

Chapter 2

Trends in evaluation and assessment

This chapter provides the context for analysing evaluation and assessment policy. First, it describes the main trends within educational evaluation. It is apparent that evaluation and assessment are increasingly being considered as levers of change guiding improvement, accountability, educational planning and policy development within school systems. Countries are developing more comprehensive evaluation and assessment frameworks, placing greater emphasis on educational measurement and indicators development, giving growing prominence to accountability uses of results, and relying increasingly on educational standards.

Second, the chapter reviews the contextual factors shaping the development of evaluation and assessment in school systems. Evaluation and assessment have gained in importance as a result of greater levels of school devolution, a stronger role for market-type mechanisms in education, the emergence of New Public Management, the growing imperative of an efficient use of public resources, the need to focus on “quality for all” and the rising importance of education in a global world. Other contextual factors influencing the development of evaluation and assessment frameworks include the rising expectations of the professionalism of teachers, more educated parents, the movement to advance the use of evidence-based decision making, technological advancements, the emergent commercial interests in education and the role of the media.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Introduction

This chapter provides the context for analysing evaluation and assessment policy. First, it describes the main trends within educational evaluation, with particular emphasis on the expansion and diversification of evaluation and assessment procedures. Second, it reviews the contextual factors shaping the development of evaluation and assessment in school systems. More specific trends within student assessment, teacher appraisal, school evaluation, school leader appraisal and system evaluation are analysed in the respective chapter.

Trends in evaluation and assessment

Expansion of educational evaluation in school systems

Increased prominence of evaluation and assessment in education policy

It is apparent that education policy is increasingly conferring a central strategic role to evaluation and assessment as indispensable tools for improvement, accountability, educational planning and policy development. In the last two decades, most countries have introduced a wide range of measures intended to improve evaluation and assessment at all levels from the student to the school system itself. These have done much to stimulate public awareness of evaluation and assessment and to develop an evaluation culture within school systems.

For example, Norwegian authorities set up a National Quality Assessment System (NKVS) for the education sector in 2004. NKVS provides access to a range of data intended to help schools, school owners and education authorities evaluate their performance and inform strategies for improvement. It includes national student assessments at key stages of education, a range of user surveys, a web-based School Portal, and a range of tools for schools' self-review (e.g. diagnostic "mapping tests"). With the establishment and development of NKVS, policy makers aimed to move policy attention away from inputs and processes to focus more on the outcomes of education (Nusche et al., 2011a). Similarly, Portugal has come far in developing the foundations of a framework for evaluation and assessment. National monitoring educational progress tests were launched in 2001 (then replaced by national examinations in 2011/12), a first cycle of external school evaluations was completed in the period 2006-11, a national system of teacher performance appraisal was launched in 2007 and the availability of national indicators on education has considerably expanded (Santiago et al., 2012a). In Mexico, as of the National Education Programme 2001-06, it was established that evaluation and assessment should be permanent and systematic, combining the involvement of internal and external agencies and be important management instruments to achieve improvement and accountability to society (Santiago et al., 2012b). In Australia, the 2008 National Education Agreement, which established a national framework for reform in education, reinforced the role of evaluation and assessment as key tools to achieve quality and equity in education, in particular with the introduction of the National Assessment Program – Literacy and Numeracy (NAPLAN) and the establishment of a set of reporting requirements for all schools (Santiago et al., 2011).

The expansion of educational evaluation results from increased demands for effectiveness, equity and quality in education so new economic and social needs are met. It is part of the effort across many countries to bring about rapid and sustained large-scale

educational reform with real gains in student outcomes (Campbell and Levin, 2009). Evaluation and assessment have also gained in importance as a result of pressures for governments to meet public accountability, ensure transparency within education systems and maintain public confidence in schooling.

Creation of dedicated agencies as part of new approaches to govern evaluation and assessment

In many OECD countries, the greater importance of evaluation and assessment in education policy has involved the creation of specifically dedicated agencies which assume a central role in the governance of the evaluation and assessment framework (see also Table 3.2 in Chapter 3). Denmark created a specific national authority to monitor compulsory education (the Quality and Supervision Agency, formerly the School Agency) in 2011 and an advisory body to evaluate priorities in compulsory education (The Council for Evaluation and Quality Development of Primary and Lower Secondary Education) in 2006 (Shewbridge et al., 2011a). In Mexico, a milestone in the development of evaluation and assessment was the creation of the National Institute for Educational Assessment and Evaluation (INEE) in 2002 by presidential decree as a public, decentralised agency to provide national guidance and direction in evaluation and assessment activities at the school level. It gained further autonomy in 2012 with the objective of reinforcing its technical leadership of evaluation and assessment and maintaining the independence of its judgement of the state of education in Mexico (Santiago et al., 2012b).

In Chile, the Quality of Education Agency was created in 2011 and started operating in 2012. It takes responsibility for evaluating the quality of learning provided by Chilean schools, including the evaluation of teachers, school leaders and school providers, in view of improving the quality and equality of education. The Agency evaluates individual schools against learning standards, makes information about the performance of individual schools publicly available, and supervises and supports schools with lower performance (Santiago et al., forthcoming). In the Flemish Community of Belgium, the Agency for Quality Assurance in Education and Training (AKOV) was established in 2009 to oversee all services related to quality improvement of education (Shewbridge et al., 2011b). In Italy, the creation in 2004 of the National Institute for the Evaluation of the Education System (INVALSI) led to the development of national student assessments since 2008. Similarly, in Ontario (Canada), the creation of the Education Quality and Accountability Office in 1996 was associated with the development of standardised student assessment to provide accountability and a gauge of quality in Ontario's publicly funded education system.

The creation of agencies dedicated to evaluation and assessment recognises the need for specialised expertise, the imperative of building adequate capacity to deliver evaluation and assessment policies and the necessity of introducing some independence *vis-à-vis* education authorities. Functions of evaluation and assessment agencies may include technical leadership (e.g. in developing evaluation instruments, guidelines, education indicators), implementation of evaluation and assessment procedures (e.g. national student assessments), the monitoring of the education system, the introduction of innovations on the basis of research results, the development of capacity for evaluation and assessment across the system, knowledge management (of results produced by evaluation and assessment activities) and the promotion of an evaluation culture.

Greater variety of evaluation and assessment activities

The expansion of educational evaluation was accompanied by considerable diversification of evaluation and assessment activities. Although educational evaluation within school systems is not a recent concern, it has traditionally focussed mostly on the assessment of students. As will be evident in subsequent chapters of this report, in recent years, countries are increasingly developing more comprehensive evaluation and assessment frameworks with more resources devoted to evaluation components other than student assessment.

For instance, within a context of growing levels of school autonomy, the responsibility for evaluation is increasingly given to the school itself, which involves greater emphasis of countries on school self-evaluation (see Chapter 6). At the same time, new accountability requirements for schools raise the importance of external school evaluation (see Chapter 6). In this context, some countries are now placing considerable emphasis on school leadership, developing reference standards for good leadership and establishing procedures to appraise school leaders (see Chapter 7). While less common in countries, there has also been a growing interest in developing formal teacher appraisal systems. Demands for instructional quality have led a number of countries to set up one form or another of teaching performance assessment (see Chapter 5). Some countries, such as the United States, are now using student standardised assessment results as an instrument to appraise individual teachers through the design of value-added models intended to measure the contribution of individual teachers to student learning (see Chapter 5). Also, as countries place greater emphasis on the monitoring of outcomes in public services, student standardised assessments play an increasingly important role in assessing learning outcomes in school systems (see below and Chapter 8). This is within a context in which countries demonstrate a growing interest in international benchmarks of student performance.

Another prominent development has been the growing importance of performance data, particularly relating to student outcomes, to inform school and classroom practices as well as system-level policies. This has generated a range of new practices in school systems related to the analysis of student outcomes. Strong emphasis is being placed on better equipping and encouraging teachers and other school agents to use data for formative assessment, on providing the incentives and means for student outcomes to be used in school self-evaluation, and on encouraging “value-added” approaches to assess the contribution of schools to student learning. Some countries have placed considerable emphasis on student information systems providing real-time access to student data on, for example, attendance, enrolment, marks and schedules. This is in a context in which education authorities have a growing concern of feeding back relevant information to school agents. The focus on student outcomes has also increasingly involved the establishment of longitudinal studies on the progression of individual students through the education system.

In the area of student assessment, in the last two decades there has been considerable policy attention to the consolidation of assessment for learning in the classroom and a growing support for the concept of *assessment as learning*, which focuses on students reflecting on and monitoring their own progress to inform future learning. In addition, technological advances have permitted student assessment to become more sophisticated, as is the case with computer-based adaptive assessment (see below).

The rise of educational measurement and indicators development

Student outcomes as the focal point for analysis

A major benefit of the stronger emphasis on evaluation and assessment has been the greater focus on improving student outcomes and achieving student learning objectives. This is reflected in the growing importance of student outcomes for system evaluation (increasingly relying on results of standardised student assessment and the international assessment of students), school evaluation (with school accountability increasingly tied to student outcomes) and teacher appraisal (with the exploration of direct links to student progress); the requirements for reporting publicly on student results; and the establishment of education national targets for student achievement including for particular groups of students. Performance in schools is increasingly judged on the basis of effective student learning outcomes. This is part of the general shift to outcome measures in the public sector. There is a greater emphasis on the use of student achievement data both to understand the balance between school, student and contextual data and to look at the school processes that appear to support improved achievements (Campbell and Levin, 2009).

The growing emphasis on measuring student outcomes

The introduction of national standardised assessments for students in a large number of countries reflects the stronger focus on measuring student outcomes. These make data on student learning outcomes available, providing a picture of the extent to which student learning objectives are being achieved, and they grant the opportunity to compare student learning outcomes across individual schools, regions of the country and over time. As put by Kellaghan and Greaney (2001), “The most remarkable development in assessment towards the end of the 20th century has probably been the growth in its use to measure the achievement outcomes of national systems of education, either considered uniquely (in national assessments) or in the context of the performance of other education systems (in international comparative studies of achievement).”

In Mexico, a ground-breaking development was the implementation of national standardised assessments: on the basis of a sample (EXCALE in 2005) and census-based (ENLACE in 2006). These made available data on student learning outcomes which, for the first time, provided a picture of the extent to which student learning objectives were being achieved. ENLACE also granted the opportunity to compare student learning outcomes across individual schools (Santiago et al., 2012b). As with Mexico, many OECD systems introduced, for the first time, central standardised assessments in core subjects in recent years, reflecting an impressive expansion of instruments to measure student outcomes. Examples include: Austria (2012), the Flemish Community of Belgium (2002), the French Community of Belgium (2009), Denmark (2009), Germany (2007), Hungary (2001), Iceland (2009), Ireland (2007), Israel (2002), Italy (2008), Japan (2007), Korea (2001), Luxembourg (2008), Norway (2004), Portugal (2001), Spain (2007) and the Slovak Republic (2004). In addition, the Czech Republic is piloting student standardised assessments in academic years 2011/12 and 2012/13 (see also Chapter 4; OECD, 2011; and Eurydice, 2009). Australia has also introduced standardised student assessment at the national level in 2008 with the National Assessment Program – Literacy and Numeracy (NAPLAN), even if centrally organised student assessments were in existence in several states and territories previously (Santiago et al., 2011). Similarly, Canada introduced standardised student assessment at the national level in 1996 through

the Student Achievement Indicators Program (SAIP), which was replaced by the Pan-Canadian Assessment Program as of 2007.

Other countries have had more extensive experience with national standardised student assessment. The United States organised the first National Assessment of Educational Progress (NAEP) in 1969. NAEP is a sample-based assessment whose results are designed to provide data on student achievement in various subjects and released as The Nation's Report Card. It may be considered the precursor to today's widespread use of student assessments as tools for holding educators accountable for student performance (Roeber, 1988, cited in Hamilton, 2003). The minimum competency testing movement of the 1970s emphasised the need to ensure that students demonstrated a grasp of basic skills and led to the first formal use of tests as tools to hold education administrators, students and teachers accountable for performance (Hamilton and Koretz, 2002). Over time there was a shift from the use of tests as measurement instruments designed to produce information to a reliance on tests to influence policy and instruction, a dual use that has continued to the present day (Hamilton, 2003). Other countries with an early experience with national student assessments include Ireland (national assessments in English reading and mathematics introduced to primary schools in 1972 and 1977 respectively), the Netherlands (standardised assessment in the majority of primary schools since 1970) and Sweden (formative national assessments in lower secondary education introduced in 1962) (Eurydice, 2009).

The proliferation of education indicators

For the purpose of monitoring education systems and evaluating school performance, data are increasingly complemented by a wide range of education indicators based on demographic, administrative and contextual data collected from individual schools. Datasets typically include information on students (type of enrolment, completion, absenteeism, age, gender, marks, socio-economic background), teachers (functions, qualifications, career status, age, gender, areas taught, teaching hours, absenteeism, remuneration), non-teaching staff (qualifications, age, gender, category), and schools (financial management, use of technology, organisation of learning). The emphasis is increasingly on output measures.

Most countries have developed comprehensive national indicator frameworks relying on data collection procedures at the school level. It is now also common practice to report statistics and indicators in education in an annual publication (e.g. *Statistical Yearbook of Education* and *Education in Figures*). In the Flemish Community of Belgium, for example, the Agency for Educational Services (AGODI, created in 2006) collects and analyses data on the state of education. In addition to the most conventional data collected at the school level (e.g. characteristics of teachers and students), it covers areas such as the operational resources and subsidies granted to schools, the re-structuring of schools and the labour market for teachers (Flemish Ministry of Education and Training and the University of Antwerp Edubron Research Group, 2010).

International benchmarking is also increasingly common. A major driver for the collection of information on national education systems has been the joint international standardised data collection by UNESCO, OECD and EUROSTAT. It is also common to have countries publish education indicators in an international perspective. In Italy, the VALSIS project involves the analysis of international education system indicators to inform the Italian system evaluation framework. One output of the project is the creation of an electronic data bank with education system indicators.

Larger and more varied uses of evaluation and assessment results

Countries are giving a more varied use to evaluation and assessment results, including as a tool for understanding better how well students are learning, for providing information to parents and society at large about educational performance and for improving school and teaching practices. As will be analysed below, an increasingly marked focus is the use of evaluation and assessment results to hold policy makers, school leaders and teachers accountable. There is also a growing use of evaluation as a system steering tool. As put by Broadfoot and Black (2004), “In recent years the importance of assessment as a policy tool has grown enormously as governments have increasingly come to realise its powerful potential as a mechanism of state control.” Evaluation procedures are now increasingly being considered as potential levers of change that can assist with decision making, resource allocation or school improvement.

There is a growing interest in using evaluation results for formative purposes. School leaders, teachers and policy makers are more and more using evaluation results to identify areas where schools are performing well, and where they may need to improve. These data may help shape policy and/or school management decisions on resource distribution, curriculum development and definition of standards, or strategies for professional development. School leaders and teachers can use evaluation data to change teaching, address ineffective programmes in their schools, and improve the functioning of the school in terms of increased student achievement (Schildkamp et al., 2012). There have also been profound changes in views on the role of classroom-based assessment. Assessment has traditionally been viewed as a tool for making summative judgements of student achievements. But increasingly, assessment is also seen as a tool for learning. Assessment, in this view, plays a “formative” role – allowing teachers to identify gaps in student learning and to adapt teaching appropriately (Looney, 2009).

The data generated by evaluation and assessment procedures is also increasingly motivating schools to engage in the corresponding analysis in view of improving student learning. Policy makers, school leaders and teachers are putting time and energy into making data accessible to schools and teachers and into using data, in training schools and teachers to use data for improvement, and in requiring schools and teachers to collect and publish data (Schildkamp et al., 2012). Greater reliance on self-evaluation by schools intensifies pressure for teachers and school leaders to possess the skills and know-how to collect and use information needed for diagnosis and performance measurement.

The growing prominence of accountability as a purpose of evaluation and assessment

Countries are increasingly using evaluation and assessment for accountability purposes. A central assumption in accountability is that substantial improvement necessitates that the school agents are held accountable for the outcomes they generate. By measuring student outcomes and holding teachers, schools and policy makers responsible for results, accountability systems intend to create incentives for improved performance and identify underperformance within school systems. Hargreaves and Shirley (2009), in their analysis of educational reforms, maintain that countries have gradually shifted from local and sampled assessments to high-stakes census testing for accountability purposes. Klenowski and Wyatt-Smith (2012) point to the political appeal of school-based accountability policies, in the sense of the clear need for politicians to be seen to deliver improved outcomes in education. Broadfoot and Black (2004) note, for example, that “decisions about assessment procedures – particularly those concerning

high-stakes testing of various kinds – are as often based on perceived political appeal as they are on a systematic knowledge on the scientific evidence concerning fitness for purpose” (as cited in Klenowski and Wyatt-Smith, 2012).

The school-based accountability movement emerged out of a desire, particularly seen in the United States and the United Kingdom beginning in the 1980s in the Reagan and Thatcher eras, to measure performance in the public and non-profit sectors (Figlio and Kenny, 2009). Elmore (2004) argues that test-based accountability has been more enduring in education than any other policy in the United States for at least the past 50 years and that it is unlikely to change in the near future. According to Hamilton (2003), the policy context in the United States is characterised by the use of tests in what may be called a test-based accountability system. These systems involve four major elements: *goals*, expressed in the form of standards; *measures* of performance (i.e. tests); *targets* for performance; and *consequences* attached to schools’ success or failure at meeting the targets. As described by Figlio (2006), in the United States, the No Child Left Behind (NCLB) Act of 2001 solidified a national trend toward increased student testing for the purpose of evaluating public schools. Under NCLB, states must develop and administer rigorous curriculum-based standardised assessments to every student in a number of year levels. These tests must be used to evaluate schools, and in the case of schools receiving federal aid for disadvantaged students, aggregate performance on these assessments is associated with substantial rewards and sanctions. In 2009, the federal government of the United States launched the Race to the Top (RTT) initiative as a competition among states to access substantial federal funding. The RTT provides funding for states which implement a range of policies such as performance-based rewards for teachers and school leaders, adoption of common nationwide standards (from the Common Core State Standards initiative), and development of high-quality standardised student assessments.

As another example, in Ireland, there has been a drive for greater accountability in recent years. Examples of the move towards accountability include the introduction of regular whole-school inspection to secondary schools in 2003, the publication of school inspection reports in 2006, and the introduction of mandatory standardised testing in primary schools in 2007. The National Strategy for Literacy and Numeracy outlines additional accountability measures such as the development of national standards of students’ achievement and the collection of national data on student achievement. The strategy also requires schools to provide parents with adequate, meaningful and clear assessment information on their child’s progress (Irish Department of Education and Skills, 2012).

The accountability uses of evaluation and assessment can take a variety of forms. First, evaluation and assessment exist in an environment where there is a growing trend of public reporting (see also Chapter 6). This consists, for example, in publishing standardised student assessment results at the school level for use by parents, government officials, the media and other stakeholders. Not only does this serve the purpose of providing information on education system performance to the general public, but the results are often used by stakeholders to take action as with school parental choice. For example, in Chile national student assessment results (System for Measuring the Quality of Education, SIMCE) are published, inform the school voucher system and have contributed to placing education on the public agenda. In Australia, NAPLAN (National Assessment Program – Literacy and Numeracy) results are published on an individual school basis on the My School website, where the public can access performance and other data on schools across Australia. English schools’ performance has also been

reported since 1988. The general move towards greater transparency of results includes the publication of school inspection reports, school annual reports and system level reports providing an assessment of the state of education.

Second, evaluation and assessment results are increasingly used to reward or sanction the performance of individual school agents. This goes alongside the expansion of school external evaluation and teacher appraisal procedures. A number of countries have now instituted systems whereby either schools or teachers receive rewards for their good performance (e.g. in the form of financial one-off packages, additional resources including opportunities for professional development, faster career advancement, opportunities for promotion), or are the subject of sanctions for underperformance (e.g. school shutdown, career stagnation, removal from post). The emergence of value-added techniques has also strengthened the interest in the use of student assessment results for evaluating and rewarding individual teachers and school leaders (Hout and Elliott, 2011).

Another development is the greater variety of accountabilities for school agents. For instance, in New Zealand, schools have multiple accountabilities – to their communities, the Ministry of Education, the Education Review Office, the New Zealand Teaching Council and the New Zealand Qualifications Authority (New Zealand Ministry of Education, 2010).

Greater reliance on educational standards

The focus on student learning outcomes has, in many countries, driven the establishment or underlined the importance of educational standards for the quality of the work of schools and school agents, and encouraged means for monitoring progress towards those standards. Educational standards refer to descriptions of what students should know (content standards) and be able to do (performance standards) at different stages of the learning process. In many countries, there is growing emphasis on the development and use of ambitious educational standards as the basis of assessment and accountability. By creating a set of standards against which student performance can be measured, countries aim to assess students against a desired measurable outcome. Examples of countries which implemented national educational standards are Australia, Austria, Belgium, Denmark, France, Luxembourg, Norway, the United Kingdom and the United States. By setting national or common standards, student outcomes can be more easily controlled for quality and they are more comparable (Wang et al., 2006). The movement towards comparing student outcomes to standards also has had a role in motivating countries to administer national standardised assessments.

At the same time, countries have adapted the key elements of standards-based systems to their own educational contexts and cultures – how they define standards, how they balance incentives and support, and how they measure school and student performance. Educational standards vary a good deal in specificity and emphasis across countries (see Chapter 4).

Internationalisation of assessment

National education debates are increasingly shaped by international comparisons, particularly of student performance in international student surveys (see also Chapter 8). These include student assessments conducted by the International Association for the Evaluation of Educational Achievement (IEA) (e.g. Progress in Reading Literacy Skills survey, PIRLS; Trends in Mathematics and Science Skills survey, TIMSS), the OECD's

Programme for International Student Assessment (PISA) or UNESCO's Latin American Laboratory for Assessment of the Quality of Education (LLECE). As explained by Bonnet (2004), "The conception of educational evaluation has changed over the years, moving from an essentially national to a more international perspective. It has also taken a new dimension with the fundamental role it now plays in Europe. In parallel with national developments countries started to show interest in comparative evaluation at the international level." Some education systems have been considerably shaken after publication of international comparative scores. For example, the first results of PISA published in December 2001 jolted Denmark and its education community. They provided evidence that one of the most expensive education systems in the world was performing at a level that, when compared to the outcomes observed in other OECD countries, was only average. The PISA results were at odds with the widely shared but poorly substantiated belief that Danish schools were the best in the world (OECD, 2008).

The growing availability of internationally comparable data on student performance has, in important ways, influenced national discussions about education and fostered education policy reforms in countries. International comparative data put countries under pressure to attain higher levels of performance building on policies identified as potentially effective in high-performing countries. A wide range of education reforms are triggered in OECD countries by student results in international assessments. For instance, in France, an application decree links the 2005 introduction of the common core competencies to the results of French 15-year-olds in PISA (Dos Santos and Rakocevic, 2012). Some countries go as far as setting education targets based on international assessments. For instance, Mexico established as an educational target in its 2007-12 Education Sector Programme, a combined score of 435 in the reading and mathematics PISA tests to be attained by 2012 (SEP and INEE, 2012). Broadfoot and Black (2004) note how assessment has become an international field: "Not only are new assessment policies and practices rapidly exported around the world, an increasing volume of assessment activity is explicitly international in being designed to compare national indicators and performance." They highlight three key themes: first, the increasing willingness of researchers to acknowledge the impact of context in the operation of particular assessment practices; second, the global scale and impact of assessment policy and practice; and third, the development of international surveys of learning and achievement.

The expansion of international assessment has also significantly contributed for some countries to introduce national standardised assessments. This was the case, for example, in Denmark, Italy, Mexico and Portugal, where there previously had been little emphasis on the measurement of student outcomes. In these countries, measured standardised student outcomes were only available through international assessments and it was deemed necessary to develop measures aligned with national student learning objectives. In some countries, national assessments were developed on the basis of methodology used by international assessments, particularly the assessment framework proposed by PISA. This has also translated into the exchange of expertise on the measurement of learning outcomes across countries.

Greater technological sophistication

The expansion of assessment, particularly the spreading out of standardised student assessment, as well as the management of the data it generates has greatly benefited from greater capacity of information and communication technologies. Improvements include more individualised assessment approaches, better assessment of cognitive skills such as

problem solving, capacity for rapidly marking large-scale assessments, reliability in marking and reduced cost to administer student assessment. For instance, in Denmark, computer-based national tests officially implemented in 2010 are adaptive in that the items are tailored to students' latent ability levels. Test items are selected sequentially according to a student's performance on the previous test items. These efficient national tests provide rapid feedback of test results to teachers the next day, which can greatly facilitate teachers' use of the test results (Danish Ministry of Education and Rambøll, 2012). In Norway, as of 2008, all primary and secondary schools were using an electronic test administration system and an electronic test execution system. It became possible to give ICT-based examinations for anyone interested, including for students sitting the national standardised tests (Norwegian Directorate for Education and Training, 2011). Other examples include the development of rapid-assessment – a computer-facilitated approach to frequent, brief formative student assessment, more sophisticated value-added models to determine a school's or a teacher's contribution to student learning, and data information systems providing new opportunities for information sharing across school agents.

Contextual developments shaping evaluation and assessment

Changing modes of school governance

Greater decentralisation and school autonomy

There is an increased prominence of evaluation and assessment as school systems decentralise with further autonomy given to intermediate levels of government (e.g. regions, municipalities) and to individual schools. There has been a general international trend towards devolution of responsibilities for budget management, staffing, educational provision, teaching content and processes, and the organisation of learning to the local level including schools. This increased autonomy has been balanced by the strengthening of accountability requirements for local education authorities and schools. This goes alongside the increasing role of central authorities in areas such as strategic steering, standard setting, support and capacity development. According to Eurydice (2007), in Europe, the expansion of school autonomy occurred mostly in the 1990s (e.g. Nordic countries, Austria, Czech Republic, Estonia, Hungary, Italy, Poland, Slovak Republic, Scotland) even if some countries pioneered it prior to the 1990s (e.g. Belgium, England, France, the Netherlands, Northern Ireland, Spain, Wales) and expansion continued in the 2000s in a more limited extent in some countries (e.g. Germany, Luxembourg, Portugal). However, it should be noted that the concepts of school autonomy differ considerably across countries. According to OECD (2012), in lower secondary education, the percentage of decisions taken at either the local or the school level exceeds 75% in the following systems: Chile, the Czech Republic, Denmark, England, Estonia, Finland, Hungary,¹ Iceland, the Netherlands, New Zealand, Norway, Scotland, Sweden and the United States.

In a context of school autonomy, greater policy attention is given to areas such as school leadership, capacity for schools to self-manage (including self-evaluation and the monitoring of the quality of teaching and learning) and ability to implement improvement processes. In addition, the greater responsibilities assumed by schools imply greater accountability requirements such as external school evaluation and public reporting of student performance.

Stronger role for market-type mechanisms

Another major trend in some countries is the growing use of market-type mechanisms in education to generate efficiencies and improve the quality of education systems. A market mechanism in education is an instrument that facilitates the co-ordination between the demand for and the supply of education services. The rationale for the introduction of market mechanisms is the expectation they will generate better outcomes in education systems than traditional regulatory instruments. Examples of market mechanisms are parental choice of schools (i.e. the ability for parents to channel public subsidies to the school of their choice, possibly a private school) and performance-based rewards or sanctions for schools and teachers. Both these mechanisms encourage competition among schools. In this context, school autonomy is seen as providing the latitude for the school to devise particular strategies to compete with other schools and demonstrate high performance publicly.

Several countries have attempted to raise educational quality by enhancing parental choice and allowing schools to compete for students. In Chile, the market-oriented education reforms of the 1980s entailed the decentralisation of public school management responsibilities to municipalities and the introduction of a nationwide voucher programme (Cox, 2005). The latter is characterised by a flat per student public subsidy for schools which are part of the voucher system (municipal schools and the majority of private schools), complemented with schemes to provide extra funding on the basis of educational disadvantage, and parents' free choice of schools. Other countries where parental choice of schools is extensive include Belgium, Ireland, the Netherlands and Sweden. In a school system significantly relying on parental choice and competition, evaluation serves two main purposes: to assure that schools are meeting the centrally defined requirements that justify their receipt of public funds; and to assure that parents have reliable information to assist with their decisions. Information about the quality of education services provided by schools (e.g. publication of student results at the school level, publication of school evaluation reports) is essential to achieve these purposes.

Some countries have also strengthened performance-based incentives for schools. For instance, in Mexico, the Incentives Programme for Teacher Quality, introduced in 2008, provides collective and individual stimuli to teachers and school leaders working in schools obtaining the highest results or most significant progressions in national student assessments (Santiago et al., 2012b). Performance-based incentives as a market mechanism require elaborate evaluation and assessment procedures to determine performance levels.

The emergence of New Public Management

The expansion of evaluation and assessment within education systems also reflect governments' efforts to "modernise" the public sector and incorporate business practices into public service management. This trend, often referred to as New Public Management (NPM) or Results-Based Management, aims to reform public sector operations by improving cost-effectiveness, measuring output and making public bodies with greater autonomy accountable to citizens and system managers (Mons, 2009). It involves greater emphasis on quality assurance and quality management in the public sector. NPM puts emphasis on leadership principles, incentives and competition between public sector agencies and private entities to enhance the outcomes and cost-efficiency of public services (Parker and Gould, 1999).

Since the late 1980s, Sweden undertook far-reaching public sector reforms to ensure a more efficient government administration. In the education sector, this led to the introduction of a system of *management by objectives*, which underlies all educational activities, including evaluation and assessment. The purpose of management by objectives is to increase efficiency in central administration by setting goals and assessing outcomes rather than focussing on input and processes (Nusche et al., 2011b). As described in Burgess and Ratto (2003), the use of explicit incentives to improve the efficiency of the public sector was an important component of the United Kingdom’s public-service modernisation agenda of the late 1990s. The White Paper “Modernising Government” of 1999 emphasised the role of financial and other incentives in promoting better performance, leading to the development of performance indicators and systems of measuring and monitoring performance.

The rising importance of education in a global world

Economic activity has become globally interconnected on an unprecedented scale. The global character of markets has become stronger through international agreements and technological advances that bring people, goods and services together ever more quickly and less expensively. This growing integration of economies has an impact on strategies for national competitiveness, innovation, employment and skills (OECD, 2013). The emergence of the “knowledge society” and the strong skill bias in technological change have increased the value of education as a determinant of social and economic outcomes; this raises the payoff to good performance and amplifies the penalty for poor performance (OECD, 2008). The quality of education is necessary to achieve economic competitiveness in a context of global economic competition.

As a result, many of the proponents of national and international assessment place assessment in the context of a global economy, particularly in the context of being able to define a country’s position in educational achievement relative to that of economic competitors, on the assumption that performance on measures of scholastic achievement has implications for economic performance (Kellaghan and Greaney, 2001). International student assessments have, to some extent, become a measure of a country’s ability to compete in the global market and to drive economic growth. This has contributed to the expansion of evaluation and assessment activities in countries.

The growing imperative of an efficient use of public resources

The efficient use of resources is a growing concern. Education is costly and getting more so. OECD countries spent on average 6.2% of GDP on education institutions in 2009; between 1995 and 2009 the education share of public expenditure increased from 11.7% to 13.0% and real expenditure per student in pre-tertiary education increased by 55% between 1995 and 2009 (OECD, 2012).

The current financial crisis has intensified the need for efficiency in the use of public funds for education. Countries such as Greece, Ireland, Italy, Portugal and Spain face severe austerity measures which include cuts in the public education budget typically involving salary cuts for personnel working in public education, the freezing of career progression in the public service, and administration for education downsized. For example, in Ireland, the *value for money* imperative has been a fundamental part of public service modernisation and this has given an additional importance to evaluation and assessment in the educational context (Irish Department of Education and Skills, 2012). With challenging financial circumstances, the emphasis is on achieving greater efficiency

from the expenditure base. Effective monitoring, appraisal and evaluation is regarded as critical for delivering on this objective with a focus on the delivery of outputs and the achievement of goals/objectives. In Ireland *value for money* is a clear part of the rationale underpinning the National Literacy and Numeracy Strategy and the actions that have been identified to implement this plan (Irish Department of Education and Skills, 2012). Evaluation and quality assurance in education have become a necessity for policy makers to demonstrate that public funds are spent effectively and that the public purposes for financing education are actually fulfilled.

A shift from quantity to “quality for all”

In the great majority of OECD countries, attendance of lower secondary education is nearly universal – enrolment rates of 5-14 year-olds reached 95.9% on average across the OECD in 2010 (OECD, 2012) –, and graduation rates from upper secondary education have considerably increased – they reached 84% on average in the OECD area in 2010 (OECD, 2012). This is shifting the attention of policy makers from quantitative expansion to achieving “high-quality education for all”. With the move towards knowledge-driven economies and societies, education has never been more important for the future economic performance and relative economic standing of countries, but also to allow individuals to perform and fully participate in the economy and society (OECD, 2007a). In this context, broad participation in education is only one side of the coin. The quality of education delivered is important to ensure that school graduates are effectively equipped to participate in the new economy and society at large, capable to learn at a higher level, and prepared to subsequently engage in lifelong learning activities to update their knowledge and skills. As a result, the issue of quality provision has received more and more interest from the various stakeholders over the past few decades. The greater stress on quality has given more prominence to evaluation and assessment activities.

Well-designed evaluation and assessment activities are expected to ensure that: each student is provided with quality and relevant education; the overall education system is contributing to the social and economic development of the country; and each school agent is performing at their best to deliver efficient education services. A corollary of this is that educational goals place increasing emphasis on equity objectives, which enlarges the scope for evaluation and assessment activities.

Rising expectations of the professionalism of teachers

The quality of learning and the successful implementation of education reforms depend crucially on teachers who are facing rising demands (OECD, 2005). The more complex and uncertain the world in which we live, the more that alternative sources of knowledge and influence are available to students, the more open schools become to diverse clienteles, and the more varied the organisational and pedagogical strategies that teachers should deploy, the greater become the levels of professional skill needed to meet them. There are growing expectations that teachers can operate in new organisational structures, in collaboration with colleagues and through networks, and be able to foster individual student learning. These call for demanding concepts of professionalism: the teacher as facilitator and knowledgeable, expert individual and networked team participant, oriented to individual needs and to the broader environment, engaged in teaching and in research and development (OECD, 2001).

This has implications for the evaluation and assessment framework. First, the standards by which teachers are appraised need to reflect the increasingly demanding

definitions of teacher professionalism. Second, teacher appraisal assumes a key role in identifying professional development needs in the process of acquiring the wider range of skills and competencies needed to meet professional expectations. Third, teacher professionalism is expected to be central in the effective implementation of evaluation and assessment policies through their understanding of evaluation and assessment procedures as well as their commitment to them.

More educated parents

The rising general education attainments of the population have a range of impacts on the world of the school, particularly of reducing the distance between schools and teachers, on the one hand, and the general public and parents, on the other. Many are now very familiar with the world of education, and are themselves qualified to levels at or greater than teachers (OECD, 2001). A result is that parents and others are more articulate and more demanding of the work of schools and teachers. This contributes to pressures for greater accountability in education, requires schools to become more transparent, and leads education systems to be more demand-driven (OECD, 2006).

Klenowski and Wyatt-Smith (2012) describe, in relation to the publication of the results of the National Assessment Program – Literacy and Numeracy (NAPLAN) in Australia, the federal government reporting of the high levels of parental support for the initiative, indicating that it believed that it serves the best interests of transparency and accountability. Parents are also gaining greater voice as countries extend opportunities for parental and student feedback through questionnaires and surveys conducted at the school, national and international levels, which assess their levels of satisfaction across a range of educational areas.

Greater sophistication of systems for the certification of learning and the recognition of competencies

In recent years, a number of countries have developed sophisticated national qualification frameworks in view of certifying learning and recognising competencies in education settings providing increasing flexibility for the individual to define his or her own learning pathway. Modularisation and credit systems, as well as qualification frameworks with several qualification levels, grant individuals with considerable flexibility in their learning, particularly at the upper secondary level within the formal school system (Dufaux, 2012). In New Zealand, for example, an elaborated qualification framework has been put in place to enable students to individualise their learning and have it formally recognised. The main qualification in secondary education is the National Certificate of Educational Achievement (NCEA), in which students are assessed against a range of National Standards. These consist of over 26 000 unit standards (vocationally based and mostly used in workplace training and the tertiary sector) and about 850 achievement standards (academically based and focused on the secondary school curriculum). Schools can design and offer their own courses mixing unit standards and achievement standards. This allows students to choose their personal learning content (Nusche et al., 2012).

The greater sophistication of certification systems, including the individualisation of learning, brings new challenges to student assessment. It requires flexible assessment instruments, accuracy in the assessment of very specific standards, and high capacity to assess and administer the qualifications framework. In these contexts, countries use a

greater variety of assessment tools such as short tests, projects, field studies, practical and aural tests (see also Chapter 4).

The growing importance of evidence-based policy

Another driver of the expansion of evaluation and assessment in school systems is the wider movement both within and beyond education to advance the use of evidence-based decision making. Data on student achievement are increasingly being used to support effective policy and practice, and to move education systems towards more evidence-informed approaches to large-scale improvement (Campbell and Levin, 2009). The public, professionals and policy makers want to know that their decisions, investments and actions are based on evidence. In education, this requires balancing schools' needs for data with external requirements and reporting (Campbell and Levin, 2009).

According to an OECD study on evidence in education (OECD, 2007b), the recent resurgence of interest in evidence-informed policy research can be explained by a range of factors such as: a greater concern with student achievement outcomes; a related explosion of available evidence due to a greater emphasis on testing and assessment; more explicit and vocal dissatisfaction with education systems, nationally and locally; increased access to information via the Internet and other technologies; and resulting changes in policy decision making. This highlights the interconnection between evidence-based policy and evaluation and assessment in education systems. Evaluations and assessments are key elements in the decision-making process. They provide the information on which accountability judgements are made and the means for steering improvement in educational practice.

Technological advancements

Information technology has developed very rapidly over the past 40 years, with computers becoming smaller, faster, cheaper, and more powerful. The ease and speed at which very large quantities of information can be rapidly accessed in a variety of settings have considerably improved (OECD, 2013). The digital revolution has drastically improved capacity to store, transmit, access and use information. The cost of transmitting information has significantly fallen, leading to the quasi abolition of physical distance. This has led to new developments in education technology – from Internet access to new teaching techniques enabled by classroom computers, which are driving changes in the education environment (ECS, 1999).

Countries are making significant investment in educational ICT infrastructure and equipment, and technology is increasingly being used to change what happens in the classroom and the school. ICT offers many opportunities to store and share data, to manage large amounts of information, to foster dialogue among education professionals, to strengthen feedback mechanisms and to improve the sophistication of evaluation and assessment procedures. The growing volume of data at all levels – student, teacher, school, local, national, and international – on education inputs and outcomes makes the monitoring of performance much easier, almost in real-time. Easier forms of communication (e.g. e-mail) improve the involvement of parents in school and their interest in following their children's progress. At the same time, teachers are able to use technology for professional development, online research and classroom and administrative data gathering (ECS, 1999). This is in addition to improvements in system-level initiatives such as the design, implementation and scoring of student standardised assessment; the development of Internet platforms to share education data among

stakeholders; and data information systems to facilitate knowledge management within evaluation and assessment frameworks.

The emergent commercial interests in education

The private sector is more and more a large player in the provision of ancillary services in education. In most OECD countries, it typically provides services such as the design and implementation of student standardised assessment, student private tutoring, online educational materials for students and teachers, textbooks and resources for school or classroom management. This is extensively the case in the United States. According to Burch (2009), the No Child Left Behind (NCLB) Act has helped private firms enter local education markets. She argues that the firms draw on political networks, new technologies and capital investments to become major suppliers to school systems for a vast array of educational services, including test score data storage, remedial instruction for the poor, online curriculum and online school management. She further claims that, triggered by high-stakes accountability policies, companies have introduced products and services that elevate the importance of standardised student assessments, private tutoring and technologies for school management (Burch, 2009).

In some countries, much of the activity of firms has concentrated on standardised student assessment, which is a growing and profitable industry. For instance, in the United States, the NCLB Act requires approximately 45 million standardised tests annually with considerable associated costs for developing, administering, publishing, scoring and reporting NCLB standardised tests (Toch, 2006). Burch (2009) emphasises that the market for test development and preparation has exploded in recent years. According to her, in 2006, the top vendors reported annual sales in the range of USD 100-600 million, with a pattern of increasing sales since the adoption of NCLB. Moreover, the testing market in the United States is dominated by only a handful of companies, which represent 90% of testing revenue (Toch, 2006). As standardised student assessment becomes a more profitable industry, companies have strong incentives to lobby for the expansion of student standardised assessment as an educational policy therefore influencing the activities within the evaluation and assessment framework.

The media as a driver of accountability in education

An important contextual influence for the development of evaluation and assessment is the role of the media in education. As data on student performance becomes readily available, as there is growing pressure for an effective use of public funds, and as the general public demands transparency in the delivery of education services, the media increasingly engages in the education public debate and makes information about student performance available, particularly school league tables. Governments are under pressure to release such information, also to prevent misinterpretation of data often presented by newspapers in simplistic ways. In addition, the attention the media devotes to education issues is also an opportunity for governments to publicise their accomplishments and feature the impact of their education policies. Given the greater mediatic impact of accountability policies, particularly those involving the measurement of student outcomes which have the potential of being reflected in school league tables, there are risks that accountability in education is, to some extent, driven by pressures from the media.

Notes

1. After two decades of decentralisation, Hungary has experienced a trend towards a larger degree of central decision-making in education. Following new legislation passed in 2011 and 2012, schools and other public educational institutions, with the exception of those maintained by the private sector and religious authorities, are subject to direct governance by central authorities (including funding allocation) from 2013 onwards. Except when explicitly indicated, information about Hungary in this report refers to the period prior to this reform.

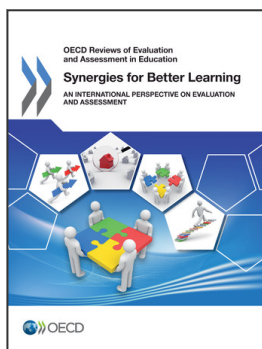
References

- Bonnet, G. (2004), “Evaluation of education in the European Union: Policy and methodology”, *Assessment in Education*, Vol. 11, No. 2.
- Broadfoot, P. and P. Black (2004), “Redefining assessment? The first ten years of *Assessment in Education*”, *Assessment in Education*, Vol. 11, No. 1.
- Burch, P. (2009), *Hidden Markets: The New Education Privatization*, Taylor and Francis.
- Burgess, S. and M. Ratto (2003), “The role of incentives in the public sector: Issues and evidence”, *Oxford Review of Economic Policy*, Vol. 19, No. 2.
- Campbell, C. and B. Levin (2009), “Using data to support educational improvement”, *Educational Assessment, Evaluation and Accountability*, No. 21, pp. 47-65.
- Cox, C. (ed.) (2005), *Educational Policies for the Turn of the Century: The Chilean School Reform*, Editorial Universitaria, Santiago, Chile.
- Danish Ministry of Education and Rambøll (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Update to the Country Background Report for Denmark*, National Agency for Quality and Supervision, www.oecd.org/edu/evaluationpolicy.
- Dos Santos, S. and R. Rakocevic (2012), *Les cadres d'évaluation en vue d'améliorer les résultats scolaires : Étude de l'OCDE – Rapport national de base pour la France*.
- Dufaux, S. (2012), “Assessment for qualification and certification in upper secondary education: A review of country practices and research evidence”, *OECD Education Working Papers*, No. 83, OECD Publishing, Paris, www.oecd.org/edu/workingpapers.
- ECS (Education Commission of the States) (1999), *Future Trends Affecting Education*, Denver, Colorado, www.ecs.org/clearinghouse/13/27/1327.htm.
- Elmore, R.F. (2004), “The problem of stakes in performance-based accountability systems”, in S.H. Fuhrman and R.F. Elmore (eds.), *Redesigning Accountability Systems for Education*, Teachers College Press, New York.
- Eurydice (2009), *National Testing of Pupils in Europe: Objectives, Organisation and Use of Results*, Eurydice, Brussels.
- Eurydice (2007), *School Autonomy in Europe: Policies and Measures*, Eurydice, Brussels.
- Figlio, D. (2006), “Testing, crime and punishment”, *Journal of Public Economics*, Vol. 90, pp. 837-851.
- Figlio, D. and L. Kenny (2009), “Public sector performance measurement and stakeholder support”, *Journal of Public Economics*, Vol. 93, Issues 9–10, pp. 1069–1077.

- Flemish Ministry of Education and Training and the University of Antwerp Edubron Research Group (2010), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Flemish Community of Belgium*, www.oecd.org/edu/evaluationpolicy.
- Hamilton, L. (2003), “Assessment as a policy tool”, *Review of Research in Education*, Vol. 27 (1), January, pp. 25-68.
- Hamilton, L. and D. Koretz (2002), “Tests and their use in test-based accountability systems”, in L. Hamilton, B. Stecher and S. Klein (eds.), *Making Sense of Test-based Accountability in Education*, RAND, Santa Monica, California, pp. 13-49.
- Hargreaves, A. and D. Shirley (2009), *The Fourth Way: The Inspiring Future for Educational Change*, Corwin, Thousand Oaks, California.
- Hout, M. and S. Elliott (2011), *Incentives and Test-Based Accountability in Education*, Committee on Incentives and Test-Based Accountability in Public Education, National Research Council, The National Academies Press, Washington, D.C.
- Irish Department of Education and Skills (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Ireland*, www.oecd.org/edu/evaluationpolicy.
- Kellaghan, T. and V. Greaney (2001), “The globalisation of assessment in the 20th century”, *Assessment in Education*, Vol. 8, No. 1.
- Klenowski, V. and C. Wyatt-Smith (2012), “The impact of high stakes testing: The Australian story”, *Assessment in Education: Principles, Policy and Practice*, Vol. 19, No. 1, February, pp. 65-79.
- Looney, J. (2009), “Assessment and innovation in education”, *OECD Education Working Papers*, No. 24, OECD Publishing, Paris, www.oecd.org/edu/workingpapers.
- Mons, N. (2009), “Theoretical and real effects of standardised assessment”, background paper to the study: National Testing of Pupils in Europe, Eurydice Network, http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/1111EN.pdf, accessed 12 July 2011.
- New Zealand Ministry of Education (2010), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for New Zealand*, www.oecd.org/edu/evaluationpolicy.
- Norwegian Directorate for Education and Training (2011), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Norway*, www.oecd.org/edu/evaluationpolicy.
- Nusche, D., D. Laveault, J. MacBeath and P. Santiago (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Nusche, D., L. Earl, W. Maxwell and C. Shewbridge (2011a), *OECD Reviews of Evaluation and Assessment in Education: Norway*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Nusche, D., G. Halász, J. Looney, P. Santiago and C. Shewbridge (2011b), *OECD Reviews of Evaluation and Assessment in Education: Sweden*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.

- OECD (2013), *Trends Shaping Education 2013*, Centre for Educational Research and Innovation, OECD Publishing, Paris.
- OECD (2012), *Education at a Glance: OECD Indicators 2012*, OECD Publishing, Paris.
- OECD (2011), *Education at a Glance: OECD Indicators 2011*, OECD Publishing, Paris.
- OECD (2008), “How do OECD countries take stock of progress and performance in education systems? Evidence and issues”, paper prepared for the Education Policy Committee Ad Hoc Workshop “Taking Stock of Education Performance: From Student Testing to System Evaluation”, EDU/EDPC(2008)33, Paris.
- OECD (2007a), *Human Capital: How What You Know Shapes Your Life*, OECD Publishing, Paris.
- OECD (2007b), *Evidence in Education: Linking Research and Policy*, OECD Publishing, Paris.
- OECD (2006), *Schooling for Tomorrow: Demand-Sensitive Schooling? Evidence and Issues*, Centre for Educational Research and Innovation, OECD Publishing, Paris.
- OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, OECD Publishing, Paris.
- OECD (2001), *Schooling for Tomorrow: What Schools for the Future?* Centre for Educational Research and Innovation, OECD Publishing, Paris.
- Parker, L. and G. Gould (1999), “Changing public sector accountability: Critiquing new directions”, *Accounting Forum*, Vol. 23, No. 2.
- Roeber, E. (1988), “A history of large-scale testing activities at the state level”, paper presented at the Indiana Governor’s Symposium on ISTEP, Madison.
- Santiago, P., F. Benavides, C. Danielson, L. Goe and D. Nusche (forthcoming), *OECD Reviews of Evaluation and Assessment in Education: Teacher Evaluation in Chile*, OECD Publishing, Paris, forthcoming at www.oecd.org/edu/evaluationpolicy.
- Santiago, P., G. Donaldson, A. Looney and D. Nusche (2012a), *OECD Reviews of Evaluation and Assessment in Education: Portugal*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Santiago, P., I. McGregor, D. Nusche, P. Ravela and D. Toledo (2012b), *OECD Reviews of Evaluation and Assessment in Education: Mexico*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Santiago, P., G. Donaldson, J. Herman and C. Shewbridge (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Schildkamp, K., M. Ehren and M.K. Lai (2012), “Editorial article for the special issue on data-based decision making around the world: From policy to practice to results”, *School Effectiveness and School Improvement*, Vol. 23, No. 2, June, pp. 123-131.
- SEP and INEE (Secretariat of Public Education and National Institute for Educational Assessment and Evaluation) (2011), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Mexico*, www.oecd.org/edu/evaluationpolicy.

- Shewbridge, C., E. Jang, P. Matthews and P. Santiago (2011a), *OECD Reviews of Evaluation and Assessment in Education: Denmark*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Shewbridge, C., M. Hulshof, D. Nusche and L. Stoll (2011b), *OECD Reviews of Evaluation and Assessment in Education: School Evaluation in the Flemish Community of Belgium*, OECD Publishing, Paris, www.oecd.org/edu/evaluationpolicy.
- Toch, T. (2006), *Margins of Error: The Education Testing Industry in the No Child Left Behind Era*, Education Sector Reports, www.educationsector.org/publications/margins-error-testing-industry-no-child-left-behind-era, accessed 12 July 2011.
- Wang, L., G. Beckett and L. Brown (2006), “Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence”, *Applied Measurement in Education*, Vol. 19, No. 4, pp. 305-328.



From:

Synergies for Better Learning

An International Perspective on Evaluation and Assessment

Access the complete publication at:

<https://doi.org/10.1787/9789264190658-en>

Please cite this chapter as:

OECD (2013), "Trends in evaluation and assessment", in *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264190658-5-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.