

ACCOUNTABILITY, ASSESSMENT AND THE QUALITY OF EDUCATION

- A Finnish perspective

Dr. Pasi Sahlberg

MONTENEGRO EDUCATION REFORM PROJECT
Western European Education Reform Experiences
Budva, May 24 to 27, 2000

Education systems are in many ways in a similar situation to businesses: they should produce higher quality services for fewer resources in unpredictable circumstances. Higher educational quality does not only mean more relevant knowledge of the new fields but rather it refers to the achievement of such personal skills, attitudes and habits of mind that are essential for overall development of our societies, and sustainable development of our planet. Most governments are considering the same question. However, what is going on in different parts of the world in terms of improving their education systems to combat the challenges of changing times is interesting and worth a closer look. What we have learnt within the education reform community is that there are no fixed solutions to the problems that exist in our schools. There are good and not so good experiences that we must study in order to find the best ways to cope with our own challenges.

This paper focuses on quality issues of education in light of national education reform that has its roots in equity, equal opportunity, and democracy. More precisely, the following paragraphs discuss the role of educational assessment and evaluation as parts of the education system. The main line of thinking is that as the education systems become more decentralised in terms of transferring authority from the top to schools and municipalities, accountability becomes a central issue at various layers of the administration. Moreover, I argue that implementation of traditional assessment and evaluation technologies is not necessarily good for improving the quality of education. Indeed, they may jeopardise the progress occurring elsewhere. My final conclusion is that the quality assurance practises in education at the system level are still rather undeveloped and, hence the models available may not be useful in satisfying the needs and requirements of any other national education system.

1. Quality assurance in the decentralised education systems

Education has always been given a high priority when we encounter life situations that are not what we used to have earlier. At the dawn of industrialisation, schools were called upon to produce qualified workers, managers and servants, during the cold war education was seen as a vehicle to promote peace, and in the late 20th century many eyes turned to teachers to raise the environmental awareness among the younger generation in order to save our planet. The purpose of general education has been challenged in a big way during the era of unpredictable political and economical reforms, and technological revolutions, and true globalisation of our lives in the past decade or so. At the start of the new millennium,

education in general and schools in particular are the key partners in the dance of change in our societies. Despite its rather low status in the political agendas of many national governments, education and learning have a key role to play in the cultural evolution of peoples.

Education is one of the few functions of societies that has been, and will continue to remain a nationally independent issue within the European Union and other economic and political alliances. The national values and traditions together with the cultural and economic characteristics create the framework for our education systems. For example, it is very unlikely that the ongoing unification of European countries would lead to a "European Curriculum", or "European Examinations" respectively. However, there are some crosscutting themes or issues that are common to all or most of Western European countries. Such grand ideas as lifelong learning, technological literacy for all, inter-cultural tolerance, and environmental education are found in any national education agendas and curricula.

In most countries it is assumed that students are prepared for quite different societies than those their parents and grandparents have experienced. We call this new era an information society or knowledge society. I would like to use the name *learning society for our time* that is so strongly characterised by the features of learning and change. Whatever name we wish to use for this time, we are witnessing the paradigm shift concerning the conceptions of knowledge and learning. Whereas in our past we believed that an educated and civilised person is the one who is able to possess and recall large quantities of knowledge, we now think that it is far more valuable to be able to process information and construct new knowledge. More specifically, such skills as gathering and analysing data, applying knowledge in new situations, solving problems, and regulating our own intellectual processes are in the centre of high quality education and learning. When the qualitative aspect of learning is emphasised as it seems to be at least in the scientific discourse and political rhetoric, it has fundamental consequences not only to curricula and the ways in which schools are organised, but also to what the pupils are expected to learn in school and how these desired outcomes are about to be assessed.

Delegation of the authority from the Ministries and central education agencies to schools and local education boards seems to be a global phenomenon. Naturally, one of the reasons for this shift of power is the belief that when decision-making takes place near the actual operations, the system is able to react to external changes faster than more cen-

tralised structures. However, transferring power from the top to the bottom makes only little sense if the resources do not follow the power. And when the financial authority and control of money is moved away from those who are in charge of spending it, a need for better monitoring of results will emerge. In brief, decentralisation of education means that schools have more power to influence in their own lives. In most cases this also means that schools have their own budgets for running their planned operations. It has become clear that this new freedom that schools have experienced during the era of self-management is followed by a new kind of accountability and responsibility.

Ultimately, the main intention of the recent legislative changes is to improve the quality of schools. Since the quality of education is not directly the topic of this paper, it is not appropriate to go into very detailed discussion this theme here. However, it is useful to bare in mind two aspects concerning the quality of schools. First, it is necessary to promote discussion on educational quality in order to understand better how to measure and improve the quality of schools. A well-known OECD report entitled *Schools and Quality* (1989) concludes, "the assessment of quality is thus complex and value laden. There is no simple uni-dimensional measure of quality. In the same way as the definition of what constitutes high quality in education is multi-dimensional, so there is no simple prescription of the ingredients necessary to achieve high quality education." Recently, school effectiveness research has clarified the role of various factors influencing school quality. Second, as Sammons (1999) argues, the variability of school quality seems to increase. The additional range of powers that is falling upon schools in general and principals in particular will increase school variability substantially. Moreover, schools will differ in their ability to cope with rapid externally induced changes. In Britain, for example, the Government's intention is that the major mechanism of quality control will be locally determined market mechanisms of parental choice. In Finland parental choice is becoming an important factor in schools profiles. However, due to the non-existing national examination and assessment system, our schools' reputation and quality image are based on informal messages and observations rather than hard evidence from examinations, tests and ranking tables. Our decentralised education system is in search of suitable quality assurance mechanism. But first it is necessary to understand what is meant by quality of schools and how to maintain the values of democracy and individual autonomy when the intended quality is measured.

2. Democracy and accountability in education

The national curricula in Scandinavia place a strong emphasis on the fostering of democratic values and on social development, in addition to the acquisition of knowledge and developing skills. The conceptual framework of assessment and evaluation is influenced by these characteristics. For example, school evaluation is generally seen as a tool for school improvement rather than external mechanisms of control or comparing one school with others (Lander & Ekholm 1998). Therefore, evaluation which embraces both single evaluation studies of selected areas and the monitoring of work done by students or teachers, is also adjusted to promote democracy within the education system, and emphasise the accountability of schools and the responsibilities of policy-makers in the overall improvement of education.

Scandinavian countries recognise education as a key system in developing democratic societies, social welfare and economic prosperity. Furthermore, it is widely accepted that democracy and related values cannot be directly taught to young people but they have to be able to experience their own lives in schools in democratic settings. The education policies of Scandinavian countries have, since the 1950's, emphasised the teachers, parents and students role in decision-making concerning how schools are operating. Local communities have received a range of new authority and teachers may have a considerable influence in curriculum decisions. Parental involvement in designing daily activities in schools has been common in most areas. In brief, one of the leading principles of developing education in Finland has been the promotion of participation of schools and parents in improving education. At the same token, the new legislation also brings to communities and schools the responsibility to evaluate their own work, its effectiveness and efficiency.

Democracy in education means that accountability and division of responsibilities have to be made clear for all parties. Evaluation and assessment become tools for management and improvement rather than those of control and inspection. Perhaps the most challenging aspect of this new management culture in schools is the lack of preparation and training of teachers and principals. The new law on education in Finland stipulates that the municipalities that are the owners of schools have to evaluate the quality of their education services and provide the State authorities with information about it. However, the focus and the structure of these local school evaluations are not fixed which leaves flexibility for local evaluators, but also increases the variability of information concerning the quality of schools.

Evaluation and assessment have been hot topics in Finnish education during the last decade. On the one hand, there is some pressure to introduce more objective assessment and examination systems to monitor the quality of schools in general and the condition of the national education in general. Probably the most frequently mentioned problem that exists today in our school system in terms of assessment is the lack of criterion-based national assessment at the end of nine-year comprehensive school. Hence, the school leaving diplomas that are based on teacher-made tests and marked by schools alone, are not able to indicate the real achievement of pupils. Nevertheless, these diplomas are the sole criteria for entering secondary education. On the other hand, however, the majority of educators, researchers and policy-makers believe that a national assessment and examination system in compulsory education would jeopardise the ideal of democratic education by shifting the focus of teaching from real life to the tests, and thus, moving the purpose of learning from understanding to memorising.

The success of democratic governance is dependent on how much there is trust in what people are willing and able to do. How well they do what they are suppose to do depends among other thing on the clarity and relevance of the achievement standards. In most cases people are able to produce good results if there is consensus and understanding of what are the expectations. In education, however, it is not always clear what pupils should know and be able to do after schooling. Improving the quality of education is not an easy task. One of the preconditions for better quality of teaching and learning is a clear understanding of what pupils should learn. Another issue is the role of assessment in improving classroom learning.

3. Can assessment improve classroom learning?

Today many people seem to believe that there is a direct link between assessment of pupils and improvement of quality of learning. Assessment is such a common theme in the world of education that the purposes of assessment have sometimes been neglected. Whenever national education systems are restructured, the goals of assessment systems have to be considered. Asp (2000) identifies four main features in today's assessment efforts.

Accountability. Evaluating the effectiveness of educational systems and programs.

Feedback. Providing specific feedback to individual teachers to inform instruction and to students to improve achievement.

Classification and certification. Putting students into categories in the name of efficiency and equity in instruction and certifying that students have achieved particular knowledge and skills.

Reform. Improving instruction on a broad scale and increasing student learning.

Ideally, assessment should serve several purposes simultaneously. Not any single assessment method is able to satisfy all the needs that were mentioned in the above features. There is a growing consensus among the education specialists that new assessments are needed to measure a broader range of abilities of pupils and to give teachers and schools better information about pupil progress and achievement (Darling-Hammond & Falk 1997). Therefore, it is not possible to find a comprehensive assessment technique that is able to serve all the various needs that it should serve. Instead, contemporary assessment systems will integrate several methods, including self-assessment of pupils and observation and inventories of students' work and learning done by their teachers.

Ensuring that students get the specific support they need requires detailed information about what students know and can do, as well as about how they learn. As our conceptions of pupil growth will be realigned to constructivist views of learning, the appropriate assessment methods become more versatile. Darling-Hammond and Falk (1997) argue that standardised tests relying on multiple-choice or short-answer formats do not adequately reflect students' performance or their abilities to solve problems in real-world situations, nor do they tap the variety of ways in which individuals think, learn, and demonstrate what they know. However, most of our teachers do not have knowledge about high-quality assessment practices. That is why it is so common among teachers to turn to the format of large-scale assessment regardless of its usefulness (Asp 2000).

It is needless to say that what is assessed is also valued in schools. Narrowly focused tests that have high stakes will obviously influence how teachers teach and, hence, what and how students learn. For example, the Minimum Competency Movement that employed multiple choice tests as the primary assessment technology in the United States in 1970's and 1980's, had a dramatic impact on classroom assessment in schools. As a consequence teachers started almost exclusively to use multiple-choice tests in assessing their students learning. This, in turn, led to a very narrow conception of educational quality and enhanced traditional teaching in classrooms.

Assessment and evaluation are interventions that always influence how the object of these interventions behaves. The traditional perspective on educational assessment is

based on the positivist view that aims at objectivity through independent data collection, control of variables, and interpretation of data using statistical methods. This has been particularly common in attempts to assess students learning in the national scale. If we accept that assessment will in any case have a steering effect on teaching and learning in school, we should also expect that high-quality assessment that emphasises learning as understanding of knowledge and development of skills and habits of mind will have a positive impact on quality of learning. This requires, therefore, that governments, their agencies, school authorities and the teaching profession study carefully their practices of formative assessment and look for ways to improve it. Black and William (1998) claim, on the basis of a solid review of research literature, that typical effect sizes of the formative assessment experiments were between 0,4 and 0,7. These effect sizes are larger than most of those found for educational innovations. Formative assessment here refers to all activities that teachers or students themselves undertake that provide information to be used as feedback to modify teaching and learning activities to meet the student needs.

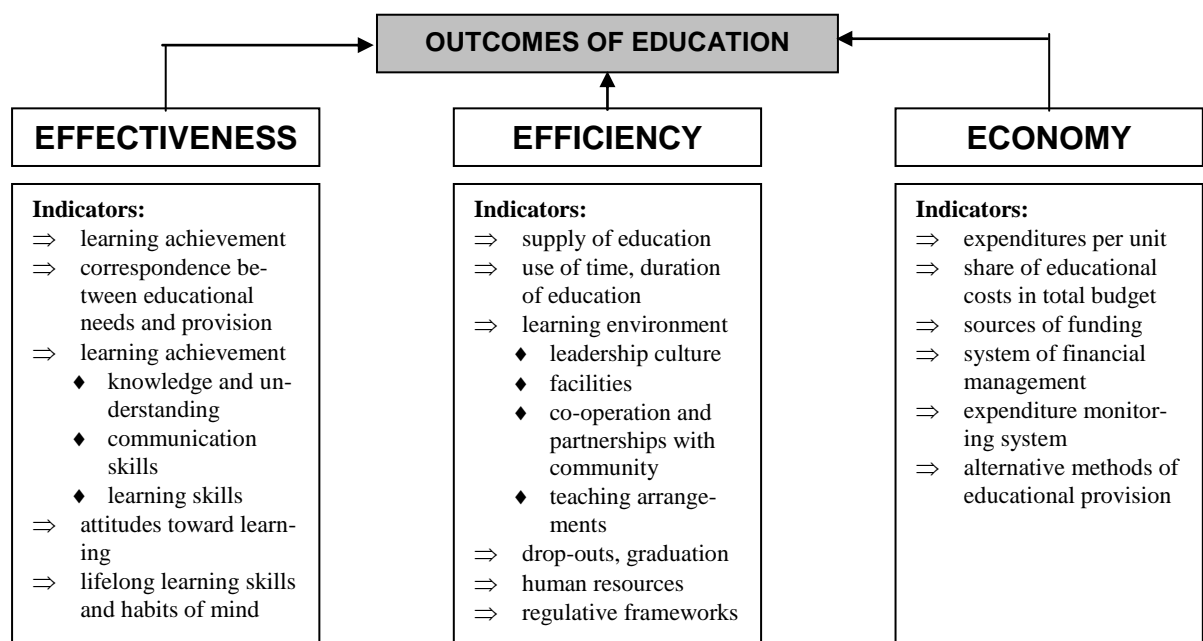
4. A conceptual framework for evaluating educational outcomes

Evaluation of educational outcomes focuses on three areas: comparative evaluation at national and international levels, evaluation of sub-sectors of an education system and educational institutions in relation to the quality of their educational services, and thematic evaluation of specific areas or functions. Comparative evaluation means that essential educational outcomes are compared with national or international findings, trends or objectives. Evaluation of sub-sectors of an education system may concern a type of educational institution or an education sector as a whole or the entire education system. Finally, thematic evaluations are focused on one or several dimensions of educational outcomes, or a certain educational theme which is under consideration.

This conceptual framework for types of evaluation was established in Finland in the mid-90's after the rise of the effectiveness movement and increased accountability at various levels of the system. At the same time three dimensions of educational evaluation were adopted as a basis for criteria of success. Education is successful when the objectives presented at national and international levels for each organisational stage of the education system, for each institution and for learning activities of each individual, have been achieved (NBE 1999). First, education is **efficient** when the functioning, flexibility and timing of education system, educational administration and teaching arrangements are as

appropriate as possible. Second, education is **economical** when the resources of education, in relation to educational objectives, have been allocated in an optimal way. Third, education is **effective** when the facilities produced by the education system promote, both qualitatively and quantitatively, the mental growth of an individual and the development of society, culture and working life. The first and second of the dimensions of educational evaluation are rather technically straightforward to handle in practice. However, the effectiveness dimension is the one that is challenging, complex and requires resources when done properly. Effectiveness of education in general and schools in particular consists of learning achievements of students and the correspondence between demand and supply of education. Determining how effective the education system is or how effectively schools operate requires multiple data collection, advanced assessment methods, and clear national goals and policies against which the actual achievements are measured.

Figure 1. An example of quality indicators of educational outcome evaluation



Educational effectiveness is evaluated and assessed in various ways in Finland. In principle, there are five educational phenomena that are measured: achievement of learning goals, development of communication skills, progress of learning skills, motivation to lifelong learning, and correspondence between need for and supply of education. At present, there is only one external national examination in general education in Finland that is both designed and marked by external examiners. This is a school leaving examination at the end

of general secondary school. No other national examinations do exist in our general education system.

One reason for the reluctance of educators and authorities to extend the national assessment of student outcomes to external examinations in compulsory education is the recognition of the need for a higher degree of flexibility within the education system. Many fear that external examinations and comprehensive assessment of learning outcomes may jeopardise the pedagogical intentions of schools towards more open and creative learning environments. Evaluation and assessment have been traditionally be seen as tools for school improvement rather than those of control and inspection. During the last decade or so the main focus of evaluation and assessment work has been on developing more accurate assessment methods that would be coherent with the education policy and the spirit of new national curriculum framework.

How do we know whether the education system is working successfully? First, the various types of evaluation are launched annually to describe the features and characteristics of selected areas of education. Recent areas of national evaluation have been special education, teaching and learning of foreign languages, and the quality of general secondary education. Second, the National Board of Education is responsible for monitoring the levels of achievement in mathematics, mother tongue and science of grade six and nine students. Assessment is based on national samples and is not able to provide any information for school-to-school comparison. Instead, the findings of these assessments are disseminated and used for curriculum development, refocusing of teacher in-service training and overall school improvement.

An education system has to be flexible in order to be able to respond to external changes and emerging needs of individuals and society. Evaluation of educational quality must not have a negative impact on the flexibility of the system. In Finland self-evaluation that schools do about their own work and self-assessment of students are becoming an integral part of the overall quality assurance in education. The increased role of the stakeholders in assessing the quality requires higher professional expertise and moral than before. Therefore, it is only possible to achieve new evaluation and assessment cultures with the assistance of increased staff development and school improvement efforts.

Finally, there are three issues that have emerged when the quality aspect of teaching and learning has become more important. First, evaluation of education in general and assessment of students learning outcomes in particular should be based on a clear consensus

of what is expected from schools and students in terms of performance and learning. Too general, and sometimes even controversial, learning or performance standards do not provide an adequate basis for measuring the quality. For example, if the curriculum does not clearly indicate what all students in any given stage should know and be able to do, assessment criteria are missing and the tests and textbooks become a curriculum. Second, the ultimate user of assessment information that is elicited in order to improve learning is the pupil. One of the major challenges in our schools is how to report students progress to themselves and their parents that would promote better performance in school and lead to higher learning of all students. Third, there is a need for considerable improvement in assessment technology and methodology in our schools. Such techniques as computer-adaptive testing, large-scale testing, classroom assessment and performance assessment may be combined into multiple method that will open up new ways of assessing complex cognitive activities (see Merzano et al. 1993). Furthermore, there is a need to develop specific value-added approaches to track more precisely student achievement. Value-added methodology measures the growth in individual student achievement from one point in time to another (Asp 2000).

I said earlier that quality is an complex issue in education. Despite rapid development of educational evaluation recently and increased emphasis on quality issues in almost every strand of life, there is not a single view of high quality education system. Moreover, it is fairly widely accepted that good education is closely interconnected with its cultural, political, social and economical contexts. Therefore, each education system that wishes to establish a quality assurance system within its national education structures, has to start this work by analysing various aspects of goodness in education and define its own idea of quality. The example of Finnish education shows that when the education system is based on such values as equal opportunities, equity, freedom of choice and cultural diversity, the conceptions of quality in education becomes versatile. In such case the quality of education should consist of elements that can be constructed from these values of the education system. As a result, in the Finnish educational evaluation the process of education in general, and the nature of teaching and learning in particular have a central position in determining whether education is of high quality or not. It is, hence, necessary that some attention is paid to clarifying the concept of quality and how it may be measured before any operational decision are made concerning establishing assessment or other quality assurance structures.

5. Two roads to quality assurance: curriculum and assessment

Every education specialist has some concern about how well schools are doing and what is the state of education in general in his or her country. As I wrote earlier there are no two alike solutions to the problem of raising educational standards or assuring the quality of education when various school systems are compared with each other. In principle, however, we may identify two broad avenues along which education policies and related practices are approaching higher performances and better success within their selected frame of criteria. These interrelated mechanisms are curriculum that provides criteria for teaching and learning respectively, and assessment that, in turn, specifies the target of learning and teaching respectively. Although these two means of raising the quality of education share several common elements, their functions and impact on daily work in schools, particularly that of teachers and pupils, may be quite different. Let us discuss some of the characteristics of these two in little bit more detail.

The basic idea of assessment-driven quality assurance is that regular monitoring of what pupils have learnt, how teachers teach and broadly speaking how schools work, will itself make each of these perform more intensively and pay attention to the intended achievements. In such environment the assessment criteria will set the requirements for teaching and provide the targets for learning. In some countries specific assessment and evaluation frameworks may guide the planning of teaching in such a way that it also serves the role of kind of curriculum. In fact, the conception of curriculum will be quite different in such setting since the factors other than those in curriculum will determine the planning of teaching and learning. It seems clear that such system will be more quality-proof since it sets more specific and objective expectations for the learning outcomes of pupils. On the other hand, the entire teaching-learning process will probably be driven by results that may in some cases, with some individuals jeopardise the openness of learning and growth within the social environments of school.

Curriculum-based education system, in turn, assumes that the quality of teaching and learning may be maintained and improved through qualitative and quantitative description of the process of education. The idea is that specifying the teacher behaviour, i.e. the content that she should teach and the activities that she should conduct with pupils, will lead to expected outcomes or something near them. The frame of reference for teaching is, thus, the curriculum, not so much the attainment targets. There is, however, considerable

differences in terms of flexibility and level of specification of the objectives, content and methods in various curriculum systems.

One possible solution for higher quality education might be the merging of the assessment-driven and curriculum-based approaches described above. This requires new types of thinking of both pupil assessment and role of curriculum in education. According to many scholars in the field performance assessment seems to be one potential practical way to integrate the purposes of assessment and learning (Glatthorn 1998). However, the practices and related technologies of such authentic assessment systems are still at the early stage of development and considerable amount of efforts are needed in order to have reliable and precise means available for all schools. The future is, nevertheless, promising since some interesting perspectives have been envisaged by many of us that may help teachers on the road to better quality schools.

6. Perspectives for the decade

I have said in this paper that evaluation and assessment have become the major terms in educational policies as well as in academic discourse. There is not a single purpose that assessment serves, nor is there one single methodology that could do it all. As Asp (2000) has put it: “In the era of accountability, where assessment defines what is important, we must align the targets. This alignment can only happen by eliminating some of them, linking others together, and clarifying the appropriate uses of particular assessments.” The key to high quality assessment in schools is to link large-scale testing and classroom assessment together. The traditional conceptions and methods of assessing student learning and school performance are not able to promote improvement in the classrooms and in the staff rooms. In addition, assessment at the classroom, school, district and state levels needs to be linked together to support and reinforce a set of clearly articulated goals.

The field of assessment is developing rapidly within the scientific community and practitioners as well. The evolution of quality management in education may vary from one system to another but it is easy to predict that there will be some common lines of development in different countries. I close this paper by forecasting some possible future trends of school assessment (see Asp 2000 for more).

1. Norm-referenced testing will decrease as accountability focuses more on what students actually know and are able to do, rather than on how much they know about general content knowledge compared to other students’.

2. Criteria for evaluating the quality of assessments and assessment systems will include the influence of assessment on learning, teaching, and school as an organisation, along with psychometric characteristics.
3. The degree to which an assessment is sensitive to instruction will become a major indicator of quality.
4. Performance assessment will play a more prominent role in both classroom and large-scale assessment as students' ability to apply knowledge in real-life situations becomes more valued as a goal of education and we are better able to assess that ability. As a consequence, self-assessment and other reflective practices will become accepted as elements of assessment methodology.
5. A new role for assessment will emerge that emphasises helping to provide more and better education for the learner. Simultaneously, assessment will become prominent over other goals such as accountability or the classification of students.
6. Curriculum in general and knowledge and performance standards in particular will play a more integrated role in the overall quality management of education. The curriculum will describe in a more specific way what students should know and be able to do while in school.

What will actually happen in assessment and evaluation remains to be seen. One thing we know, however, without a doubt: assessment will develop and change as the years go by. So will curriculum. One of the turning points of both these aspects of education will be the moment when the majority of educators believe that teaching is not talking and that learning is not listening (Sahlberg 1999). Interesting question is: Will our pupils and student realise it before us? If that happens, then we have to do something that we are not perhaps ready yet. My message is: Let us work on it right now.

References

- Asp, E. 2000. Assessment in education: Where have we been? Where are we headed? In: Brandt, R. (Ed.) Education in a new era. Alexandria: Association for Supervision and Curriculum Development.
- Black, P. & William, D. 1998. Inside the black box. Raising standards through classroom assessment. Phi Delta Kappan, October.
- Darling-Hammond, L. & Falk, B. 1997. Using standards and assessment to support student learning. Phi Delta Kappan, November.

- Glatthorn, A. 1998. Performance assessment and standards-based curriculum: The achievement cycle. Larchmont; NY: Eye On Education Inc.
- Hargreaves, A. & Evans, R. 1997. Teachers and educational reform. In: Hargreaves, A. & Evans, R. (Eds.) Beyond educational reform. Buckingham: Open University Press.
- Marzano, R., Pickering, D. & McTighe, J. 1993. Assessing student outcomes. Alexandria: Association for Superviaion and Curriculum Development.
- Lander, R. & Ekholm, M. 1998. School evaluation and improvement: A Scandinavian view. In: Hargreaves, A., Lieberman, A., Fullan, M. & Hopkins, D. (Eds.) International Handbook of Educational Change. Dordrecht: Kluwer Academic Publishers.
- NBE 1999. A framework for evaluating educational outcomes in Finland. Helsinki: Evaluation 8, National Board of Education.
- OECD 1989. Schools and quality. Paris: OECD.
- Open Society Institute 2000. Education in Montenegro. Needs assessment.
- Sahlberg, P. 1999. Building bridges for learning. Brussels: European Youth Forum.
- Sammons, P. 1999. School effectiveness. Coming of age in the twenty-first century. Lisse: Swets & Zeitlinger.
- Teddlie, C. & Reynolds, D. 2000. The international Handbook of school effectiveness research. London: Falmer Press.
- Whitford, B.L. & Jones, K. 1998. Assessment and accountability in Kentucky: How high stakes affects teaching and learning? In: Hargreaves, A., Lieberman, A., Fullan, M. & Hopkins, D. (Eds.) International Handbook of Educational Change. Dordrecht: Kluwer Academic Publishers.