learning and development for pupils.

The Open File

The Open File we present here contains two types of text: (a) articles based on European situations, describing ways of evaluating schools that are consistent with their independence; and (b) papers illustrating how the independence of schools is coupled with their evaluation in four different national settings.

Needless to say, any operation to evaluate schools must, in order to succeed, be acceptable to teachers, whatever the type and method of evaluation adopted. The panoply of evaluation methods is such today that any may be envisaged as long as its scope, limits, risks, advantages and disadvantages are fully grasped. There is nevertheless one factor that can never be overlooked if the evaluation is expected to help improve education and learning, and that is the trust of teachers. If they are not convinced of the benefits that they, in the exercise of their profession, can derive from evaluation, there is a potential risk of wasting resources

and time, producing inaccurate data and sending out the wrong signals. Evaluation of schools cannot therefore be carried out without the teachers. It is important to take the full measure of this elementary axiom in the knowledge that the conviction of teachers in this regard can by no means be taken for granted and that in general evaluation, despite all the advances made, in particular in recent decades, is still not an integral part of educational practices or, even more so, of school management. All the articles in the file hinge on this requirement: how do we reconcile evaluation mechanisms with teachers' concerns? How do we design evaluations that are acceptable to the key players in education? What happens when evaluation practices that disregard the teachers' expectations and ignore their daily professional concerns are forced on schools? To what experiences may we refer in order to appraise factors of success and failure in evaluation procedures? Is there a strong connection between typologies of evaluation and teacher behaviour?

The article by *Abi-Saab and Alt*, the first in the European part of the file, describes France's experience with secondary school indicators (IPES). Among the cases described in this file, this is the one that proposes the closest linkage between a source outside the schools and the schools themselves. School-specific data, processed statistically and reduced to a set of key indicators, are supplied to the schools, together with a special computer application, so that they can be used as benchmarks to assist teachers and school heads in analysing outcomes.

In the same vein, the text by *Saunders* presents studies carried out in the United Kingdom to measure the value added to schools, in other words to correct and relativize the raw evaluation data by incorporating into them the particular variables of a school. This approach meets a demand for fairness, as it is not possible to put on an equal footing schools which are obviously very different from each other in a number of ways. Teaching in a school in a disadvantaged area is not the same as teaching in a school in a prosperous area: the pupils have a different cultural and social capital, the schools do not enjoy the same facilities and the staff itself varies considerably from one school to the next. These parameters must, then, be taken into account when a school is being evaluated. Saunders sums up the characteristics of the value-added-based approach and makes recommendations drawn from practical experience regarding the measurement and use of added value. In addition, she shows the way in which schools make use of the data on their added value that are measured by specialist centres and made available to them. The French and British experiences, though inspired by different theories, resemble each other in that they are characterized by the transmission to schools of tailored statistical information which should or could provide input for self-evaluation.

The experiences in the Netherlands and in Geneva described in the papers by *Hendriks et al.* and *Favre*, mark an important new stage, that of the internalization of evaluation practices and the disappearance from the scene of specialist external bodies providing *ad hoc* data. This can be seen in the ZEB0 project, which endeavours to provide the self-evaluation

approach with credible and valid instruments. In other words, it is not enough to carry out self-evaluation; it must be scientifically sound, leaving no room for improvisation and approximation. The question then is knowing how to equip schools to be fully in command of the evaluation operation by making use of the potential that independence can free up within school communities. The solution consists in ensuring that the school itself has control over the evaluation instruments and can identify with their purpose and use. In the Dutch experience, it is the stakeholders themselves—administration, teachers and pupils—who decide what information to gather, who gathers it and who processes it.

The Geneva study described by Favre falls within the context of a major reform of primary education in the Canton of Geneva. However, it concerns a phase that is upstream of evaluation and self-evaluation practices. This means that the instrumental dimension of the analysis is lost but in exchange it has a density and range that were missing altogether from previous exercises, since it covers the history of the school and an analysis of its organization and the elements that go into making it an educational community. The object is to describe the particular organizational rationale into which evaluation and self-evaluation necessarily fit. This rationale helps to explain, at least in part, teacher resistance to evaluation and any departures from the original purpose of the evaluation. As a result, the Geneva study may open up another field of study, namely the relationship between organizational learning and school evaluation and self-evaluation practices, or the ways in which the stakeholders' own knowledge of their school as an organization can be used for purposes of evaluation.

In the second part of the file there are four studies describing national policies dealing with the problem of regulating the independence of schools. The situations described in this section are very different. *Fiske and Ladd* shed light on New Zealand's experience in the context of an ambitious reform launched since 1989 and aimed at decentralizing the administration of the education system, giving considerable independence to schools, but without setting up a restrictive evaluation and self-evaluation mechanism to offset the effects of decentralization and act as an instrument to regulate the system as a whole.

Jansen describes developments in South Africa, where the requirement for school evaluation has taken the form of officially requiring schools to carry out their own evaluation, the object being to regulate schools and ensure that they align themselves on the objectives of national education policy. The independence of schools and the concern to embed and strengthen good practices within schools that led them to obtain increasingly satisfactory results are overshadowed by the imposition, in an almost underhand way, of a sophisticated control mechanism entailing the hijacking of self-evaluation. It is legitimate in this case to talk of hijacking. The independence of schools is to all intents and purposes emptied of meaning and scope.

The experience of Chile described by *Casassus* provides insights into an interesting new approach to school evaluation which aims to reduce teachers' distrust of evaluation schemes and secure their commitment to a process which they can only partly call their own

in order to ensure that, in the end, evaluation represents an opportunity to improve education. The Chilean approach could be said to be halfway between the impasse of the South African model analysed by Jansen and the extreme solution of a dilution of self-evaluation in a self-referential practice that would be found in the Geneva model described by Favre if it were seen as a form of self-evaluation. The school evaluation system introduced in Chile is not centred on teachers but on the school as a recognized community through incitements and rewards provided for and paid when the school is managed effectively by the teaching team.

Lastly, the paper by *Schmelkes* gives a particularly clear illustration of the difficulties and deadlock encountered when seeking to reconcile the independence of schools with evaluation, as we have already observed in South African education policy. The balance between the independence of schools and the evaluation of education is a delicate one; it is to be found neither in the external nor in the internal evaluation approach. The New Zealand solution puts a premium on independence but does not provide an evaluation model that takes account of the need to move a large group of schools forward, together if possible, in order to avoid the exacerbation of disparities in the education on offer. In France, the Netherlands and the United Kingdom ways and means of combining complex evaluation mechanisms for the benefit of schools are being tested. In Mexico, in a situation which is far less advantageous economically, the independence of schools is a theme that policies address with caution—there are reforms along those lines, as in New Zealand and Chile, but closely-controlled experiments, warranted by the enormous diversity in the supply of education available in the country. The fear that independence might lead to greater, rather than less, injustice in education is in fact a basic theme in this debate: we still do not have proof that independence—even when it is implemented partially or selectively—makes for a better distribution of educational opportunity, but equally we know that there is no definite proof that the model of centralized management established to ensure the uniform and therefore fair development of education throughout a country has been successful either. In the second case, there is abundant data, even though in many countries—not only in Mexico—the authorities disregard them or do not even know that they exist.

In conclusion, even if on a theoretical level it is acknowledged that the self-evaluation of schools has not only a formative value but should be part of the raft of new forms of regulation of education systems, in practice, the coexistence—although it would be more accurate to say intermeshing—of independence and evaluation has proved highly problematic. However, as shown by the experiences presented in this open file, the intermeshing of central and local regulation is possible, but cannot be done on the cheap: it requires considerable investment in terms of scientific and financial input, but above all in terms of clearly formulated objectives that enable schools to know from the outset what to expect, without excessive illusions but also without suspicion.

Notes

1. See, for example, the open file 'The Evaluation of education systems', *Prospects* (Paris), vol. XXVIII, no. 1, March 1998.
2. J. Coleman, et al, *Equality of educational opportunity*, Washington, DC, United States Department of Health, Education, and Welfare, Office of Education, 1966.
3. C. Jencks, *Inequality: a reassessment of the effect of family and schooling in America*, New York, Basic Books, 1972.
4. *Ibid.*, p. 16.
5. M. Rutter, et al., *Fifteen thousand hours: secondary schools and their effects on children*, London, Open Books, 1979.
6. For a synoptic study of this research trend, see C. Szaday, X. Büeler and B. Favre, *Schulqualität und Schulentwicklung: Trendbericht* [School quality and development: trend report], Bern, PNR33 Direction; Aarau, Swiss Co-ordination Centre for Education Research, 1996.
7. ISIP was set up in autumn 1982 and was operational until 1986. One hundred and fifty people from fourteen countries took part in this project on the improvement of the functioning of schools. The following fields were dealt with: (i) self-analysis of schools; (ii) the school head and internal agents of change; (iii) the role of support systems; (iv) research and evaluation in evaluating the functioning of schools; (v) the formulation and implementation of policies to improve school functioning; (vi) the development of a theoretical framework to improve school functioning.
8. United States of America, Department of Education, *What works: research about teaching and learning*, Washington, DC, 1986.
9. International Academy of Education and International Bureau of Education series 'Educational practices'. See: <http://www.ibe.unesco.org>
10. Robert Ballion, L'évaluation du système de formation initiale [Evaluation of the pre-service training system] in: *Eduquer pour demain : acteurs et partenaires* [Educating for tomorrow: actors and partners], Paris, La Découverte and La documentation française, 1991, p. 143.
11. French Republic, State Secretariat for Planning, *Eduquer pour demain : acteurs et partenaires* [Educating for tomorrow: actors and partners], introduction by René Raymond, foreword by Lionel Stoléru, Paris, La Découverte and La documentation française, 1991.
12. R. Ballion, *ibid.*, p. 143.
13. Herbert J. Walberg and G. Haertel, eds., *The international encyclopaedia of educational evaluation*, Oxford, Pergamon Press, 1990.
14. J.-L. Derouet, L'établissement scolaire, des principes aux réalités [The school, from principles to reality], *Sciences humaines* (Auxerre, France), no. 111, December 2000.
15. See in particular the summary paper by M. Duru-Bellat and D. Meuret, Nouvelles formes de régulation dans les systèmes éducatifs étrangers : autonomie et choix des établissements scolaires [New forms of regulation in foreign education systems: independence and choice of schools], *Revue française de pédagogie* (Paris), no. 135, 2001, p. 173-221, and Th. L. Good and J. Braden, *The great school debate: choice, vouchers and charters*, London, Lawrence Erlbaum Associates, 2000.
16. See the article by Sally Thomas, Value-added approaches of school effectiveness in the United Kingdom, *Prospects* (Paris), vol. XXVIII, no. 1, March 1998, p. 91-108.