OECD Education Working Papers No. 24

Assessment and Innovation in Education

Janet Looney

https://dx.doi.org/10.1787/222814543073

OECD

Unclassified **EDU/WKP(2009)3**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

English - Or. English

DIRECTORATE FOR EDUCATION

EDU/WKP(2009)3
Unclassified

ASSESSMENT AND INNOVATION IN EDUCATION

OECD Education Working Paper No. 24

By Janet Looney

High-stake assessment can be an obstacle to the introduction of innovative practices in education and to the development of social and behavioural competences that are crucial for innovation (but typically fall out of the scope of existing assessments).This report reviews the evidence on the impact of high-stake summative assessment on innovation in education and explores possible ways to reconcile high-stakes assessments and examinations through innovative approaches to testing.

This report was prepared in collaboration with the OECD Secretariat for the project "Innovation Strategy for Education and Training" undertaken by the Centre for Educational Research and Innovation (CERI) (www.oecd.org/edu/innovation). It contributes to the human capital pillar of the OECD Innovation Strategy (www.oecd.org/innovation/strategy).

Contact: Stéphan Vincent-Lancrin, OECD, EDU/CERI
E-mail: Stephan.Vincent-Lancrin@oecd.org, Tel: +33 1 45 24 92 29

English - Or. English

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

OECD DIRECTORATE FOR EDUCATION

OECD EDUCATION WORKING PAPER SERIES

This series is designed to make available to a wider readership selected studies drawing on the work of the OECD Directorate for Education. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language (English or French) with a short summary available in the other.

Comment on the series is welcome, and should be sent to either edu.contact@oecd.org or the Directorate for Education, 2, rue André Pascal, 75775 Paris CEDEX 16, France.

The opinions expressed in these papers are the sole responsibility of the author(s) and do not necessarily reflect those of the OECD or of the governments of its member countries.

Applications for permission to reproduce or translate all, or part of, this material should be sent to OECD Publishing, rights@oecd.org or by fax 33 1 45 24 99 30.

---

www.oecd.org/edu/workingpapers

Applications for permission to reproduce or translate
all or part of this material should be made to:

Head of Publications Service
OECD
2, rue André-Pascal
75775 Paris, CEDEX 16
France

Copyright OECD 2009

SUMMARY

Do some forms of student (and school) assessment hinder the introduction of innovative educational practices and the development of innovation skills in education systems? This report focuses on the impact of high-stake summative assessment on innovation and argues that it is possible to reconcile high-stakes assessments and examinations through innovative approaches to testing.

While necessary, assessment based on high-stake examinations often acts as an incentive to teach or study "to the test". It may thus limit risk-taking by teachers, students and parents, for instance. The problem may be amplified if a system of accountability and incentives uses the results of these examinations and tests to assess teachers and schools. What should be done to ensure that the systems used to assess education systems do not stifle the risk-taking inherent to innovation – and that they foster innovation skills in students?

This study proposes three main ways of combining assessment and innovation: 1) developing a wide range of performance measurements for both students and schools; 2) rethinking the alignment of standards and assessment; 3) measuring the impact of assessments on teaching and learning.

One way of influencing teaching and learning might be to modify high-stake testing. Systems will adapt to this, and both teaching and learning will focus on acquiring the right skills. Rather than testing the content of learning, standards could relate to cognitive skills such as problem-solving, communicating and reasoning – with test/examination developers adapting those skills to subjects such as mathematics, science or literary analysis. Similarly, more use might be made of innovative assessment methods based on information and communication technologies, inasmuch as these may feature simulation or interactivity, for instance, at a reasonable cost. Focusing the assessment on cognitive processes rather than content would leave more scope for teachers to put in place innovative teaching/learning strategies. This does, however, assume a high standard of professionalism in teachers and an adequate system of continuing training and knowledge management.

As a single type of assessment cannot fully capture student learning, one effective strategy might also be to multiply the number of measurements and thus relieve the pressure on students and teachers to perform well in a single, high-visibility, high-stake test. At the same time, this larger number of measurements could provide the necessary input for systems based on accountability, diagnosis and assessment of the effectiveness of innovative practice.

Finally, assessing the technical standard of tests and examinations is an integral part of their development, but it is less common to address the impact they have on teaching/learning or the validity of how their results are used. Since assessment is an integral part of the education process, it is just as important to assess tests and examinations as it is other educational practices in order to achieve improvements and innovation in educational assessment, but also in educational practice.

RÉSUMÉ

Certaines formes d'évaluation des élèves (et des écoles) font-elles barrière à l'introduction de pratiques pédagogiques innovantes et au développement des compétences individuelles pour l'innovation au sein des systèmes éducatifs? Ce rapport se concentre sur l'impact de l'évaluation à fort enjeu (*high-stake summative assessment*) sur l'innovation et argue qu'il est possible de réconcilier ce type d'évaluation et d'examens grâce à des approches innovantes de l'évaluation.

Bien que nécessaire, l'évaluation reposant sur des examens à forts enjeux donne souvent des incitations à enseigner et à étudier dans le but premier de réussir à l'examen. Elle peut ainsi limiter la prise de risque des enseignants, des étudiants, des parents, etc. Ce problème peut être amplifié si un système de responsabilisation et d'incitations utilise les résultats de ces examens et tests pour évaluer les enseignants et les écoles. Comment faire pour que les systèmes d'évaluation des systèmes éducatifs n'empêchent pas la prise de risque inhérente à l'innovation et permettent le développement des compétences pour l'innovation des étudiants ?

Cet examen propose trois pistes principales pour combiner évaluation et innovation : 1) développer une multiplicité de mesures de la performance des étudiants et des écoles ; 2) repenser l'alignement entre objectifs (standards) et évaluation ; 3) évaluer l'impact des évaluations sur l'enseignement et l'apprentissage.

Une manière d'influencer l'enseignement et l'apprentissage pourrait consister à modifier les examens à forts enjeux. Les systèmes vont s'y adapter, et l'on étudiera et l'on enseignera pour acquérir les bonnes compétences. Plutôt que de tester les contenus d'apprentissage, les objectifs pourraient porter sur des compétences cognitives telles que la résolution de problèmes, la communication et le raisonnement – avec une adaptation par les développeurs de tests et examens de ces compétences aux mathématiques, à la science, à l'analyse littéraire, etc. On pourrait de même davantage tirer profit des méthodes innovantes d'évaluation basées sur les technologies d'information et de communications, dans la mesure où celles-ci peuvent intégrer des simulations, de l'interactivité, etc., à un coût raisonnable. Axer l'évaluation sur des processus cognitifs plutôt que sur les contenus laisserait davantage de liberté aux enseignants pour mettre en place des stratégies innovantes d'enseignement et d'apprentissage. Cela suppose cependant un haut niveau de professionnalisme des enseignants et un système de formation continue et de gestion des connaissances adéquats.

Dans la mesure où un type d'évaluation ne peut pas capturer pleinement l'apprentissage des étudiants, une stratégie efficace pourrait aussi être de multiplier les mesures afin de baisser la pression qui pèse sur les étudiants et les enseignants de bien réussir un seul test à haute visibilité et fort enjeu. Cette multiplicité de mesures pourrait dans le même temps fournir l'information nécessaire à des systèmes de responsabilisation, de diagnostic et d'évaluation de l'efficacité des pratiques innovantes.

Enfin, alors que l'évaluation de la qualité technique des évaluations et examens fait partie intégrante de leur développement, l'impact qu'ils ont sur l'enseignement et l'apprentissage ou la validité des usages que l'on fait de leurs résultats sont plus rarement examinés. Dans la mesure où l'évaluation fait partie intégrante du processus pédagogique, l'évaluation des tests et examens est tout aussi importante que celle des autres pratiques pédagogiques pour le perfectionnement et l'innovation de l'évaluation dans l'éducation, et aussi des pratiques pédagogiques.

TABLE OF CONTENTS

ASSESSMENT AND INNOVATION IN EDUCATION 4

1. Why innovate? 4

2. What counts as innovation in education? 5

2.1. Innovations in teaching, learning and assessment 5
2.2. Many roles for assessments and examinations 7

3. Do high-stakes summative assessments and examinations undermine innovations? 8

3.1. Alignment 9
3.2. The incentives of high-stakes assessments and examinations and the problem of "teaching to the test" 10
3.3. Tests worth teaching to? 12
3.4. The impact of high-stakes testing and expectations on student (and teacher) engagement and motivation 13
3.5. Implementation challenges 14

4. Is it possible to reconcile tensions between innovation and high-stakes assessment? 16

4.1. Promoting innovative approaches to assessments and examinations 16
4.2. Rethinking alignment 18
4.3. Using multiple measures of student and school performance to evaluate quality and outcomes 20
4.4. Creating a better balance between high-stakes and school empowerment 21

5. Conclusion 21

REFERENCES 23

ANNEX 1: NATIONAL ASSESSMENTS AND EXAMINATIONS IN OECD COUNTRIES 30

ANNEX 2. OVERVIEWS OF OECD COUNTRY POLICIES 34

ASSESSMENT AND INNOVATION IN EDUCATION

Janet Looney[1]

1. Why innovate?

1. Innovation is a key driver of economic and social progress in OECD countries. If absent, innovation growth stalls; economies and communities stagnate. As this paper is being written in early 2009, countries are facing a major global financial and economic crisis, and the need to promote innovation appears to be stronger than ever. Economists are emphasizing the importance of large-scale investments in innovative programmes, including for education, which will help revive economies in the shorter term and will also have a lasting impact. At the same time, Schumpeter's classic definition of innovation as "creative destruction" – that is, "… destroying an old combination and realizing a new one" – still pertains (Schumpeter, 1942). In some senses, the current crisis also represents a major opportunity for "creative destruction".

2. Innovation has been a topic of considerable interest in the education sector for some time. Indeed, successful innovation depends upon the human creativity, knowledge, skills and talents that are nurtured and developed, in large part, through education. How can schools and teachers better cultivate these capacities? And, importantly, how can educational systems develop their own capacities for innovation?

3. Beyond the current economic crisis, trends behind the drive for innovation in education and learning include:

- social and economic pressures to raise achievement levels and to ensure greater equity of outcomes for *all* students;
- changes in work, social and family life;
- rapidly advancing technologies;
- the need to motivate and engage students.

4. Education systems across OECD countries have made two particularly significant changes in response to these trends. The first is the focus on "lifelong learning". Schools are charged with instilling in students a desire to learn, and developing their skills and capacities for "learning-to-learn" (for example, skills for problem solving and critical analysis, as well as greater autonomy, initiative, flexibility, and so on). The second significant change is the focus on equity of student outcomes. Certainly equity of access and opportunity has been an important principle for modern education. But the focus on improving equity of outcomes – that is, helping all students to achieve to higher levels – is relatively new. The shift is, in part, a response to the need for more workers with higher skills in knowledge economies. There is also a growing recognition that education has a positive impact on the well-being of communities, families and individuals (OECD, 2007).

5. Top-down policies that have dictated the content and methods of teaching and learning for much of the twentieth century are seen as being inadequate to the tasks of modern schools. OECD countries now encourage greater school autonomy and responsiveness to local contexts and individual student needs. Schools with greater autonomy, it is believed, will also have more freedom to innovate, and to deepen knowledge about what works, for whom, and under what circumstances.

6. Assessment is vital to the process of innovation. Those implementing innovative projects need to assess their effectiveness and to make necessary adaptations. Evidence on the impact of new approaches is also essential for successful dissemination. At the policy level, assessment ensures that schools, while exercising greater autonomy, are still meeting centrally defined standards.

7. This report is about the intersection of high-stakes assessments and examinations with innovation in education. While assessment is an essential component of innovation, high stakes assessments have the potential to undermine it. High stakes – such as financial sanctions or the threat of school closure – are intended to provide incentives for teachers to focus on central standards, and to provide information for school improvement. Yet, these high stakes also discourage risk taking necessary for innovation, and may often encourage teachers to "teach to the test". Highly competitive university entrance examinations also have a powerful impact on classroom teaching and learning. The mismatch between goals set out in central standards and measured in high-stakes tests or examinations, and those emphasised in innovative educational programmes – may mean that it is difficult for these innovative programmes to show proof of their strengths.

8. Getting assessment right will be one of the most important priorities for education systems in OECD countries over the near term. Assessment is important not only for purposes of accountability (and accountability will rightly remain a high priority), but also for building the evidence base on "what works in innovation". Those systems that best support diversity and complexity, and that have effective ways to evaluate learning will have the richest opportunities to deepen knowledge.

9. The following three major sections of this report will examine these issues in more depth. The next section of the report provides a brief overview of "what counts as innovation" in OECD education systems. The report then turns to examination of whether and how high-stakes assessments and examinations may undermine educational innovation. The final major section of the report explores ways to reconcile assessments, examinations, and innovation.

2. What counts as innovation in education?

10. The OECD refines the Schumpeterian definition of innovation as "creative destruction", differentiating between product innovations (goods or services), and innovations in processes (methods, practices, and organisation) (OECD and Eurostat, 2005). Educational innovations generally fall within the latter category and include new or significantly improved approaches to classroom-based teaching, learning and assessment, as well as changes in the organisation, or governance, of systems.

2.1. Innovations in teaching, learning and assessment

11. In describing what counts as innovation in teaching, learning and assessment, it may be helpful to first depict what are thought of as more "traditional" practices. Traditional approaches to teaching are often characterised as involving "direct transmission": teachers communicate knowledge in a clear and structured ways, and demonstrate and explain solutions. While teachers may also focus on the development of skills for critical analysis and reasoning, students have few opportunities to experiment with or apply new knowledge in different situations. Knowledge is presented within traditional, discrete categories (language, mathematics, science, arts, and so on), and there may be few opportunities to link learning

across subjects. There is little attention paid to student motivation, or to developing skills for "learning to learn" (OECD, forthcoming).

12. Teaching and learning approaches considered as innovative, on the other hand, are generally characterised as being "student-centred", or "constructivist". The development of thinking and reasoning skills (skills for "learning-to-learn"), and the ability to synthesise knowledge across the curriculum are emphasised. Teachers take on different roles, moving away from "front-of-the classroom" teaching to more active engagement in the learning process with students; students conduct their own inquiries and develop solutions. Teachers and students may tailor programmes according to the learning needs and interests of individual students (OECD, forthcoming).

13. 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 judgments of student achievements. Decisions as to whether students should advance to the next level, go on to vocational or university programmes, and so on are based on the results of "summative assessments". 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. This approach fits well with goals of OECD countries to promote lifelong learning (which relies upon skills for learning-to-learn, including skills for self-assessment).

14. These broad themes in teaching, learning and assessment are seen as supporting student "learning for the 21st century", and are increasingly embedded in national policies in OECD countries. For example:

- Several countries promote the development of key competences in their national curriculum frameworks [Austria, Belgium (the French and Flemish communities), the Czech Republic, France, Finland, Germany, Hungary, Italy, Luxembourg, New Zealand, Norway, Switzerland, and the United Kingdom]. These include skills, knowledge and attitudes for "learning-to-learn", team and independent work, skills for communication, and so on.

- Cross-curricular, or integrated studies are increasingly common [Belgium (the Flemish community), Japan, Korea, the Netherlands, Poland, Switzerland, Sweden, Turkey and the United Kingdom]. National curriculum frameworks may encourage teaching of basic literacy and numeracy skills within the context of other subjects. Vocational and academic streams may be integrated. Or students may have the opportunity to participate in interdisciplinary projects. The emphasis is on helping students to develop the capacity to synthesise knowledge from a range of disciplines.

- Several countries promote constructivist, student-centred approaches to teaching in curriculum guidelines and professional development programmes [Canada, Finland, Ireland, Japan, Korea, Turkey].

- Several countries have policies promoting the use of formative assessment (the frequent, interactive assessment of student understanding and progress to identify learning needs and adapt teaching) as a way to build students' skills for self-assessment and learning-to-learn, and to raise levels of achievement, particularly for low performing students [Australia, Canada, Denmark, Finland, Ireland, Italy, New Zealand, Switzerland, Turkey, and the United Kingdom] (OECD, 2005).

- There are initiatives for the "personalisation of learning", including greater freedom of choice within curriculum [France, Hungary, Japan and the United Kingdom].

- National policies may support the development of specialist schools and programmes, promoting specific subject interests as well as new approaches to teaching and learning [Australia, the Netherlands, and the United Kingdom].

- There is a universal emphasis on the importance of developing students' ICT literacy, as well as its use as a tool for learning. Governments have invested significant levels of funding in ICT and related teacher professional development.

15. Education stakeholders in the private and non-profit sectors also offer strong support for innovation in teaching, learning and assessment – there is no lack of energy in this regard.

16. Innovative approaches to education are increasingly part of the common discourse. Yet, within schools and classrooms, innovations tend to happen only at the margins and may not be sustained over time. For example, video studies conducted for the international Trends in Mathematics and Science Study (TIMSS)[2] (International Association for the Evaluation of Educational Achievement [IEA], 1999) showed that traditional approaches, including teacher-led structured activities such as lectures and students working on assignments, were dominant in almost all countries. Teachers, as well as other education stakeholders, may also have multiple, and sometimes conflicting, expectations as to "what education is for". Goals to achieve greater equity of educational outcomes for students co-exist with merit-based competition. Education systems support skills for higher-order learning[3], yet innovative programmes promoting such skills are sometimes seen as lacking in rigour. These tensions play out in conflicting policies and practices related to assessment and innovation.

2.2. Many roles for assessments and examinations

17. The great majority of OECD countries have decentralised education systems to encourage greater school autonomy and innovation. National or regional education authorities ensure a degree of coordination, standardization and transparency across systems. Schools are held accountable for helping all students to meet centrally developed standards and fulfil goals within the curriculum. Several countries also require external school evaluations (usually on a tri-annual basis), and/or school self-evaluation (usually annual). In systems that encourage competition, information provided by assessments may also help parents to decide where to send their children to school. (For a detailed overview of OECD country approaches to assessment, examination and evaluation, see Annexes 1 and 2 of this report.)

18. As previously noted, systems may attach "high stakes" to school performance on assessments and examinations (Canada, the United Kingdom and the United States being the most prominent countries to promote high stakes assessments for schools; in France, Japan, Korea, and Hungary, university entrance exams are considered as high stakes for students).

19. High stakes assessments for accountability are intended to strengthen incentives for teachers to focus their attention on helping all students to meet standards. The "stakes" refer to the possibility of school shut down or reconstitution, teacher job loss, or financial sanctions for schools or individual teachers. Publication of the results of student performance of different schools on external assessments and evaluations, as happens in a majority of OECD countries, also add to the stakes.

20. Examinations that determine which school track students will follow, whether they will graduate or qualify for entrance to university also may hold very high stakes for students. Examination results may have a long-term impact on students' life chances, including employment and income. While technically examinations with civil effect for students do not hold high stakes for schools, a school's reputation may rest on the success of its graduates.

Box 1. A summary of terms related to assessment and evaluation

A range of terms is used to indicate the different functions of assessment and evaluation in education:

Formative assessment refers to the frequent, interactive assessment of student understanding and progress to identify learning needs and adapt teaching. Formative assessment differs from summative assessment in that the information gathered is used to shape improvements, rather than serve as a summary of student achievement.

Summative assessments include tests and examinations that seek to provide summary statements of student achievements and capabilities. Summary assessments may be conducted at the end of a unit of study, to determine student readiness for promotion from one grade to the next.

Summary assessments include **examinations with civil effect for students,** for example, graduation/completion of a programme, or formal admission to vocational or upper secondary school, or university. These examinations are sometimes also used for school accountability or to provide feedback to schools and teachers.

National or regional assessments are tests of student achievement, but are conducted primarily to evaluate the performance of a system as whole, to hold schools accountable for student results, and/or to provide feedback to schools and teachers. These tests do not have civil effect for students (that is, graduation, university admission). The results of national or regional assessments may be published.

Evaluations are judgments of organisational effectiveness. They may include internal school/self-evaluations, and external evaluations or school inspections. (For the purposes of this report, assessment refers to judgments of individual student performance, while evaluation refers to organisational or school performance.) Evaluations may be administered at central or at regional/local levels. The aim of evaluation is to provide feedback to schools and teachers to improve practice, and/or to inform parents and the community of school performance.

Source: OECD (2005, 2008).

21. The results of assessments, examinations and evaluations may also be used to shape improvements in school management and classroom teaching and learning. School leaders may use assessment or evaluation results to support decisions about how to direct resources. Teachers may spend more time with students in areas where they are underperforming. Ideally, school leaders and teachers may also use information from assessment and evaluations to refine school and classroom-based innovations, to measure the impact of new methods on learning, and to build the evidence base on "what works". Assessments and examinations are thus seen as a way to empower schools. Yet these different roles for assessment – for accountability and for school empowerment – may sit together uneasily. The sanctions associated with high stakes assessment regimes may overwhelm messages regarding school empowerment.

3. Do high-stakes summative assessments and examinations undermine innovations?

22. The key question addressed in this report is whether and to what extent high-stakes assessments and examinations may undermine educational innovation. As noted above, assessment and evaluation are critical to the functioning of education systems. And some studies have shown that students in programmes focused on higher-order learning – as emphasised in innovative programmes –perform as well or better on external assessments than students in more traditional programmes (described in more detail below). But this section will show that the weight of the evidence points to the difficulty for schools and teachers of initiating and/or sustaining innovation in high-stakes environments. Indeed, high-stakes tests do little to support and may often undermine educational innovation.

23. The challenges are both social and technical. Many of the technical problems with high-stakes assessment are well established, but it is worth reviewing them here, as discussion often rests within community of assessment experts, and issues are not openly debated in policy realms or in the broader

public. The social challenges are closely related to the technical – both policy-makers and the public maintain faith in "hard numbers", and their capacity to communicate quality.

24. In this section we explore in more detail:

- the problem of alignment;
- incentives of high-stakes assessments and examinations and the problem of "teaching to the test";
- tests worth teaching to?
- the impact of high-stakes assessments and examinations on student (and teacher) motivation, engagement and morale;
- implementation challenges.

3.1. Alignment

25. Alignment of standards, curriculum and assessment is critical in the standards-based assessment systems now common across OECD countries. In a well-aligned system, standards for student performance are comprehensive, and clearly define the content and cognitive processes students are expected to demonstrate at different stages. The curriculum covers learning objectives identified in the standards, and tests test students' knowledge and skills in relation to those standards. Effective assessments also take into account relative emphasis of different elements of standards and curriculum. If systems are poorly aligned, it is impossible for schools and teachers to draw valid conclusions about student performance, or to adapt teaching to better meet identified needs – the central goal of Standards Based Assessment systems (Linn, 2001, 2005; Haertel and Herman, 2005).

26. The early experiences of the New American Schools (NAS) initiative in the United States provide one dramatic example of how high-stakes assessments may undermine reform efforts. The initiative was launched in the 1990s to help lower performing schools adopt innovative, evidence-based approaches to teaching and learning. NAS school designs were adopted in approximately 3 000 schools across the United States The success of the designs was to be measured according to student performance on high-stakes tests, as well as through independent evaluation of the RAND Corporation. University-based researchers also followed the implementation as well as outcomes of the initiative. The studies yielded several lessons related to developing the capacity of schools and teachers. Of particular interest for this report were findings related to the impact of high-stakes assessment on innovation and development. The RAND evaluation of the first years of the initiative concluded that:

> The typical outcome measures used in public accountability systems provide very limited indicators of student and school performance. The use of state- and district-mandated assessments is likely to present a constant hindrance to understanding the effect of innovative approaches unless alternative indicators and assessments are developed in ways that align well with the reforms' goals. The high-stakes testing regimes currently in vogue and the overwhelming emphasis given to scores on state- and district-mandated tests as the measures of student improvement do not bode well for many innovative reform efforts.

27. At the same time, RAND researchers noted the vital importance of monitoring progress, school and teacher self-evaluation, and assessment of outcomes aligned with goals of the reform (RAND, 2002).

28. Alignment is also important in systems with high-stakes university entrance examinations. While education systems may set out reforms and promote innovation, high-stakes university examinations may create a "shadow curriculum" (Andrew, 2007). Sinka and Köpataki (2008), for example, report that in Hungary "[t]he accepted explanation for not implementing changes in classroom practice was that due to the unchanged examination requirements the tense pace of preparing for the secondary school-leaving entrance exams did not make it possible…."

29. No system can achieve perfect alignment. If standards are poorly designed or not specific enough, teachers are more likely to focus on tests, thus narrowing the focus of teaching (Stecher *et al.*, 2000). Baker (2004) points to the complexity of school systems as a barrier to tight alignment – including the number of links across different components of school systems, the diversity of regional contexts, differences in organizational contexts of schools, the range of teacher and school leader capabilities, the level resources devoted to professional development and other support for teachers. Given this complexity, it is very difficult to develop clear and explicit relationships across standards, instruction and assessments. Moreover, all assessments and examinations contain some degree of error, and as Haertel and Lorié (2004) assert, can only provide "an imperfect estimate of student performance". At the same time, teachers' and school leaders' perceptions of the quality of assessments and examinations make a difference in their willingness to refer to test results to improve their teaching (Russell and Sloan McCombs, 2006).

30. There may also be a misalignment of views on the role and purpose of assessments. While current rhetoric focuses on the importance of helping all students to succeed, many of those now leading education systems – policy-makers, school leaders, teachers (as well as parents) likely have grown up in and had deep experiences of systems where assessment was a tool for sorting and selecting students. Moreover, many find a cognitive dissonance in high-stakes systems that promote the idea of success for all, but do not provide sufficient support to schools and teachers in achieving this goal.

3.2. The incentives of high-stakes assessments and examinations and the problem of "teaching to the test"

31. While systems with high-stakes assessments and examinations have indeed succeeded in focusing teachers and school leaders' attention on high priority standards (or the requirements of university entrance examinations), they also provide incentives for teachers to "teach to the test". Teaching to the test might mean that teachers *re-allocate* time and *re-align* priorities in order to spend more time on content likely to be covered in the tests, thus narrowing curriculum. They may also *coach* students in test-taking skills (Cizek, 1998; Koretz *et al.*, 2001; Koretz, 2005; Popham, 2002; Smith and Rottenberg, 1991; Stecher, 2002).

32. To the degree that high-stakes assessments and examinations feature knowledge and skills deemed to be central to student learning, re-allocation and re-alignment are desirable (and this is the aim of Standards Based Assessment systems). The problem is that no single test can measure proficiencies in any given domain exhaustively. This is because they take place over a very limited period of time and therefore cannot cover all learning priorities. Assessments are meant to serve as proxies for wider achievement. But when teachers focus narrowly on content most likely to be on the assessment, they no longer serve this purpose.

33. Moreover, in most cases, large-scale assessments do not test students' ability to draw upon more complex skills for problem solving, collaboration, and so on. These are, of course, the kinds of skills commonly emphasised in educational innovations. Tests are often limited to formats that are easiest to score and most cost-efficient to implement, such as multiple choice or short essay questions, so students may only draw upon a limited set of skills.

34. Research in the United States strongly suggests that educators, particularly in schools that are under the greatest pressure to show improvement, are teaching to the test, not the standards (Stecher *et al.*, 1998; Stecher *et al*., 2000; Pedulla *et al.*, 2003). Accountability tests, thus, are the lens through which the standards are interpreted and serve to define the standards. In addition, if standards are not well-designed, or if they lack specificity, teachers will also have an incentive to teach to the test (Koretz, 2005; Linn, 2005).

35. Coaching, another way in which teachers respond to external assessments and examinations, is never a desirable approach. Coaching may focus on substantive and/or non-substantive aspects of tests. For example, if teachers notice certain patterns in tested content and then prepare students to focus on that content, they are engaging in "substantive" coaching. Substantive coaching may border on cheating if teachers provide students with advance access to tests, or provide inappropriate assistance during the examination session. When teachers help students to develop test-taking tricks (*e.g.* how to recognise distractors in multiple choice tests, while not necessarily learning how to recognise the correct answers), they are engaging in "non-substantive" coaching[4] (Koretz, 2005; Popham, 2002).

36. Re-allocation, re-alignment and coaching all lead to score inflation. The term "score inflation" refers to test score gains that overstate improvements in student learning, and therefore make it difficult for those interpreting results to identify meaningful gains in student learning. This is true even when the test focuses on content considered as central to the domain. To the degree that teachers believe it is necessary to re-allocate time, align content, or coach students, they are favouring test performance over achievement of learning goals[5]. But in high-stakes environments, teachers may find that they are under pressure to jettison innovative programmes in order to focus on test preparation.

37. According to the RAND Corporation "Implementing Standards-Based Accountability" (ISBA) study, conducted between 2004 and 2006, teachers are more focused on standards, but also using less desirable strategies to improve test scores. The study included a survey of teachers of mathematics and science, district superintendents and school leaders in the states of California, Georgia and Pennsylvania and case studies[6]. The study found that:

- The majority of teachers surveyed reported that they had increased their focus on topics emphasised in state test, as well as on the style and formats of tests. They also spent time on developing students' test taking skills. (These approaches are associated with score inflation, and can reduce the validity of inferences based on test results) (Stecher and Hamilton, 2006). Percentages of teachers who said that they focused their own tests on certain formats ranged from approximately 20 to 40% across the three states (Hamilton and Berends, 2006).

- A positive finding was that about two-thirds of mathematics teachers in this three state survey reported that they focused more on central standards when looking for more effective ways to teach as a result of the state assessment system (Stecher and Hamilton, 2006). Some teachers and principals praised the focus on standards as an effective way to ensure that all students were exposed to the same material. In most cases, however, fewer than half of the teachers surveyed agreed that the tests were good measures of the standards, raising concerns that teachers deferred to tests, rather than standards, to determine what should be taught (Hamilton and Berends, 2006).

- In all states included in the study, teachers reported increased instructional time for core subjects of state assessments, mathematics and English/language arts. This pattern did not hold for science (with the exception of lower secondary schools in Pennsylvania). Increases in non-tested subjects were minimal (Hamilton and Berends, 2006).

- Teachers were concerned that the need to cover large amounts of material meant that there was little time for students to master new knowledge and skills (Hamilton and Berends, 2006).

- Many school districts had implemented pacing plans, outlining the order and schedule for teaching to ensure that standards were covered. Not all teachers view the pacing plans favourably (Stecher and Hamilton, 2006).

38. For those schools and teachers that manage to sustain innovative teaching programmes, the news may be positive. Some studies have shown that students in programmes featuring teaching of higher-order skills perform as well, or outperform peers in programmes focused on test preparation, or transmission modes of teaching. Newmann *et al.* (2001) in a three-year study of schools in the American mid-west, found that students in classrooms featuring constructivist approaches (that is, disciplined inquiry, as opposed to reproduction of knowledge) out-performed students in more traditional classrooms[7]. Boaler (2002) conducted a three-year study of mathematics programmes in two secondary schools in England. The schools were matched in terms of students' prior achievement. However, one school emphasised the development of higher-order thinking and student responsibility for learning, while the other emphasised practice of assessment items. The students in the "reform" school outperformed their peers on the national school-leaving examination (the General Certificate of Secondary Education, or GCSE), on average, by one third of a grade.

39. Nuthall and Alton-Lee (1995) conducted three observational studies of student learning in science and social studies in elementary and lower secondary school classrooms in New Zealand. Students took short-term and long-term (12 months) achievement tests. Selected students were asked to describe how they answered test items and to recall relevant learning for these tests. Nuthall and Alton-Lee found that students in environments where learning was aimed at developing critical thinking and analysis demonstrated longer-term retention and better capacity to apply knowledge in their daily lives.

40. In a study focused on the impact of formative assessment on student performance on external tests, Wiliam and colleagues (2004) worked with 24 mathematics and science teachers in six English schools. Following an initial six-month planning period, the teachers introduced formative assessment practices at the beginning of a new school year. Researchers measured prior attainment for students in the classes featuring formative assessment, and for at least one comparison group. Those students in the classroom featuring formative assessment out-performed their peers in the comparison groups by approximately one-half of a GCSE grade (which is based on an eight-point scale, A*-A-B-C-D-E-F-G).

41. These studies are helpful for demonstrating that effective educational innovations are not necessarily at odds with external assessments (and it should be noted that the interventions featured in the studies discussed above did not introduce radically different approaches to curriculum, such as integrated studies or cross-curricular learning, and were therefore more likely to be effectively aligned with the external assessments). Ultimately, however, the question is whether high-stakes assessments and examinations provide incentives for teachers to strengthen their focus on central standards and improve teaching, or instead, to focus more narrowly on improving test performance at the expense of deeper learning. The evidence is that many teachers are more likely to take the latter approach.

3.3. Tests worth teaching to?

42. The problem of teaching to the test can be resolved, some have argued, by ensuring that the test is worth teaching to. Certainly it is vital to ensure that tests are well-designed and aligned with broader goals for curriculum and teaching and learning. But even when systems are well aligned (and as has been noted, no system can be perfectly aligned), even the highest-quality test will only cover a subset of knowledge. Moreover, large-scale standardised tests based on short answers and/or multiple-choice question (MCQ)

may include biases[8]. Nor can these formats provide insight on students' higher order skills, such as problem solving or reasoning processes. Tests that focus only on those aspects of a subject that are easy to measure – for example, recall of facts and figures – dramatically narrow the scope of learning.

43. Performance-based assessments are also subject to score inflation if contents are not frequently changed, or if teachers ignore development of skills that are not part of the test (Koretz, 2005). So, for example, while open-ended essay questions will draw on a broader range of student skills and critical thinking abilities, if instructors spend most of their time focused only on those subjects that are most likely to appear on the test, or spending time on essay writing skills to the exclusion of speaking and listening skills, they are narrowing curriculum. No single test can capture student performance across an entire subject domain.

3.4. The impact of high-stakes testing and expectations on student (and teacher) engagement and motivation

44. Another set of potential problems relates to the impact of high-stakes testing on student engagement, motivation and performance – vital for countries' efforts to promote lifelong learning. McDonald (2001) reviews evidence on the prevalence of test anxiety and its impact on test performance for children in compulsory education in the United Kingdom and the United States[9] He concludes that "fear of exams and test situations is widespread and appears to be becoming more prevalent possibly due to the increasing frequency of testing within education systems" (p. 98). Even assessments that do not have any impact on the student's record (such as assessments for accountability), students are nevertheless likely to pick up on pressures felt by teachers. McDonald finds that while correlations between test anxiety and examination performance are modest, the influence on the number of children passing or failing an examination may be considerable.

45. Putwain (2008), in a study of 615 students in the final two years of compulsory education in the United Kingdom, found that students reported the lowest levels of anxiety and earned the highest marks in examinations with "mid stakes". Cole *et al.* (2008), in a study of 1 005 undergraduate students in mid-size public universities, found that students' efforts and performance on low-stakes assessments was related to their perception of the usefulness or importance of a test.

46. While there is evidence that high expectations may serve to motivate students and raise levels of achievement (Bransford, Brown and Cocking, 1999; Cizek, 2001; Covaleskie, 2002) expectations that all students will meet goals at the same pace are unrealistic. For example, the United States' No Child Left Behind (NCLB) legislation of 2001 established the goal for all schools to meet standards (established at the state level) by 2012. States establish annual yearly progress targets and specify the percentage of students in each school to meet proficiency, as measured in state tests. Schools that consistently meet targets may be eligible for rewards, and those that do not may be subject to sanctions. Many have questioned whether the expectation of 100% proficiency within 12 years of the passage of the legislation is realistic.

47. The RAND Corporation's ISBA described above included data on the impact of assessments on student and teacher morale. The survey found that:

- Many teachers expressed concern about NCLB's high expectations. Teachers noted challenges in addressing the needs of students from lower socio-economic backgrounds, students' inadequate basic skills and prior preparation, the wide range of student abilities within classrooms, and a lack of support from parents.

- Teachers reported that lower achieving students were sometimes pulled out of classes such as band, physical education or art in order to increase time spent in tested subjects. Many of those

surveyed were concerned that students would lose motivation or disengage from school as a result. (These findings are consistent with the Center on Education Policy's (2006) findings, based on a national sample of districts) (Hamilton and Berends, 2006).

- Because of the NCLB's requirement that school performance be measured according to a percent-proficient metric, schools focus on those students who are close to the passing score. Raising these students' scores has a significant impact on ratings of school performance. Approximately half of the teachers across this three-state survey agreed that higher achieving students were not getting appropriately challenging instruction as a result of the intense focus on the students in the middle (see Hamilton and Stecher, 2006, Hamilton and Berends, 2006). Higher-achieving students may also lose motivation under these circumstances.

- The same RAND survey found majorities of teachers reporting that high-stakes tests had a negative impact on the morale of school staff (Hamilton, Berends and Stecher, 2005).

48. The United Kingdom-based Assessment Reform Group (ARG) has been particularly concerned with the impact of high-stakes tests on student motivation. A review commissioned by the ARG review identified 19 studies providing empirical evidence related to this question. The studies found strong evidence of a negative impact on student motivation (although the impact varied according to student characteristics as well as the conditions of learning). Moreover, the tests had a negative impact on the self-esteem of students who did not perform well on the tests (ARG, 2002).

49. Harlen and Deakin-Crick (2003) found that, following the introduction of the National Curriculum tests in England, low-achieving students reported lower self-esteem than higher achieving students, whereas prior to the tests there was no correlation between self-esteem and achievement. The 11+ tests in Northern Ireland have been found to have a particularly strong impact on underachieving students. Harlen and Deakin-Crick argue that tests that have high stakes for schools but not for students may also have a strong impact on student self-esteem. Tests may also increase the gap between high and low achievers. High achievers are confident, use more appropriate test taking strategies, and may enjoy taking tests. Low achievers, on the other hand, are de-motivated by constant evidence of their low achievement.

50. Critics in both Japan and Korea have expressed concern that intensely competitive university entrance examinations are too stressful for students. Lee and Larson (2000), found that higher rates of clinical depression among Korean adolescents (as compared to their American peers) were related to the competitive college entrance examination. Using time-sampling data for 56 high school seniors in Korea and 62 seniors in the United States, they found that Korean adolescents spent much more time on schoolwork and less time in discretionary activities than their American counterparts, and that their "affect states" across their daily activities (schoolwork, social activities and passive leisure activities) were more negative, and related to higher depression, as compared to their American peers.

3.5. Implementation challenges

51. Challenges of implementation are also related to the "how to" of improving teaching and learning. Assessment reformers have argued that in some countries, high-stakes Standards Based Assessment systems have placed too much focus on educational outcomes, but ignored the processes needed to achieve those outcomes (ARG, 2002; Black, 2001; Black and Wiliam, 1998). This is particularly important if systems are to help promote success for all students, and not just the select few. This is not, it should be emphasised, an argument to return to a strong, prescriptive focus on educational inputs. Rather, the focus is on providing teachers with more opportunities for professional development, as well guidelines and tools.

52. To be sure, many education systems do call for the use of assessment and examination data to provide feedback to school leaders in teachers. In the case of standards-based assessments, test results may provide valuable information on student performance, and help pinpoint learning needs of subgroups or individual students. In some countries, schools receive feedback based on the results of student examinations for entrance to higher levels of education[10]. These data may be useful for helping teachers adapt instruction for incoming groups of students. However, data from university entrance examinations are likely to be of limited use for those engaged in developing and implementing educational innovations. School leaders and teachers also need skills to interpret – and respond to – assessment and examination results.

53. The usefulness of assessment data for improving instruction or shaping innovative classroom practices depends upon several factors. Data need to be broken down to a useful level of detail. Data also need to be provided in a timely manner. For example, data from summative examinations administered toward the end of students' tenure at that particular school, will have little impact for the cohort just tested. Results from national assessments that are delivered late in the school year may be useful for policy decisions related, for example, to resource allocation, but will be of only limited use to instructors. (In France, as well as in French-speaking Belgium, national assessments are conducted at the beginning of the school year and are intended solely to serve purely diagnostic purposes.)

54. School leaders and teachers also need skills to interpret, and to respond, to data. They may not have knowledge of effective alternatives – including new and innovative approaches – to improving instruction. While most OECD countries encourage innovation, there is much less attention to systematic dissemination of new knowledge on effective practices, or investment in capacity building of individual teachers or teacher networks (Hargreaves, 2003; Fullan, 2001; Newmann *et al.*, 2000).

55. School leaders and teachers must also deal with real tensions between the accountability goals of high-stakes assessment and the use of data to improve teaching and learning (the empowerment agenda). In principle, goals for assessment for accountability and assessment for learning are compatible. But in countries with high-stakes testing, the political and public focus on accountability is much stronger, and efforts to promote a culture of assessment for learning lag behind. Moreover, it appears that the culture of high-stakes testing pervades schools, even though external tests are usually administered only once for students in primary school, and once for students in secondary school. For example, American researchers Hamilton and Berends have found that even teachers in non-tested grades made changes in their teaching to improve overall school performance, and perhaps prepare students for future testing (Hamilton and Berends, 2006).

56. Those engaged in implementing innovative education programmes face additional pressures. Innovation, by definition, involves risk-taking. In high-stakes environments, however, school leaders and teachers may be less willing to take risks, particularly if they face potential sanctions as a result of poor test performance. In the United Kingdom, the Qualifications and Curriculum Authority (QCA) concludes, on the basis of the school evaluations for which it is responsible, that while schools may wish:

> …to innovate and/or encourage creativity within the curriculum, many …are inhibited from doing so through anxiety over attaining and/or sustaining high levels of achievement. The factors most commonly identified as inhibitors are challenging targets and league tables (QCA, 2002 [cited in Gipps, 2003]).

57. Most new programmes experience an "implementation dip" – that is, student performance gets worse before improving. Improvements in student achievement may take as long as five years in primary schools, and longer in secondary schools (Fullan, 2001). Teachers working in innovative programmes will need extra support to understand where they may need to adjust practices.

58. To conclude this section, there are several ways in which high-stakes assessments and examinations may undermine innovative educational programmes. While high-stakes do help to focus many school leaders' and teachers' attention on central standards (or on the requirements for school exit or entrance to higher levels of education), they are also more likely to teach to the test. Focus on a single high-stakes test may, in turn, lead to a significant narrowing of curriculum. This problem is compounded when standards are poorly designed, or not well-aligned with curriculum, assessment or examinations.

59. Innovative programmes are likely to face additional barriers. Systems may achieve effective alignment for traditional curriculum, but miss important learning goals emphasised in innovative programmes. In addition, high-stakes may discourage risk-taking necessary for innovation, and may de-motivate students (particularly underachieving students) and teachers. Teachers may find it impossible to balance the pressures of implementing new and innovative programmes and high-stakes tests. Those that do persist with pilot projects may find it more difficult to champion new and effective approaches more widely.

4. Is it possible to reconcile tensions between innovation and high-stakes assessment?

60. This section examines ways in which systems might better reconcile assessment and innovation. Over the last several years, researchers have made a number of advances in testing methodology. Several OECD countries have also introduced a range of approaches to measuring school and student performance. There are opportunities providing opportunities for learning about effective assessment and examination approaches across countries.

61. It is worth emphasizing that accountability systems in OECD countries, while imperfect, have had several positive impacts. Stakeholders in education systems have more clearly defined and updated national or regional goals for education. They have also set out expectations for greater equity of student outcomes. All schools are expected to help children to meet high standards. There is a stronger focus on using data gathered to improve teaching and learning.

62. The task for OECD countries is to preserve the best features of assessments and accountability, while addressing shortcomings. There are a number of ways (not mutually exclusive) to better reconcile innovation and assessment. Some important steps include:

- promoting innovative approaches to assessments and examinations;
- rethinking alignment;
- using multiple measures of student and school performance to evaluate quality and outcomes;
- creating a better balance between high-stakes and school empowerment.

63. The longer-term success of these strategies will require a fundamental change in thinking about the role of assessments and examinations, and the balance of accountability and school empowerment and improvement.

4.1. Promoting innovative approaches to assessments and examinations

64. Potentially, one of the most effective ways to reconcile educational innovation and high-stakes assessments is through the introduction of innovative approaches to testing. These may include new methodologies for assessment, as well as new technologies that can measure complex skills and reasoning processes.

65. The most important advances are at the conceptual level. Traditional tests, Pellegrino *et al.* (1999) have argued, do not provide information on higher order thinking because they focus on independent "traits", or fragments of knowledge, rather than on the student's understanding of the inter-connections and patterns. Continuing this line of argument, Black (2001) suggests that data from assessment are of little use unless they are understood in terms of a model of pupil's learning. Otherwise, it is difficult to draw inferences about a student's learning needs or potential.

66. Mislevy and colleagues (1998) suggest that tests should be developed with important inferences users will want to make established at the beginning of the process (as is the case with tests that focus on content knowledge). For example, test developers might want to measure students' cross-curricular skills, such as problem solving, communication, and reasoning processes. The next step would be to identify how students use these skills in different subject domains (mathematics, science, analysis of literature, and so on). Test developers may then consider the kinds of performance tasks that provide evidence of student capabilities, including performance-oriented tasks. Since this approach focuses on cognitive demands, rather than specific content, it would open up opportunities to apply innovative strategies for teaching and learning.

67. Several innovations in assessment take advantage of recent advances in ICT. ICT-based assessments can now provide efficient feedback, measure students' reasoning processes and other complex cognitive skills, score "open-ended" performances such as student essays, and even examine how students go about thinking through problems[11]. While it has always been possible for teachers or external assessors to perform these functions, ICT offers the possibility for large-scale and more cost-effective assessment of complex skills (Mislevy *et al.*, 2001). Technology-based assessments may also incorporate simulation, interactivity, collaboration and constructed response formats. For example, students may use the multi-media functions of ICT to show how they would perform a physics experiment or some other problem-solving task. Students may use Internet–based programmes to "predict-observe-explain" specific concepts. Or, they may develop concept maps using online tools to show their understanding of processes. The student's map is scored by comparing it against an expert map (Bennett, 2001). In another example of ICT-based assessment, students are scored on their use and judgment of information on the Internet. Students may obtain feedback on their work in real time. Test developers have found that students respond positively to Internet tasks (searching, and judging the quality of on-line information) and are engaged, even in difficult, open-ended tasks (Lewis, 1998).

68. Increasingly sophisticated ICT programmes that score "open-ended performances", such as essays, are under development (see Chung and Baker, 2003; Chung *et al.*, 2001; Herl *et al.*, 1999). These programmes use natural-language processing, artificial intelligence and/or information retrieval technologies to detect of textual features of essays (for example, variety in use of syntax, quality of content and organisation of ideas). These ICT models are still in the relatively early stages of development however, and while they may facilitate scoring of large-scale assessments, cannot replace human raters. Further studies are also needed to determine the validity and generalisablity of different automated essay scoring tools (Wang and Brown, 2007).

69. The cost of designing and implementing effective ICT-based assessments and examinations has been an important barrier to wider use. However, as noted at the beginning of this paper, all OECD countries have invested significant levels of funding in ICT and related teacher professional development over the past several years. ICT-based assessments may help ensure that schools realise the value of these investments. Moreover, school systems will also realise savings if ICT-based assessments and examinations provide better information on the quality of teaching and learning, and provide information to support school improvement and effective innovation.

70. Of course, ICT-based technology is not a pre-requisite for large-scale assessment of students' open-ended performance. The most effective ICT programmes will be modelled on processes followed by human raters (as illustrated in the case of Senior Certificate assessments in Queensland, Australia as described in Box 2). These assessment processes also serve as valuable professional development for the raters, helping to fulfil goals for improvement and empowerment.

Box 2. Queensland, Australia

There is no whole-cohort external testing or examining in secondary schools in Queensland. In 1972, Queensland abolished external examinations and replaced them with a system of moderated internal assessments.

School-based assessments for the Senior Certificate (year 12) are currently moderated for those subjects that count towards university entrance. Moderation aims to ensure consistency, comparability and equity of teacher judgments of student performance standards. The moderation processes for the Senior Certificate involve subject-based panels of expert teachers providing advice to schools on the quality of their assessment program and their judgments of quality of student performance based on sample portfolios. The system involves follow-up where panels identify difficulties. There is negotiation of the final results to be recorded on the Senior Certificate. Results are expressed in terms of five relative grades or 'levels of achievement' expressed in terms of standards descriptors (referred to as "exit standards").

Source: Sebba and Maxwell (2005).

4.2. Rethinking alignment

71. Standards Based Assessment approaches discussed in this paper all depend on "tight alignment" of standards, curriculum and external tests. Tight alignment requires that content standards clearly specify all instructional goals, and that each goal be validly measured. But, as Linn points out, such "complete congruence" is not realistic. Most studies on alignment focus instead on the degree to which it is approximated, rather than fully achieved (Linn, 2005). Moreover, tight alignment, as we have seen in this report, tends to undermine innovative programmes where goals for learning go beyond those specified in standards and given priority in high-stakes tests.

72. Baker has suggested that the concept of alignment probably should not be interpreted so literally. Rather than looking for complete congruence across standards, curriculum and assessment, raters might use broader frameworks to judge the quality of student performance. For example, they might use rubrics or exemplars outlining criteria and illustrated expected standards of performance at different levels (these tools may also be used formatively, to identify student needs and to adapt teaching and learning strategies). Standards for judging student performance are therefore clear, but are also broad enough to be used in a variety of situations, including in innovative programmes. The Equipped for the Future (EFF) model in the United States takes this approach (see Box 3 below) (Baker, 2002, 2004; EFF Assessment Consortium, 2005). EFF has been developed primarily for adult foundation skill learners, but could easily be applied to school-level assessment.

Box 3. Equipped for the future

The Equipped for the Future (EFF) programme was initiated in 1990 to help States clarify objectives for adult literacy programmes, to strengthen the validity of assessments, and to ensure better alignment across teaching, learning and assessment. The EFF framework includes sixteen integrated content standards defining core knowledge and skills for literacy. The framework describes progressive levels of performance for each of the standards, and identifies a variety of assessment tools. The tool may be used in a variety of learning contexts.

One of the primary advantages of EFF is that it provides tools for learning as well as for assessment. Both instructors and learners may use the EFF to track progress toward learning goals (a formative approach to assessment). Learners report that they better understand what they are doing and why, and are more active in planning their own learning, identifying criteria to judge the quality of their performance, and assessing their accomplishments. The EFF website provides a variety of assessment tools and specific examples of how they may be used (http://eff.cls.utk.edu/toolkit/examples.htm).

Adult literacy instructors have also been positive about the impact of the EFF on their own professional development. They have used the framework to monitor their own performance, to develop plans for programme improvement.

Source: Center for Literacy Studies (2009) : http://eff.cls.utk.edu/; Stein (1995).

73. Another approach to alignment might be to use *random assessment* to track performance of the education system, rather than to hold individual schools accountable for student performance. With this approach, a small percentage of randomly selected schools participate in assessments that measure student learning across the curriculum. Advocating this methodology, Wiliam (2001) suggests that random assessments should include a large number of tasks, although not all students would be tested the same task. Teachers would not be able to narrow their focus to those areas most likely to appear on the assessment (that is, to teach to the test), because the entire curriculum would be assessed. Systems would still monitor individual schools through self- and external-school evaluations, but the pressure on high-stakes testing for accountability purposes would be relieved.

Box 4. Assessment and evaluation in Finland

At the national level, monitoring and evaluation focus on the extent to which the objectives set in statutes, education policy decisions and national core curricula are achieved. Finland does not have an inspectorate. Moreover, there are not any national exams except for the matriculation examination to finalise upper secondary general education. This is why the National Board of Education conducts national assessments. The aim is that each comprehensive school participates in national assessments of different subjects every third year for a random sample of students. The objective is to develop a general picture of the state of education.

Each school that participates in the assessment receives feedback on their own results, as well as the national averages. At the national level, the results of these evaluations shape the education strategies and the core curricula. The national evaluation system of education consists of three sections:

- Learning outcomes;
- Long-term educational trends;
- Evaluation of selected issues.

Source: Voogt and Kasurinen (2005).

74. This approach would also provide more accurate information about teaching and learning, and where policy makers need to direct attention. For systems promoting school choice, parents and student would be more likely to look to a variety of data sources for information on school quality, and would therefore have a better view on school performance than that provided in comparative tables listing assessment and examination scores. Currently, Finland and Korea take a random approach to national assessment of student performance.

4.3. Using multiple measures of student and school performance to evaluate quality and outcomes

75. In a high-stakes environment, it is particularly risky for systems to rely on a single test score as a measure of school, teacher or student performance, no matter how well-designed the test. A test can provide an important snapshot of student performance, but no single test can provide enough information to fully understand how instructional strategies are influencing student learning. Moreover, students will perform differently on different days; different assessment tools will provide different ways for measuring student capabilities and needs, and may therefore yield different results.

76. As noted above, in standards-based assessment systems, high-stakes tests are used not only to hold schools accountable for student performance, but are also intended to provide data teachers may use to improve instruction. These data are most useful when they are used in combination with data gathered at the school and classroom level. Schools and systems that keep good records, and school leaders and teachers who know how to interpret results will be able to respond more effectively. Teachers who have more data available will be better able to identify the source of misunderstandings or misconceptions (insufficient prior knowledge, thoughts not well-organised, etc.) (Baker, 2003). Data are also important for monitoring implementation (a particularly important function for innovations), and where instructors may need to adjust their practice, or to re-think the strategy of the innovation.

77. *School evaluations*, which include any kind of systematic description of school performance provided within an evaluative framework, also provide vital data. They include school self-evaluation or external evaluations, audits, benchmark data, or external inspections (OECD, 2008). School self-evaluations may be carried out by school staff. School leaders and teachers may need training to gather and interpret school and programme level data. Nevertheless, staff benefit from their direct involvement in the evaluation process (Monsen, 2002; Simmons, 2002; Lander and Ekholm, 1998). Results may also be used to provide information on school performance to parents and local communities, and to improve performance (OECD, 2008). (See Annex 1 for a summary of OECD countries approaches to assessment, examinations and internal and external school evaluations. Annex 2 provides additional details.)

78. Several OECD systems also provide support for classroom-based *formative assessment* as a way to improve teaching and learning. These may include “progress tests”, aligned with high-stakes assessments (Stecher and Hamilton, 2006). Other countries provide guidelines, tools and professional development in support of classroom-based formative assessment[12] (OECD, 2005). These classroom-based formative assessments are aimed directly at improving the quality of teaching and learning, and there is evidence that they may support substantial learning gains, particularly for underachieving students. Black and Wiliam (1998), in their highly influential review of literature described the gains associated with formative assessment as “among the largest every reported for educational interventions”. Teachers using formative assessment identify the factors behind variations in student achievements, and adapt teaching to meet diverse student needs. The focus is on helping all students to achieve at high levels. Students also develop skills to assess their own and their peer’s work, and develop strategies for “learning-to-learn”. These are the skills vital for problem-solving and the kind of higher-order thinking emphasised in innovative programmes.

79. Multiple measures of student and school performance help to ease the pressure on teachers and students to perform well on a single, high-stakes test. (The same principle holds true for high-stakes university entrance examinations: universities that solicit a variety of information on student capabilities, such examination results, marks, interviews, recommendations, will relieve the pressure on students to perform well on a single high-stakes examination). But if multiple measures are really to increase validity and reliability of educational measurement, then systems need to address a number of questions. For example, how much redundancy should systems include? How should different assessments be weighted? How can complementary measures help systems better fulfil goals to improve policy, school leadership and classroom instruction? (Baker, 2003; OECD, 2005).

80. As with high-stakes assessments and examinations, teachers, school leaders and policy-makers need training to effectively interpret and make use of data gathered in various forms and formats. Education stakeholders must also support the need to invest in a system of multiple measures. Certainly the cost of implementing multiple measures is higher. But the cost of relying upon faulty or inadequate information is much higher.

4.4. Creating a better balance between high-stakes and school empowerment

81. The strategies outlined above do not suggest that those systems that now have high-stakes assessments and examinations simply lower or eliminate those stakes (although many countries that do not have high-stakes assessments and examinations perform well on international assessments, such as the OECD's Programme for International Student Assessment). Such a recommendation would not likely gain any real traction in those countries that now use high-stakes assessments, as policy makers and the general public continue strong support for this approach. In the case of school and university entrance examinations, high-stakes will persist in the form of competition for limited places in prestigious institutions.

82. Rather, the three strategies outlined above – the development of more effective assessments and examinations, more sophisticated approaches to alignment, and the use of multiple measures of school and student achievement – can go a long way toward improving the quality of information gathered and incentives given to all stakeholders. These strategies may also relieve some of the incentive to "teach to the test".

83. A vital fourth strategy is to ensure schools and teachers are empowered to take the risks entailed in innovation. School leaders and teachers may need extra support as they work to keep new programmes on track while also responding to the demands of high-stakes testing. New programmes may experience growing pains, and teachers will need to be able to identify problems and make appropriate adjustments. Indeed, as Foray and Hargreaves (2003) have argued, innovation within the education sectors occurs through "learning-by-doing", where "…individuals learn as they go along, assess what they learn and hone their practices over time" (p. 8). They refer to this as a "humanistic" mode of knowledge production. The goal should be to learn from failure and to innovate for improvement.

5. Conclusion

84. This report has focused on impact of high-stakes assessments and examinations on educational innovation. It has found significant evidence that such assessments and examinations undermine innovation. Teachers may respond to the incentives of high stakes by teaching to the test, rather than focusing on the goals of the innovative programme. Countries with highly competitive examinations for university entrance refer to the "shadow curriculum" shaped by these examinations. These problems are intensified when tests are poorly designed, or misaligned with the goals of innovative programmes.

Moreover, there is evidence that high-stakes tests promote student anxiety, particularly underachieving students, further undermining goals to motivate students' desire to engage in lifelong learning.

85. Some studies have found that students in classrooms featuring the kind of higher order skills emphasised in innovative programmes may perform as well or better than their peers in traditional classrooms on external tests. However, newer innovative programmes may not show positive results for the first five years, and may even experience an "implementation dip". It is difficult for programmes to sustain their commitment to innovative programmes when they are under pressure to show positive results quickly.

86. It is possible to reconcile high-stakes assessments and examinations through innovative approaches to testing. For example, tests that measure student's understanding of inter-connections and patterns will provide more opportunities for teachers introduce use innovative strategies in the classroom. ICT-based assessments provide ways to measure complex skills and reasoning processes on a larger-scale. There are also effective approaches using practitioner panels to score open-ended student performances on a large-scale basis. Systems may also re-think approaches to alignment of tests with standards and curriculum, again allowing opportunities to score students on open-ended performances. The use of multiple measures of student and school performance may also relieve some of the pressure of high-stakes assessments, and also ensure that data are valid and reliable. Systems will have access to richer information, and better insight regarding the impact of innovative programmes on teaching and learning.

87. Finally, if systems are to balance accountability with goals to empowering schools and teachers to innovate, they will need to make serious and sustained investments in supporting school leaders and teachers. This may involve support for professional development, networks, and deeper research on innovation. It will also require a greater risk-tolerance and ability to accept some level of failure as teachers learn more about what works and doesn't work. This holds true for the development of new and innovative assessments and examinations, as it does for classroom practice.

REFERENCES

Almond, R.G. and R.J. Mislevy (1999), "Graphical Models and Computerized Adaptive Testing", *Applied Psychological Measurement*, Vol. 23, pp. 223-237.

Andrew, C. (2007), "Seventh National Curriculum: Policies for Pre-College Education" in *Dynamic Korea: Education Policies and Reform*, Group Project: EPS530Z- Spring 2007. http://globalizationandeducation.ed.uiuc.edu/Students%20Projects/GSEB/2007/South%20Korea2007.pdf.

Assessment Reform Group (2002), *Testing, Motivation and Learning*, Faculty of Education, University of Cambridge, www.assessment-reform-group.org.uk.

Baker, E. (2002),"The Struggle to Reform Education: Exploring the Limits of Policy Metaphors", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Baker, E. (2003), "Multiple Measures: Toward Tiered Systems", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Baker, E. (2004), "Aligning Curriculum, Standards, and Assessments: Fulfilling the Promise of School Reform", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Bennett, R. E. (2001), "How the Internet will Help Large-Scale Assessment Reinvent Itself", *Education Policy Analysis*, Vol. 9, No. 5. http://epaa.asu.edu/epaa/v5n9.html.

Black, Paul, (2001), "Dreams, Strategies and Systems: Portraits of Assessment Past, Present and Future", *Assessment in Education: Principles, Policy and Practice*, Vol. 8, No. 1, pp. 65-85.

Black, P. and D. Wiliam (1998), "Assessment and Classroom Learning", *Assessment in Education: Principles, Policy and Practice*, Vol. 5, No. 1, pp. 7-74.

Bloom, B.S. (ed.) (1956), *Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain*, Longmans, Green, New York and Toronto.

Boaler, J. (2002), *Experiencing School Mathematics: Traditional and Reform Approaches to Teaching and their Impact on Student Learning*, Lawrence Erlbaum Associates, Mahwah, NJ.

Bransford, J.D. Brown, A. L. and Cocking, R. R. (eds.) (1999), *How People Learn: Brain, Mind, Experience, and School*. National Academy of Sciences, National Academy Press, Washington D.C.

Center on Education Policy (2006), "From the Capital to the Classroom: Year 4 of the No Child Left Behind Act", Washington, D.C.: Author. http://www.cep- dc.org/nclb/Year4/Press/.

Center for Literacy Studies (2009), Equipped for the Future Collection, University of Tennessee, http://eff.cls.utk.edu/.

Chung, G.K.W.K. *et al.* (2001), "Knowledge Mapper Authoring System Prototype", (Final deliverable to OERI), University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Chung, G.K.W.K. and E.L. Baker (2003), "Issues in the Reliability and Validity of Automated Scoring of Constructed Responses" in M.D. Shermis and J. Burstein (eds.), *Automated Essay Scoring: A Cross-Disciplinary Perspective*, pp. 23-40, Lawrence Erlbaum, Mahwah, NJ.

Cizek, G. J. (1998), *Testing in American schools: Getting the right answers*. Lawrence Erlbaum Associates, Mahwah, NJ.

Cizek, G.J. (2001), "More Unintended Consequences of High-Stakes Testing" *Educational Measurement: Issues and Practice*, Vol. 42, No. 1, pp. 75-83.

Cole, J.S. *et al.* (2008), "Predicting Student Achievement for Low-Stakes Tests with Effort and Task Value", *Contemporary Educational Psychology*, Vol. 33, No. 4, pp. 609-624.

Covaleskie, J.F. (2002),"Two Cheers for Standardized Testing", *International Electronic Journal for Leadership in Learning*, Vol. 6, No. 2. Retrieved 18 February 2009 from http://www.ucalgary.ca/~iejll/volume6/covaleskie.html.

Crocker, L. (2005), "Teaching FOR the Test: How and Why Test Preparation is Appropriate" in R. P. Phelps (ed.), *Defending Standardized Testing*, Lawrence Erlbaum, Mahwah, NJ, pp. 159-174.

Edwards, W. (1998), "Hailfinder: Tools For and Experiences With Bayesian Normative Modeling", *American Psychologist*, Vol. 53, pp. 416-428.

Foray, D. and D. Hargreaves (2003), "The Production of Knowledge in Different Sectors: A Model and Some Hypotheses", *London Review of Education*, Vol. 1, No. 1, pp. 7-19.

Fullan, M. (2001), "The Return of Large-Scale Reform", *Journal of Educational Change*, Vol. 1, pp. 5-28.

Gipps, C. (2003), "Educational Accountability in England: The Role of Assessment", Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, April 21-28, 2003.

Haertel, E. H., and W.A. Lorié (2004), "Validating Standards-Based Test Score Interpretations", *Measurement: Interdisciplinary Research and Perspectives*, Vol. 2, pp. 61-103.

Haertel, E. and J. Herman (2005), "A Historical Perspective on Validity Arguments for Accountability Testing", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Hamilton, L.S. and M. Berends (April 2006), "Instructional Practices Related to Standards and Assessments", Document No. WR-374-EDU.

Hamilton, L. S. and B.M. Stecher, B. M. (2006), "Measuring Educators' Responses to High-Stakes Testing", Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 7-11, 2006.

Hamilton, L. S. *et al.* (2005), *Teachers' Responses to Standards-Based Accountability*, RAND, Santa Monica, CA.

Hargreaves, D. (2003), *Education Epidemic: Transforming Secondary Schools Through Innovation Networks*, Demos, London.

Harlen, W. and R. Deakin Crick (2003), "Testing and Motivation for Learning", *Assessment in Education: Principles, Policy and Practice,* Vol. 10, No. 2, pp. 169 - 207.

Herl, H.E. *et al.* (1999), "Reliability and Validity of a Computer-Based Knowledge Mapping System to Measure Content Understanding", *Computers in Human Behavior*, Vol. 15, pp. 315-334.

International Association for the Evaluation of Educational Achievement (IEA) (1999), *1999 Video Study of Eighth-Grade Mathematics Teaching*, http://timssandpirls.bc.edu/ [video studies available at: http://nces.ed.gov/timss/video.asp]

Jacob, B. (2002), "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools", Working Paper W8968, National Bureau of Economic Research, Cambridge, MA.

Jensen, F.V. (1996), *An Introduction to Bayesian Networks*, Springer-Verlag, New York.

Klein, S. P. *et al.* (2006), "What Do Test Scores in Texas Tell Us?", Issue Paper IP-202, RAND, Santa Monica, CA.

Koretz, D. *et al.* (2001), "Toward a Framework for Validating Gains under High-Stakes Conditions" University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Koretz, D. (2005), "Alignment, High Stakes, and the Inflation of Test Scores", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Lander, R. and M. Ekholm (1998), "School Evaluation and Improvement: A Scandinavian View" in A. Hargreaves, A. Lieberman, M. Fullan and D. Hopkins (eds.), *International Handbook of Educational Change*, Kluwer Academic Publishers, Dondrecht, the Netherlands, pp. 1119-1134.

Lee, M. and R. Larson (2000), "The Korean 'Examination Hell': Long Hours of Studying, Distress and Depression", *Journal of Youth and Adolescence*, Vol. 29, pp. 249-271.

Lewis, A. (1998), "Assessing Student Achievement: Search for Validity and Balance", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Linn, R. L. (2001), "The Design and Evaluation of Educational Assessment and Accountability Systems", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Linn, R.L. (2005), "Issues in the Design of Accountability Systems", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

McDonald, A. (2001), "The Prevalence and Effects of Test Anxiety in School Children", *Educational Psychology*, Vol. 21, No. 1, pp. 89-101.

Mislevy, R.J. *et al.* (2001), "Making Sense of Data From Complex Assessments", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Mislevy, R.J. (1994), "Evidence and Inference in Educational Assessment", *Psychometrika*, Vol. 59, pp. 439 – 483.

Mislevy, R.J. *et al.* (1998), "A Cognitive Task Analysis, with Implications for Designing A Simulation-Based Performance Assessment", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Monsen, L.I. (2002), "School-Based Evaluation in Norway: Why Is It So Difficult to Convince Teachers of its Usefulness?" in D. Nevo (ed.), *School-Based Evaluation: An International Perspective*, JAI Press, Oxford, pp. 73-88.

Newmann, F. *et al.* (2000), "Professional Development That Addresses School Capacity", Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 24-28, 2000.

Newmann, F. *et al.* (2001), *Authentic Intellectual Work and Standardized Tests: Conflict or Coexistence?*, Consortium on Chicago School Research, Chicago, IL.

Nuthall, G. and A. Alton Lee (1995), "Assessing Classroom Learning: How Students Use Their Knowledge and Experience to Answer Classroom Achievement Test Questions in Science and Social Studies", *American Educational Research Journal*, Vol. 32, No.1, pp. 185-223.

OECD and Eurostat (2005), "Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data", Third Edition, OECD and Eurostat, Paris.

OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris.

OECD (2007), *Understanding the Social Outcomes of Learning*, OECD Publishing, Paris.

OECD (2008), *Education at a Glance: OECD Indicators*, OECD Publishing, Paris.

OECD (forthcoming), *Teaching and Learning International Survey*, OECD Publishing, Paris.

Pedulla, J.J. *et al.* (2003), *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers*, National Board on Educational Testing and Public Policy, Boston.

Pellegrino, J.W. *et al.* (1999), "Addressing the 'Two Disciplines' Problem: Linking Theories of Cognition with Assessment and Instructional Practice, *Review of Research in Education*, Vol. 24, pp. 307-353.

Popham, W.J. (2002), "Right Task, Wrong Tools", *American School Board Journal*, Vol. 189, No. 2, p. 18-22.

Putwain, D. (2008), "Do Examination Stakes Moderate the Test Anxiety-Examination Performance Relationship?", *Educational Psychology*, Vol. 28, No. 2, pp. 109-118.

Qualifications and Curriculum Authority (QCA) (2002), *Report on QCA's Monitoring, Evaluating and Developing the Curriculum 2001-2002*, QCA, London.

RAND Research Briefs (2002), "A Decade of Whole-School Reform: The New American Schools Experience, RAND RB-8019 (2002).

Russell, J. and J. Sloan McCombs (2006) "Superintendent, Principal, and Teacher Understanding of and Attitudes about Accountability", WR-376-EDU, April 2006.

Schumpeter, J.A. (1942), *Capitalism, Socialism and Democracy*, Routledge, London and New York.

Sebba, J. and G. Maxwell (2005), "Queensland, Australia: An Outcomes-based Curriculum", in *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD, Paris, pp. 191-204.

Simmons, H. (2002), "School Self-Evaluation in a Democracy", School- Based Evaluation in D. Nevo (ed.), *School-Based Evaluation: An International Perspective*, JAI Press, Oxford, pp. 17-34.

Sinka, E. and M.M. Köpataki (2008), *National Report on Development of Education, Hungary*, developed for the UNESCO International Bureau of Education, Budapest.http://www.ibe.unesco.org/National_Reports/ICE_2008/hungary_NR08.pdf.

Smith, M. L., and C. Rottenberg (1991), "Unintended Consequences of External Testing in Elementary Schools", *Educational Measurement: Issues and Practice*, Vol. 10, No. 4, pp. 7-11.

Stecher, B.M. and L.S. Hamilton (2006), "Using Test-Score Data in the Classroom", WR-375-EDU, April 2006, RAND Education.

Stecher, B. M. *et al.* (2000), "The Effects of the Washington State Education Reform on Schools and Classrooms: Initial Findings", DB-309-EDU, RAND Santa Monica, CA.

Stecher, B. M. *et al.* (1998), "The Effects of Standards-Based Assessment on Classroom Practices: Results of the 1996-1997 RAND Survey of Kentucky Teachers of Mathematics and Writing", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles.

Stecher, B. M. (2002), "Consequences of Large-Scale, High-Stakes Testing on School and Classroom Practices", in L. S. Hamilton, B. M. Stecher, and S. P. Klein (eds.), *Making Sense of Test-Based Accountability in Education*, RAND, Santa Monica, CA, pp. 79-100.

Stein, S.G. (1995), *Equipped for the Future: A Customer-Driven Vision for Adult Literacy and Lifelong Learning*, National Institute for Literacy, Washington, D.C.

Voogt, J. and H. Kasurinen (2005), "Finland: Emphasising Development instead of Competition and Comparison", in *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris, pp. 149-156.

Wang, J. and M.S. Brown (2007), "Automated Essay Scoring Versus Human Scoring: A Comparative Study", *The Journal of Technology, Learning, and Assessment"*, Vol. 6, No. 2, retrieved 11/03/09 from http://www.jtla.org.

Wiliam, D. *et al.* (2004), "Teachers Developing Assessment for Learning: Impact on Student Achievement", *Assessment in Education: Principles, Policy and Practice*, Vol. 11, No. 1, pp. 49-65.

Wiliam, D. (2001),"What is Wrong with Our Educational Assessments and What Can Be Done About It?", *Education Review*, Vol. 15, No. 1., pp. 57-62.

NOTES

1 Vital Insight. Contact: janetlooney1@yahoo.fr. The author wishes to thank Stéphan Vincent-Lancrin of OECD/CERI and Gordon Stobart of the London Institute of Education for their helpful suggestions on earlier drafts of this report.

2 The Trends in International Mathematics and Science Study (TIMSS) is an international assessment of the mathematics and science knowledge of fourth and eighth-grade students around the world. TIMSS was first administered in 1995 and every four years thereafter. The TIMSS video study conducted in 1995 examined teaching in eighth-grade mathematics classrooms in Germany, Japan and the United States. A 1999 follow-up study included classrooms in Australia, the Czech Republic, Hong Kong, Japan, Switzerland, the Netherlands and the United States.

3 Bloom's (1956) taxonomy of skills within the cognitive domain of learning form the foundation for what is now referred as an order of thinking skills. The lower level skills include knowledge, comprehension and the application of knowledge. Higher level skills include analysis, synthesis and evaluation. Higher order skills may be seen as more complex combination of lower skills.

4 This is not to suggest that students should not receive any preparation for examinations. For example, Crocker (2005) proposes that it is appropriate to help students understand how they should behave in the testing milieu. She also suggests that teachers discuss effective problem-solving strategies for tests. Teachers should use a variety of formats in their own tests – not just those used in the major assessments. Crocker's general guidelines emphasise the importance of maintaining focus on the development of broader cognitive skills and content knowledge.

5 While only a small number of studies have focused on the validity of test score gains in high-stakes testing, they have usually found substantial evidence of score inflation. For example, Jacob (2002) examined the impact of a high-stakes tests in Chicago Public Schools in 1996-97. Using a panel of student-level, administrative data, the study found that student math and reading achievement increased sharply following the introduction of high-stakes assessments. Younger students participating in a low-stakes state-administered examination did not increase performance. Jacob's item-level analysis suggests that the achievement gains were driven by increases in test-specific skills, as well as student effort. Teachers also responded to the high-stakes tests by increasing special education placements, preemptively retaining students, and shifting time away from low-stakes subjects like science and social studies. In another study, Klein *et al.* found gains on high-stakes tests in the United States may be three to five times greater than on tests with low stakes (for example, the National Assessment of Educational Progress, or NAEP) (Klein *et al.*, 2006).

6 These three states were chosen for the RAND survey because they have very different approaches to meeting the requirements of the NCLB, and vary in characteristics of student populations. The survey was conducted in elementary and middle schools in approximately 25 school districts. Districts were selected according to a stratified random sampling procedure and included rural, suburban and urban districts. Approximately 100 schools per state were randomly selected within these 25 districts. All school leaders, teachers of mathematics and science in Grades 3, 4, 5, 7 and 8 received surveys (survey questions were tailored for each of these groups and according to the state policy context). High schools were excluded from the survey as testing requirements are not as extensive. Survey findings can be generalised to the entire state from which responses were gathered, based on sampling and non-response weights. The researchers also conducted case studies in six school districts (two in each state).

7 Newman *et al.* found significant standardised effect sizes: 0.43 (reading), 0.52 (writing) and 0.64 (mathematics).

[8] Tests that systematically favour one group over another are defined as biased (for example, girls *vs.* boys, higher *vs.* lower socio-economic status, majority *vs*. minority populations). Critics have charged that test makers and facilitators are often over-represented by a particular socio-demographic group (for example, middle class whites) who share values, habits and language, which may be reflected in the design and content of tests, thereby increasing the chance of test bias.

[9] McDonald describes emotional reactions associated with assessment as follows: "Emotional reactions typically accompany situations where our performance is being measured or assessed. If at any stage of an evaluation we feel unprepared, unsure of our ability, or feel we have not performed to our best, we may experience feelings of unease, apprehension, distress or depression. Alternatively, believing that we are well prepared and able to perform well on an evaluation will be associated with more positive feelings such as anticipation, excitement, exhilaration and pride. Test anxiety is most closely, although not exclusively, associated with the former, more negative set of emotions" (p. 90).

[10] The OECD's 2008 *Education at a Glance* (EAG) found that national student examinations with civil effect for students have a moderate or high influence on the performance feedback provided to schools and teachers in France, Iceland, Ireland and Scotland. Other countries do not have national student examinations, or do not use them for the purpose of evaluating schools. Note that EAG survey covers only lower secondary schools.

[11] Mislevy and colleagues (2001) provide an example of how ICT-based assessment of complex tasks might work in a computer-based simulation for problem solving in dental hygiene (assessing the status of a new patient). Their framework included structures for modeling knowledge and skills (the complex of knowledge, skills and other attributes to be assessed), designing tasks (behaviors or performances that should reveal those constructs), and detecting and synthesizing evidence (tasks or situations should elicit those behaviours). The team identified patterns distinguishing levels of expertise in the field, and built the assessment around this collection of indicators. They then applied probability distributions on levels of knowledge and competence (for example, a hygienist who demonstrates a high proficiency in gathering and using information is more likely to have a high or medium proficiency in assessing the status of a new patient). The model integrates new information as the student proceeds through tasks of the simulation. Interested readers may find additional information on the statistical model used (a Bayseian inference network, or Bayes net) in Jensen (1996); on model development, in Edwards (1998) and on assessment (Mislevy, 1994 and Almond and Mislevy, 1999).

[12] Countries participating in the 2005 OECD study, *Formative Assessment: Improving Learning in Secondary Classrooms* had developed specific policies to promote classroom-based formative assessment. Countries and regions participating in this study were: Canada, Denmark, England, Finland, New Zealand, Scotland and Queensland, Australia. The study did not include an inventory of formative assessment policies across all OECD countries.

ANNEX 1: NATIONAL ASSESSMENTS AND EXAMINATIONS IN OECD COUNTRIES

	Periodical national assessments	National examinations with civil effect for students	External evaluations by an inspectorate	Internal school/self-evaluations	Financial implications/ Likelihood of school closure	Publication of results
Australia	Yes	No	1 per 3 y	1 per y	No	Results of national assessments are published
Austria	No	No	None	None	a	a
Belgium (Flemish Community)	Yes	No	1 per 3y+	None	External evaluations have a high influence on financial rewards and sanctions, moderate influence on the likelihood of school closure	Results of national assessments and external evaluations are published
Belgium (French-Speaking community)	m	m	m	m	m	m
Canada	m	m	m	m	m	m
Czech Republic	a		1 per 3y	1 per y	External evaluations have a low influence on financial rewards and sanctions, high influence on the likelihood of school closure	Results of external evaluations are published
Denmark	No	Yes	m	m	m	The results of national student examinations are published, including in tables that compare school performance

	Periodical national assessments	National examinations with civil effect for students	External evaluations by an inspectorate	Internal school/self-evaluations	Financial implications/ Likelihood of school closure	Publication of results
England	Yes	No	1 per 3y	1 per y	National periodical assessments and external evaluations have a high influence on the likelihood of school closure	The results of periodical national assessments are published, including in tables that compare school performance The results of external evaluations are also published.
Finland	Yes	No	a	m	m/a	No
France	Yes	Yes	a	m	m/a	m/a
Germany	m	Yes	m	m	m	m
Greece	m	m	m	m	m	m
Hungary	Yes	No	a	1 per 3y +	No	The results of internal evaluations are published
Iceland	No	Yes	1 per 3y	1+ per y	a	The results of external evaluations are published
Ireland	No	Yes	1 per 3y+	None	The results of national examinations and external evaluations have a moderate influence on the likelihood of school closure	The results of external evaluations are published
Italy	Yes	Yes	None	None	No	The results of national examinations and of periodical national assessments are published
Japan	No	No	m	m	m	The results of internal evaluations are published

	Periodical national assessments	National examinations with civil effect for students	External evaluations by an inspectorate	Internal school/self-evaluations	Financial implications/ Likelihood of school closure	Publication of results
Korea	Yes	No	1 per 3y	1 per y	External evaluations have a low influence on the provision of financial rewards or sanctions	The results of national periodical assessments and external evaluations are published
Luxembourg	Yes	No	None	1 per y	None	No
Mexico	Yes	No	m	1+ per y	m/a	No
Netherlands	No	No	1 per y	m	m/a	The results of national examinations and external evaluations are published
New Zealand	No	No	1 per 3y	1 per 3y	External evaluations have a low influence on the likelihood of school closure	The results of external evaluations are published
Norway	Yes	Yes	a	None	m/a	The results of national examinations are published
Poland	m	m	m	m	m	m
Portugal	No	Yes	1 per 3y+	1+ per y	a	The results of national examinations and external evaluations are published
Scotland	Yes	Yes	1 per 3y+	1 per y	National examinations have a low influence on the provision of financial rewards or sanctions	The results of national examinations, national periodical assessments and external evaluations are published
Slovak Republic	m	m	m	m	m	m
Spain	No	No	None	None	a	a

	Periodical national assessments	National examinations with civil effect for students	External evaluations by an inspectorate	Internal school/self-evaluations	Financial implications/ Likelihood of school closure	Publication of results
Sweden	Yes	No	1 per 3y+	1 per y	External evaluations have a low influence on the likelihood of financial rewards or sanctions and of school closure	The results of both internal and external evaluations are published
Switzerland	No	No	m	m	m/a	m/a
Turkey	Yes	Yes	1+ per y	1 per y	External evaluations have a low influence on the likelihood of school closure	The results of national periodical assessments are published, including in tables that compare school performance. Internal evaluations are also published
US	m	m	m	m	m	m

Source: OECD, 2008.

Note 1:

None: No requirements

M: Data not available

A: Data not applicable because the category does not apply

1 + per year: More than once per year

1 per y: Once per year

1 per 3y: Once every three years

1 per 3y+: Once every three + years

Note 2: This table includes only national level information for lower secondary schools. The choice to focus on lower secondary schools allows analysis of assessment and evaluation policies and data gathered for the OECD's Programme for International Student Assessment, which is conducted with students at 15 years of age, when students are generally in the last year of compulsory schooling across OECD countries).

ANNEX 2. OVERVIEWS OF OECD COUNTRY POLICIES

This annex provides very brief overviews of OECD country policies related to governance and accountability; recent reforms and support for innovation; student examinations (including school exit examinations and entrance examinations to secondary schools and universities); and school evaluation. Due to space constraints, it does not include full descriptions of all government agencies concerned with education, the structure of compulsory school systems, teacher training and professional development, special education, school finance, or private education.

The information contained in this annex was drawn from the UNESCO International Bureau of Education's "World Data on Education" (WDE) reports for 2006[13]. The full reports are available at http://www.ibe.unesco.org/Countries/WDE/2006/index.html.

AUSTRALIA

Governance and accountability **–** The national agency is the Ministerial Council for Education, Employment, Education, Training and Youth Affairs (MCEETYA). The Australian system is decentralised, and the federal government has implemented a system of standards-based assessment. Standards are developed at the national level. States and territories determine policies and practices and delegated responsibilities give to individual schools in matters such as curriculum, student assessment, personnel, and resource allocation. Ministers from the Australian commonwealth, states and territories meet biannually to discuss mutual concerns and to co-ordinate activities.

Recent reforms and support for innovation – The Australian country report describes teaching and learning activities as involving a range of strategies such as group work, peer tutoring and assessment, experiential learning and opportunities for learners to negotiate the process and content of their learning. Ongoing assessment and feedback (in formal and informal assessment) are an integrated part of the teaching and learning cycle, hands-on experiential learning and opportunities for learners to negotiate the process as well as the content of learning.

Individual jurisdictions and/or schools may receive special funding to develop pilot projects. Innovations developed in individual jurisdictions may be pick up and implemented in other jurisdictions or nationally. Schools may choose to develop a focus, such as environmental education, learning technologies, language, performing arts, sport or some other area of specialty. These schools must typically undergo a selection process. Projects also typically provide professional development opportunities as part of the development brief along with clearly specified review and evaluation mechanisms.

Exit and entrance examinations – Each state and territory has developed its own accreditation process for certifying school completion at Year 12 and ranking students for entry into tertiary institutions. Several states also certify students at Year 10 state-based or school-moderated assessment processes.

Assessment/examination formats – No information.

Evaluation – States and territories develop annual strategic plans or similar documents, setting out priorities for schooling. The priorities guide system-wide planning, as well as schools, district or regional offices and the central office of the education department directorates in annual planning, budgets and in programme monitoring.

AUSTRIA

Governance and accountability – The national agency is the Federal Ministry of Education, Science and Culture. In 1995, schools were granted autonomy, with opportunities to develop specific profiles and areas of emphasis.

Between 2003 and 2007, Austria piloted education standards for Grades 4 and 8 in the subjects of German and mathematics, and English (Grade 8). Standards for the natural sciences (applied on a transdisciplinary basis), were introduced as pilot projects in Autumn 2008. The standards describe competences for both students and teachers are described in concrete terms so that they can be developed into specific tasks, and assessed.

Recent reforms and support for innovation – Decisions concerning innovations and/or reforms of the school system are made by the Federal Parliament.

The School Reform Commission, which serves an advisory body to the federal Ministry, has been discussing a comprehensive reform, including the aims, contents, structures, and methods of education. The curriculum is increasingly focused on the development of fundamental, sustainable competences at different stages of education. The Centre for School Development of the Federal Ministry of Education provides analysis and advice on school development and innovation; national and international research findings; evaluation of innovation and pilot projects having an impact on national education policies; and other development activities.

ICT is seen as a driving force for development processes, innovations and quality enhancement in all areas of the education system – including the teaching of comprehensive digital competences, individual support.

Exit and entrance examinations – Students take and examination for graduation from secondary school. The examination provides access to university, higher technical colleges and academies (students may be required to take additional examinations in some subjects), and civil service.

Assessment/examination formats – The school-leaving examination includes seven partial examinations (some written, some oral) in at least four different subjects; compulsory written examinations in German, a foreign language and mathematics; and compulsory oral examination in an additional modern foreign language.

Evaluation – School self-evaluations and external school inspections. Regions draft education plans.

BELGIUM (FLEMISH COMMUNITY)

Governance and accountability – The Flemish Minister of Education and Training is the central executive authority within the Flemish Community. The Department Minister is responsible for the development of an education framework. The system within Flanders is decentralised, and schools are held accountable for student outcomes.

Recent reforms and support for innovation – Priorities for education are set out in *A Forward-looking Education Policy* (13 July 1999) and in the *Policy Documents: Education and Formation* (11 January 2000). The curriculum must be approved by the Minister, with the advice of the inspectorate. Priorities include: scaling-up of partnership between schools; the establishment of new local councils and participation councils; renewal of teacher education; strengthening equal educational opportunities (including school choice); the introduction of ICT; strengthening the Centres for Educational Guidance and links with the job market.

Goals and developmental objectives have been have been divided into course-related and cross-curricular goals and objectives, with the aim of having as many pupils as possible achieve these minimum targets.

Developmental objectives are to be pursued, but do not necessarily have to be achieved by all pupils. Cross-curricular goals include: learning to learn, social skills, public responsibility, health education, environmental education, musical and creative training, technical and technological training. The government has launched several initiatives targeting teachers, heads of organizing authorities to facilitate implementation of target based approaches.

Exit and entrance examinations – All students with a secondary education diploma are eligible for admission to university; no entrance examinations are required.

Assessment/examination formats – The Flemish inspectorate may organise assessments in certain learning areas, but because the nationwide central testing is legally forbidden, assessment results may not be used to compare school quality.

Evaluation – External evaluation by the Inspectorate. Research projects are followed by monitoring groups (consisting of scientific experts, those who are responsible for policy, educators, as well as the researchers who are working on the project concerned). Monitoring data are used to adjust projects, if necessary.

BELGIUM (FRENCH-SPEAKING COMMUNITY)

Governance and accountability – The central agency is the Ministry of the French-speaking community (in 1997, the Ministry of Education, Research and Training and the Ministry of Culture and Social Affairs were combined). Within the French-speaking community, there are 23 "socio-pedagogic districts". Each of these districts is overseen by a group of school leaders; this group is led by an elected member who represents the district at the Pedagogic Council of the French-speaking community.

National standardised assessments are organised at the beginning of the year and the beginning of a cycle in all schools to help teachers understand students' level, and to adapt teaching appropriately. A representative sample of classes is analysed and shared with teachers, allowing comparison of the results of their class to the sample. Teaching plans are then developed and sent to each teacher. An inspection assures follow-up, and continuing professional development of teachers.

Recent reforms and support for innovation – Educational objectives in the French-speaking community are to: to promote the self-confidence and personal development of each student; help all students develop knowledge and competences for lifelong learning; take an active role in the economic, social and cultural life; promote equity.
The *Contract for the School* guides the education community and its social partners through 2013, and has the following priorities: hiring of more teachers; mastery of foundation competences for each student; development of job skills while in school; better preparation of teachers; strengthened dialogue between schools and families.

Exit and entrance examinations – No information.

Assessment/examination formats – No information.

Evaluation – There is an inspectorate service charged with overseeing teaching at all levels in the French-speaking community, and inspectorate charged with teaching or maternal and primary schools. Researchers and directors of the inspectors general, and an education unit in the Ministry of Statistics organises an external assessments of student achievements at the beginning of the third and fifth years of primary school, and the first year of secondary school.

CANADA

Governance and accountability – The Canadian Federal government does have direct power over education, but does have some influence over policies, standards and objectives. The national Council of Ministers of Education, Canada, (CMEC) was established in 1967 by the ministers of education to promote discussion on matters of mutual interest, collaboration and consultation among the provinces and territories, and to liaise with the federal government. Individual provinces have implemented high-stakes assessments as part of standard-based assessment reforms.

Recent reforms and support for innovation – The Canadian country report notes that the prevalent focus of education is on progressive or child-centred education. At the same time, the traditional/teacher-directed model has regained some momentum with the increased emphasis on the development of core skills, and large-scale provincial testing in targeted subject areas at particular grade levels have been introduced. The CMEC has also set priorities for Aboriginal education and literacy.

Recent major changes in Canadian education centre on: accountability, quality of education, accessibility, mobility and responsiveness to learners' needs. There have been several initiatives aimed at reorganization and restructuring, funding reform, curriculum renewal, early childhood development, performance assessment, the integration of ICT into the curriculum, cooperation and shared services, and school-to-work transition.

During the final two years of secondary education, students have fewer compulsory subjects and can choose more optional courses in specialised programmes. In general, academic or university preparatory tracks and vocational or work-related training are in the same institutions.

Exit and entrance examinations – No information.

Assessment/examination formats – No information provided on assessments implemented at province level. The national School Achievement Indicators Programme (SAIP) was launched in 1989. The SAIP is given to a random sample of students. Achievement is described over five levels, representing a continuum of competences. Assessments include written components consisting of multiple-choice and short-answer questions. The science assessment includes seven practical hand-on tests to measure science inquiry and problem-solving skills.

Evaluation – Policies for school inspections and quality assurance vary from province to province

---

CZECH REPUBLIC

Governance and accountability – The national agency is the Ministry of Education, Youth and Sports (MEYS). The Czech system was decentralised in a series of reforms between 1990 and 1993 to reflect social, economic and political changes since 1989. Standards and instruments for assessment of educational outcomes as part of a coherent system are now under development.

Recent reforms and support for innovation – The National Programme for the Development Education in the Czech Republic (or White Paper) was approved in February 2001. The National Programme calls for the development of long-term plans, which are formulated at national and regional levels and include medium-term development priorities. Since 2004, schools have been allowed to develop their own programmes, within the national Framework Education Programmes (FEP). The framework programmes are approved by the MEYS in agreement with relevant ministries, trade unions and representatives of employer organisations. The FEP focus is on the development of competences.

The Czech Republic recently introduced the "integrated secondary school", combining vocational and technical schools. Secondary schools also offer extension programmes, primarily allowing non-maturita vocational school graduates to earn the maturita. Higher professional studies were introduced as a new educational level in 1995.

Upper secondary schools, or gymnasia, are oriented towards students' cultural and personal development. Gymnasia support active role of students in choosing subjects for study, and for independent study (individualised study or independent investigative and research projects).

Exit and entrance examinations – The maturita examination previously provided a way to monitor schools' educational success. However, with increasing diversity in quality and difficulty, it no longer serves this purpose. The examination system has not been standardised or centralised. The system is now undergoing reform in order to bring greater uniformity to the system.

Assessment/examination formats – Diverse formats for entry into secondary school and for the maturita examination at completion of secondary schools across the country.

Evaluation – There are school self-evaluation instruments at various levels of governance, but there is no coherent system. External evaluations are conducted by the Czech School Inspectorate (CSI), which was established in 1991. The CSI publishes annual reports setting out the findings from inspection activities in schools.

DENMARK

Governance and accountability – The national agency is the Ministry of Science, Technology and Innovation. The Danish system of education is decentralised. The Minister of Education sets the targets for student learning. Curriculum guidelines are seen as recommendations, but nearly all schools choose these as their binding curricula

A range of compulsory national tests have been introduced to track student outcomes using ICT-based technology.

Recent reforms and support for innovation – In 2007, the government presented the "National Strategy for Lifelong Learning", building on the Globalisation Strategy ("Progress, renewal and development: Strategy for Denmark in the global economy"). Education, lifelong learning, research and innovation are emphasised.

A major reform of the gymnasium/upper secondary school was adopted in 2005. The reform: updates the knowledge, proficiency and skills to better prepare students for higher education; updates the common reference framework for cultural and general educational concepts; encourages cooperation between teachers in different subject areas; introduces four programmes with clearly different profiles; and provides greater flexibility. The reform introduces a common six-month course, followed by two and a half years of specialised studies, and increases the focus on natural science.

The government has made significant investments in ICT, including for in-service teacher training, support for Internet-based teaching materials, ICT-based assessments and examinations, and e-learning.

Exit and entrance examinations – There is a national matriculation examination.

Assessment/examination formats – National assessments are computer-based and adaptive. If a student answers a question incorrectly, then he or she is given an easier question; if the student answers correctly, he or she is given a more difficult question. No two students receive the exact same test, so it is not possible to make internal comparative evaluations of tests within a specific class.

For the school leaving examination, written questions are produced by the Ministry of Education's Department of Upper Secondary Education. For each subject, there is an examination committee of hand-picked teachers and the Department's subject advisers. Oral examinations are held in subjects selected at random. The results of the examinations are averaged with the mark the students achieve for the year.

Evaluation – The Danish Evaluation Institute was established in 1999 to undertake systematic and mandatory evaluations of the teaching and learning at all levels of the education system.

FINLAND

Governance and accountability – The national agency is the Ministry of Education. Many of the former powers of the central education administration in Finland have been devolved to the local education authorities and to individual schools. National assessments of student outcomes are given to random samples of students; the key function is to identify areas requiring further improvement within the entire school system.

Recent reforms and support for innovation – The government's 1999-2004 plan for the development of education and research (most recent data available) emphasised: the development and implementation of an extensive Information Society Programme, with an emphasis on content, teacher training and use of information networks; education in mathematics and natural sciences; strengthening programmes for language teaching and internationalization; and, quality improvement. Lifelong learning is also an important theme.

Teachers choose their preferred teaching methods. The Basic Education Act promotes student self- assessment and independent goal-setting. The country report notes that traditional (*e.g.* teacher-led lectures) are still common, in addition to student-centred approaches (*e.g.* independent learning and teamwork). The report describes the task of

schools as being to teach students to acquire knowledge independently, and to tolerate contradictions.

Upper secondary education includes thematic subject modules, integrating different elements of the curriculum. The modules may be joint-taught and/or include project work.

The 1999-2004 Plan also emphasised greater cooperation between general upper secondary and vocational schools, for example, in the development of joint study programmes and regional education and training. The aim is to promote entrepreneurship education and training.

Exit and entrance examinations – Students participate in a national matriculation examination.

Assessment/examination formats – The national matriculation examination consists of four compulsory and one or more optional subjects. No information on examination formats is available.

Evaluation – School self-evaluation. The National Board of Education is responsible for national evaluations of schools. Monitoring and evaluation focus on the extent to which the objectives set in statutes, policies and core curricula are achieved. In addition, the national evaluation system supports schools and teachers in reform efforts.

FRANCE

Governance and accountability – The national body is the Ministry of Education. The French system decentralised some functions in 1983. In 2004, new legislation transferred ownership of education to the French regions and to local institutions.

The curriculum is developed at the national level. National standardised assessments were introduced in 1989. The assessments, which are conducted at the beginning of the school year, are not used as a tool for school accountability, but as for diagnostic purposes. The national Ministry of Education analyses a representative sample of the assessment results to develop a picture of pupil achievement across the country. The Ministry also gathers information on student attitudes and values in questionnaires.

Recent reforms and support for innovation – The New Contract for Schools was established in 1994. Under the Contract, each school should: clarify its mission, re-focus teaching on mastery of fundamental skills, adapt teaching to the needs of heterogeneous students; and develop students' skills for lifelong learning, as well as their creativity and adaptability. School-based innovations are encouraged, within the framework of general education and regulations, and national objectives for education.

The central curriculum increasingly emphasises objectives for learning, rather than content. Teachers have some flexibility on setting timetables, and may choose methods for teaching. The national Ministry has also placed focus on building capacity for initiative (among school leaders, teachers, parents). The central education framework outlines seven competences: French; a living foreign language (including mastery of basic competences for reading, writing and speaking); mathematics and scientific and technological culture; ICT; humanist culture; social and civic competences; autonomy of spirit and of initiative. Students have more choices.

Exit and entrance examinations – Students at the age of 15 may take the *brevet des colleges*, or for those following a vocational course, the *certificat d'aptitude professionele* (CAP) and the *brevet d'études professionelles* (BEP). The *brevet* was required for entry to secondary school, but is no longer required. Many students nevertheless participate in the examination, as a way to prepare for later examinations.

The main requirement for entry to higher education is the *baccalauréat* taken at the end of the third year of upper secondary school. There are three types of *baccalauréat*: academic technical and vocational. For students wishing to enter the *grands écoles*, students take highly competitive entrance examinations. Students who do not succeed on these examinations may apply to the regular university system. Some universities accept scores on the *grands écoles* entrance examinations in place of the *bac*.

Student performance on the *baccalauréat* is published in newspapers.

Assessment/examination formats – With the exception of foreign language, all subjects are assessed, primarily through a series of examination papers, each lasting from 3 to 3.5 hours, taken over four days. For most universities, entry is based solely on the *bac.* Questions for the *bac* are set at the regional level. While students are taking the examination at the same time, they are not on the same questions. The examination is marked by teachers from other schools.

Evaluation – The Ministry of Education's Office for Assessment and Forecasting (established in 1987) evaluates cost, financing, organization, assessment of achievement assessment, school effectiveness, classroom teaching practices innovation projects, and so on.

Schools conduct self-evaluation. Centrally developed computer software, which includes a set of standard indicators, is made available to secondary schools. Schools can judge their own performance (exam performance, resources, school management and environment) against national levels. Schools are also encouraged to develop their own specific indicators based on local characteristics and needs.

GERMANY

Governance and accountability – The system is highly decentralised. At the central level, the Ministries of the Länder are headed by a Minister (in Berlin, Bremen and Hamburg: Senator). The Federalism reform of 2005 further reduced the federal role in the education system, and transferred legislative powers to the Länder. The Ministries of Education, Cultural Affairs and Science in each Länder develop policy, legislation and administrative regulations.

The education system is shifting from input-based control (through political and administrative regulations) to greater control of outputs. The core elements of this approach are: educational standards, binding for all Länder, and educational reporting.

Recent reforms and support for innovation – In 2003-04, standards relating to general objectives for competences and knowledge students should attain at different levels were introduced. They include standards for German and mathematics at the end of Grade 4. Standards have also been adopted for the general education school-leaving certificate (obtained on completion of Grade 9 at the *Hauptschule*/*Hauptschulabschluss*), and the general education school leaving certificate obtained on completion of Grade 10 at the *Realschulen* (*Mittlerer Schulabschluss*). Standards are based on areas of competence in different subjects as well as basic knowledge.

General trends across the Länder include: interdisciplinary learning; independent learning; "sustainable" knowledge; core knowledge; specialist knowledge; concepts and methods of work in different disciplines. Länder are currently concerned with reducing the quantity of required educational content. New core curriculum, also referred to as "basic knowledge", and establish essential areas of knowledge and behaviour. All Länder have introduced foreign language.

Exit and entrance examinations – Students take the *Abitur* examination at the end of the academic upper secondary school (usually Grade 13). The examination qualifies students for entry into general higher education in all fields of study. The four subjects covered in the *Abitur* examination must include mathematics or a foreign language.

Assessment/examination formats – No information.

Evaluation – In recent years, all Länder have been focused on developing quality assurance measures for school systems and individual schools. Approaches include implementation of standardised school assessments at level of the Länder (to compare performance in parallel classes within a school); external moderation or correction of final examinations (for example, a second correction of central final examination papers, or random monitoring by the school supervisory authority of work submitted for the *Abitur*). There are also focused evaluations on special topics, and internal and external evaluations of the schools (for example, through mutual participation in *Abitur* examinations of other Länder or the exchange of *Abitur* work between the Länder).

Other methods of quality assurance include statistical surveys (carried out by the Federal Statistical Office and by the statistical offices of the Länder), and evaluations carried out by educational research in institutes subordinate to

federal and/or Land ministries.

GREECE

Governance and accountability – The Greek education system is centralised, although some administrative structures have been decentralised. The national agency is the Ministry of Education and Religious Affairs.

Recent reforms and support for innovation – Current debates in Greece focus on the possible extension of compulsory school from 9 to 12 years, strengthening of decentralised administrative structures, increased involvement of local authorities in financial decisions, and renewal of curricula and pedagogy.

The Ministry of Education and Religious Affairs is planning an ICT-based National School Network, including the development of educational software (consortiums of Greek universities have developed more than 100 multimedia titles).

Exit and entrance examinations – Students who pass the matriculation at the end of lower secondary school receive a school-leaving certificate. The examination can be re-taken without having to repeat the year, if necessary.

Assessment/examination formats – No information

Evaluation – Schools are required to conduct self evaluations. The Hellenic Quality Assurance Agency was established in 2006.

HUNGARY

Governance and accountability – The national agency is the Ministry of Education and Culture. The national government has granted increasing levels of autonomy to schools since the early 1993. There is a two-tier system of national and local regulation. The 1995 Amendment of the Education Act further strengthened the professional autonomy of the educational and teaching institutions and the role of the local governments. The National Core Curriculum is developed by the central government, along with pedagogical programmes and local curricula.

Surveys of student performance based on standardised testing methods are conducted in several examination centres. Data gathered from surveys based on sampling procedures are available for several years.

Recent reforms and support for innovation – The National Core Curriculum (NCC), which was introduced in 1995, has been a key element for education reform. It provides a framework for the development of local curricula. The local curricula take the views of parents, students and school leaders into consideration. Since school-level development was completely new in public education, the central government took a significant role in building teacher competence to develop curricula and plan pedagogical work.

Between 2005 and 2007, the Hungarian government introduced new legislation and drew upon European and national funding sources to promote a series of new initiatives, including the introduction of key competences for lifelong learning (such as basic competences, foreign language skills, labour market competences, ICT, mathematics and science literacy, life skills and entrepreneurship). In addition, Hungary promotes the development of innovative schools and teaching methods (such as forest schools, green schools in the countryside, experiential learning).

The country report acknowledges ongoing challenges regarding the highly selective nature of the Hungarian educational system, and its low success in compensating for various social inequities. The report notes there has not been a tradition of personalising learning, or assessing student performance in other than summative ways.

Exit and entrance examinations – There is a highly competitive matriculation examination at the end of upper secondary school.

Assessment/examination formats – Achievement of standard quality criteria is measured through a central written assessment containing the same tasks based on the same requirements for all candidates. Students also participate in an oral examination. Examinations include essay tasks that measure skills in interpreting and using resources, using the appropriate terminology, exploring and context.

The matriculation examination is held in the school and is administered by an examining board made up of teachers from the school and chaired by a delegate from the educational authority. They are corrected and marked by teachers from the school. The upper secondary school-leaving certificate allows for but does not guarantee admission to higher education; colleges and universities have their own entrance examination requirements. The prior school leaving examination was not well-aligned with curriculum reforms, and created significant barriers to change. The new secondary school-leaving examination (first administered in 2005) is a standardised exam inasmuch it is based on standardised requirements and is administered according to standardised specifications. However, it does not meet the requirement of standardization concerning piloted and calibrated tasks.

Evaluation – The National Public Education Evaluation and Examination Center was established in 1999. An independent quality scheme, called "Comenius 2000" was imitated in 2000. The Comenius programme involves institutional self-evaluation, adaptation of pre-existing quality assurance systems at the local level and capacity development, and, the diffusion of ideas.

ICELAND

Governance and accountability – The Ministry of Education, Science and Culture issues the National Curriculum Guidelines for compulsory education. Iceland has increasingly decentralised responsibility in the education system to local levels. Teachers choose teaching methods, their instructional aims and the conditions under which they teach. There are national assessments in the fourth and seventh Grades in Icelandic and mathematics.

Recent reforms and support for innovation – General and vocational training were merged into one comprehensive system in the 1970s. The Upper Secondary School Act specifies four types of programmes of study, *i.e.* academic programmes leading to matriculation, vocational programmes, artistic programmes and a general programme of study. According to legislation, all programmes of study can lead to further education.

Upper Secondary School reforms introduced in 1996 aim at improving standards, and increasing flexibility and diversity in post-compulsory education. Legislation increased options for subject specialisation based on programme choice, and well-defined entrance requirements for different programmes.

The Ministry of Education, Science and Culture's Project Plan for *e-Learning 2001-2003* introduces 'distributed education'. Students will be able to choose among different services offered by a variety of educational institutions. The updated Plan for 2005-08 includes professional development for teachers on the use of ICT, development of educational software, and the build-up of ICT infrastructure. Distributed learning promotes new practices, strengthened cooperation among schools, and the breaking down of traditional boundaries.

Exit and entrance examinations – There are nationally coordinated examinations in four subjects (Icelandic, mathematics, English and Danish) at the end of compulsory school (Year 10). The examinations are marked centrally by a group of teachers who are selected by the Educational Testing Institute.

Assessment/examination formats – No information.

Evaluation – Schools are required to conduct self evaluations. Every five years each school's methods of assessment are to be evaluated by an outside party.

IRELAND

Governance and accountability – The central agency is the Department of Education and Science. Ireland has increasingly devolved educational services to school or institutional level and is responding to demands for more accountability. Student performance is are measured in the National Assessment of Mathematics Achievement (NAMA) of mathematics in primary schools at fourth year level.

Recent reforms and support for innovation – Primary school curriculum is based on a child-centred approach, as initially outlined in a 1971 review. The review was a radical change from traditional subject-centred, didactic approaches. Subsequent reviews reinforced (1990 and 1999) the child-centred principles.

In 1999 the primary school curriculum was completely revised. New emphases of the curriculum include a focus on learning styles, the role of ICT, and the integration of assessment for diagnosis and adaptation of teaching to meet children's needs (formative assessment). Primary school teachers participated in in-service training during the phased implementation of the revised curriculum.

Recent reforms aim at meeting increasing demand for secondary and tertiary education, and greater accountability. A 2008-13 plan aims at developing an e-learning culture in all schools, embedding ICT across the curriculum.

Exit and entrance examinations – Students normally take the Leaving Certificate Examination at the age of 17 or 18, after five to six years of post-primary education. This examination may be used for entry qualification for a range of third-level institutions, including the universities and as a selection test for entry to many kinds of employment. The Leaving Certificate has a dominant influence upon the curriculum, methodology, assessment and organization of secondary schools.

Assessment/examination formats – No information.

Evaluation – External evaluations are conducted by the national inspectorate.

ITALY

Governance and accountability – The national Ministry of Public Education has the overall responsibility compulsory education. The education system is decentralised. Legislative Decrees in 2004 and 2005 called for the creation of a national body for the assessment and evaluation of the educational system, a general framework and regulations on standards (*i.e.* minimum learning outcomes) for lower secondary and upper secondary schools.

Recent reforms and support for innovation – National standards are under review to better define knowledge, skills and competences students should have acquired by the end of each segment of their schooling. Recent legislation stresses lifelong learning. Education is to support the development of general and specific competences, consistent with student ability and choices.

Since 1974, schools have had the freedom to innovate in programmes and in teaching methodology. A 1974 Decree also established several institutions to support innovation, including the Regional Institutes for Research, Experimentation and In-service Training (IRRSAE, subsequently transformed into the Regional Institutes for Educational Research, IRRE); the European Centre for Education (CEDE, subsequently transformed into the National Institute for the Evaluation of the Education System); and the Educational Documentation Library (BDP, subsequently transformed into the National Documentation Institute for Educational Innovation and Research, INDIRE; as of 27 December 2006 the INDIRE has been replaced by the National Agency for the Development of School Autonomy, and the IRRE by Regional Offices of the National Agency).

A February 2004 Decree makes teaching of ICT compulsory (beginning in the first year of primary school) and foreign language (primarily English, beginning in lower secondary school). Schools are required to teach core curricula, defined at the national level; regions have flexibility to define other parts of the educational programme.

Exit and entrance examinations – Both the primary and secondary cycles end with a national matriculation examination.

Assessment/examination formats – The examination at the end of upper secondary school consists of two written tests and an oral examination, held before a Board of Examiners appointed by the Ministry of Education. The oral examination includes two subjects and includes a discussion of the written tests.

Evaluation – External evaluation of schools are carried out by INVASI.

JAPAN

Governance and accountability – The central education agency is the Ministry of Education, Science, Sports and Culture (MEXT). The country has increased levels of local autonomy, and plans to implement national academic achievement tests at elementary, junior high and high school levels to track student progress. At the elementary level, assessments are to be administered in Grades 5 and 6 in Japanese, social studies, mathematics and science. At the junior high level, assessments are to be administered in all grades in Japanese, social studies, mathematics, science and foreign language.

Recent reforms and support for innovation – Japan has introduced a series of reforms since the mid-1980s aimed at addressing problems related to uniformity and rigidity of the system. In 2002, the Ministry introduced a revised course of study for primary and lower secondary schools, and in 2003 introduced revisions for freshman in upper secondary schools. "Integrated Study" programmes (introduced in 2002) in lower secondary schools to develop issues outside of the curriculum framework. Students also have more choice of subjects in upper secondary schools. The new curricula aim to reduce stress in schools, and promote and nurture "zest for living", helping students to develop skills for independent thinking and learning-to-learn. The curricula also encourages teaching approaches that put the focus on individual children (through tutorials and group study), adaptation of teaching to student's level of understanding, and a greater range of elective courses with advanced content. In 2002, the school week was reduced from six to five days. Reforms have also aimed to reduce student workload.

The government introduced an ICT Reform strategy in 2006 (to be partially implemented by April 2009), with the aim of enabling students in primary and lower secondary schools to use ICT, and to cultivate "information ethics". The reform provides for teacher training in ICT use.

Exit and entrance examinations – Local public upper secondary schools select students on the basis of scholastic achievement tests given by the Board of Education of the prefecture or municipality running the school and records of each applicant. The central Ministry of Education (MEXT) encourages prefectural governments to adopt diversified selection methods for entry into upper secondary schools.

Admission to universities and junior colleges is granted on competitive basis. New students are selected by means of scholastic achievement tests (including the nation-wide examination jointly administered by the National Center for University Entrance Examinations and national, local public and private universities) and other tests. Assessments submitted by upper secondary schools may also be taken into account. Entrance examinations for prestigious universities are highly competitive.

Assessment/examination formats – No information.

Evaluation – No information.

KOREA

Governance and accountability – The national agency is Ministry of Education and Human Resources Development. The education system is decentralised. The Korean country report notes that while the national curriculum allows secondary schools to introduce their own courses and tailor provision, 10 of 16 provinces present almost the same guidelines as found in the national curriculum and do not specify courses they offer. As a result, there appear to be no major differences across regions.

The National Assessment of Scholastic Achievement (NASA) measures attainment of objectives in the school curriculum. NASA samples 1% of all students at different school levels and across regions. The results of NASA are analyzed and published by the National Board of Educational Evaluation (NBEE) annually.

Recent reforms and support for innovation – The Korean government undertook a major reform of the education system, based on recommendation of the Commission on Education Reform (PCER). The reform emphasises lifelong learning (including open education through ICT), shifts to "learner-centred provision", diversifies educational programmes and schools and areas of specialization, and extends student and parent choice. The university entrance examination was revised to address diverse aptitudes of individual students. The reform also introduced greater school autonomy and accountability, and an administrative and financial support system based upon objective evaluation.

The central government sets out curriculum for each school level and sets criteria for the development of textbooks and instructional materials. Curricula are revised periodically. The common basic curriculum developed in the Seventh Curricular Revision, introduced in 1997, includes core subjects, and autonomous activities, special activities. Cross-curricular learning themes are taught mainly through the "discretionary activities" and integrate related curricular subjects and special activities. They also serve to link local community and family. The Seventh Curricular revision sets out guidelines for the reduction of student workload, based upon the "minimum necessary learning elements".

While references materials are developed centrally, there are also efforts to decentralise the curriculum decision-making. For example, the national curriculum allows each high school to develop its own courses. However, ten out of sixteen provincial education offices of follow guidelines of the national curriculum.

Exit and entrance examinations – Lower secondary school graduates who opt for academic upper secondary schools submit an application to the Provincial Offices of Education. Following an examination administered by the local offices of education, students are assigned to different high schools through a random method. In non-equalised areas, students directly apply to high schools they wish to enter, and upper secondary schools select students according to their own criteria.

Students take the highly competitive College Scholastic Ability Test (CSAT). Colleges and universities may also take teacher-documented assessments into account in their admissions decisions.

Assessment/examination formats – Questions in the NASA are posed using mainly the multiple-choice and free-response format. English listening comprehension test items and science performance assessment tasks are required as part of English and science tests, respectively. The results of each school and individual student's performances are not released to the public. No information is provided on the format of the CSAT.

Evaluation – The Centre for School Development at Korea Educational Development Institute (KEDI) monitors daily educational practices in the schools and provides assistance.

LUXEMBOURG

Governance and accountability – The national Ministry of Education and Training, as well as the Ministry of culture, higher education and research are responsible for administration of the education system.

Recent reforms and support for innovation – Education reforms implemented at the beginning of the 1990s did not lead to improvement in student achievement. Reforms at the end of the 1990s therefore concentrated on the following objectives: improve continuity between primary and secondary school; rethinking assessment of competences; improvements in initial teacher training; engagement of schools in development projects. ICT has been progressively integrated at all levels of education.

The student population in Luxembourg is increasingly diverse. The education system has therefore placed emphasis on supporting the success of all students, and reducing the impact based of students' social and cultural origin on the educational path they follow. The country report emphasises the need for teachers to have greater flexibility in order to respond to student needs. Teachers have about one-third of their time free in order to support student needs. Other priorities include the need to raise the level of qualifications for all and to increase the number of students receiving all kinds of qualifications. Schools are to guarantee all children the same minimum level of knowledge and competence. Programmes are adapted at the local level.

Research in education is coordinated by SCRIPT – the Coordinating service for research and innovation in pedagogy and technology established in 1993 under the Ministry of Education. SCRIPT develops annual action plans based on the priorities defined by the Minister, supports different projects, and participates in implementation of some priority actions, and coordinates professional development for teachers.

Exit and entrance examinations – Secondary school students take a school-leaving examination. Tests take place over eight half days. The tests are organised at the national level. Successful students earn their diploma, and have access to higher education.

Assessment/examination formats – School leaving examinations are written although recently some subjects include an oral component. Student marks are also taken into account.

Evaluation – Primary school inspectors advise local areas in regard to organizational and pedagogical issues. There is no external inspection in secondary schools; school directors report to the central administration regarding teaching and learning.

SCRIPT conducts evaluations of the education system. SCRIPT coordinates basic statistics on the education system, as well as the implementation of indicators.

THE NETHERLANDS

Governance and accountability – The national agency is the Ministry of Education, Culture and Science. The school system is decentralised. Curriculum objectives are centrally developed and adapted at the local level.

Common standards for literacy and numeracy for all students across different educational streams were developed in 2007. The standards describe the learning outcomes of students at 12 years (end of primary education); 16 years (entry into vocational education); and, 18 years (entry into higher education), A each level, two standards are specified: a common, basic standard, and an advanced standard. Pilots are currently being held, and will be revised based on findings.

Recent reforms and support for innovation – Revised core objects for primary school were published in 2006. There are now fewer objectives (down to 58 from 103); the objectives are also more global nature. They describe the desired results of learning. Schools choose teaching methods and select or develop their own materials.

New curricula are also intended to give school greater freedom to develop distinct profiles. Schools may offer specialised subjects combining: science and technology; science and health; economics and society; culture and society. A 2007 Act of Parliament adapts rules, allowing students greater freedom to choose subjects, and opportunities to deepen knowledge and skills. Several committees are to make recommendations on further curricular reform, with particular emphasis on science.

The Netherlands supports innovative schools centred on the development of students' needs and interests. However, most schools differentiate within more traditional forms of whole class teaching. In 1999, the government introduced "Preparatory vocational education", with different "learning pathways". Graduates of level 4 courses may continue to higher professional education.

The Ministry of Education, Culture and Science promotes the use of ICT in schools to prepare students for the labour market and to facilitate learning (anywhere, anytime).

Exit and entrance examinations – During the final two years of secondary school (or, for vocational programmes, the final three years) students prepare for their school-leaving examinations. Most of the school timetable in the last few years is devoted to subjects on the examination.

The school-leaving examinations include an internal school exam and a national exam. The national exams are the same for all students following the same type of education.

The teachers calculate the final grades for each student in each subject by taking the average of the marks obtained in the internal or school exam and the national exam. All the final grades together decide whether a student has passed the school-leaving examination.

Assessment/examination formats – Under the new system of fixed subject combination, the school exam consists partly of written tests with both open-ended and single-answer questions as well as practical assignments. Some of these practical assignments are graded. For others, pupils only need to show that they have carried out the assignment satisfactorily. The school exam in also includes a combined project. This is a long-term practical assignment involving two or more of the specialised subjects studied. The idea is to test students' knowledge, understanding and skills, and make them think about the links between different elements of the curriculum. The national exams are held in May of each year. When the results of these exams are published, every student is given the chance to re-take one subject in June, and if necessary, they can re-take the remaining subjects in August.

NEW ZEALAND

Governance and accountability – The system is decentralised. The Ministry of Education develops curricula for primary and lower secondary schools, sets learning objectives, and monitors student performance against those objectives.

The New Zealand Curriculum Framework identifies five aspects of formal assessment: school-based assessment; records of school achievement; transition point assessment; national monitoring of standards; and assessment for qualifications.

Recent reforms and support for innovation – The document *Education priorities for New Zealand* (May 2003) promotes the development of "twenty-first century skills" (foundation skills, leadership and innovation skills, and skills and enthusiasm for lifelong learning). The document also emphasises the importance of reducing systematic underachievement in education. The government supports Maori language learning with additional funding for Maori immersion or Maori language programmes.

Curricula are also being revised to reflect key competences, integrating essential skills, knowledge, attitudes and values (in line with the OECD's Defining and Selecting Competences – DeSeCo). Nationally agreed "unit standards" serve as "building blocks" towards a qualification. Each standard belongs to one of the eight National Qualification Framework levels, which signal progressively more difficult or complex outcomes required of learners. The levels describe (rather than prescribe) qualifications. Learners receive credit toward qualifications as they attain different outcomes. Outcomes at each level are registered in a personalised Record of Learning.

Transition Point Assessment is a national initiative now under development. It is intended to help schools to identify the learning needs of cohorts of students at each new phase of schooling, and may also be used by to assist in targeting school resources. National standards will need to be determined before this initiative can be launched.

In addition, initiatives for standards-based assessment (assessment against standards which are described in some form) draw upon international models, such as the assessment framework for the United Kingdom's General National Vocational Qualification (GNVQ) and the competency-based national assessment model being developed in Canada. National educational standards will be monitored on a four-year rolling cycle by assessing around 3% of 8 and 12 year-olds. The four-year cycle will cover all essential learning areas and essential skills in the curriculum.

Exit and entrance examinations – The National Certificate of Educational Achievement (NCEA), the senior secondary school national qualification, has been in place since 2002. There are three NCEA levels: Level 1 (Year 11 or fifth form); Level 2 (Year 12 or sixth form); and Level 3 (Year 13 or seventh form). A student is awarded a National Certificate when he/she has been successfully assessed against National Qualifications Framework standards and has accumulated sufficient credits. The NCEA replaced University Bursaries.

Assessment/examination formats – No information.

Evaluation – The Education Review Office (ERO) is a Crown agency responsible for reporting publicly on the quality of education.

National Monitoring of Standards is an externally administered assessment activity currently under development. The aim is to provide a broad picture of the achievements of representative samples of New Zealand students at successive points in time, so that trends in educational performance can be identified.

NORWAY

Governance and accountability – The primary national agency is the Ministry of Education of Education and Research. Decision-making has been decentralised to local levels, and schools have been granted greater autonomy.

The National Quality Assessment System (NQAS), was established in 2003 within the framework of the Knowledge Promotion Reform. Information about learning results, learning environment, use of material and human resources, and rate of progress at the upper secondary level provides the basis for feedback and improvement. Results from national assessments, final examinations at the end of Grade 10 and examinations at the end of upper secondary education are part of the NQAS. In June 2008 the government presented a White Paper on Quality in Education proposing to further develop and expand the national assessment system. The proposed system would add national tests in reading, mathematics and English at Grade 9 in lower secondary education (in addition to tests now in place for Grades 5 and 8. Mandatory diagnostic tests in reading and math would be introduced in the first Grade of upper secondary school. The improved monitoring system is intended to give reliable feedback and indicate where to target interventions.

Recent reforms and support for innovation – Norway developed a common core curriculum for all stages of primary, secondary and adult education in the 1990s. In 2004, the government initiated "The Knowledge Promotion Reform", focusing on primary and secondary education. New subject curricula have been introduced, and are adapted across levels in order to create coherent learning paths. The curricula are relatively ambitious, and it is expected that goal achievements of learners will vary. Teachers are expected to differentiate learning according to the needs of individual students. Curricula do not specify teaching methods or materials. The Reform also encourages more practical learning and collaboration between schools and external institutions and enterprises, so that students will have more exposure to working life.

Implementation of the Reform began in all schools in 2006/07. It emphasises: basic skills (reading, oral and written expression; mathematics; and ICT, to be integrated in across subjects/grades; an increase in the number of lessons in primary school (particularly Grades 1 to 4); the development of clear objectives and competence levels; greater flexibility in the organization of education.

National tests assessing pupils' basic skills in reading, writing, English, and mathematics (Grade 4 and 7) were introduced in the 2003-04 school year. Results from national tests are expected to give the teachers information for adapting teaching to individual pupils. New tests were also implemented in the 2007-08 school year.

Each school is responsible for developing its own innovations, assisted by local and regional authorities if necessary.

Exit and entrance examinations – Students take a matriculation examination.

Assessment/examination formats – At the end of Grade 10, students are required to take written national examinations in one of three subjects: Norwegian, mathematics or English. Most pupils also must take an oral examination, which is organised locally and may be in any of the academic school subjects

Evaluation – No information.

POLAND

Governance and accountability – The national agency, the Ministry of National Education prepares core curricula. Schools independently decide on the contents of the curricula, their scope and methods for implementing them. These decisions are made within schools, by a body of teachers, after consultations with parents and taking into account pupils' needs, and taking into account local conditions

Recent reforms and support for innovation – The Ministry of National Education initiated reforms in 1989, but only partial changes took effect in the early years. A comprehensive reform – including for vocational and continuing education – was introduced to reflect provisions in the new Constitution and the reform of state administration (1997). In 1999, the government launched a new six-year primary education programme, and established a general lower secondary school. Implementation of the reforms continued through 2004-05.

Recent reforms were also intended to respond to the lack of equal access to all levels of education, and the low percentage of students completing secondary and tertiary education; the need to adapt vocational education to the changing market economy; problems associated with the predominance of traditional "transmission" methods of teaching, rather than the development of students' skills and personality; and the need to establish closer links between schools, families and communities.

Innovations in curriculum have included: integrated teaching, subject blocks, and educational pathways. Schools may go beyond the traditional division of the teaching time into lessons. Teachers may also develop an programme if it complies with the core curriculum and has been approved by all teachers in the school.

The reform of the school system has introduced an external standardised test upon the completion of the primary school. This test is one of the conditions for graduation from primary school, but it is not used for selection purposes. An examination assessing students' skills and knowledge in the humanities, mathematics and natural sciences is administered in the third and final year of the gymnasium. The test assesses reading, writing, reasoning, use of information and practical application of knowledge. The results are used for information only, and do not give basis for any selection or ranking of pupils. At the beginning of 1999, the Central Examination Board (CEB) and eight Regional Examination Boards (REBs) were established to develop standards for examinations (in the case of the CEB) and to administer external examinations and tests (in the case of REBs) at the primary and secondary levels.

Exit and entrance examinations – The new matura examination was introduced in 2004-05. The external, standardised general examinations give students access to higher education.

Assessment/examination formats – The new matura examination includes an external written component (prepared and assessed by an external examination commission) and an internal oral component (prepared and assessed by school teachers).

Evaluation – The Ministry of National Education, local Kurators and school head teachers and directors provide pedagogical supervision to: assess performance quality, taking into account the student, parent and teacher views; provide assistance to teachers and head teachers; and, monitor compliance with current school legislation.

PORTUGAL

Governance and accountability – The Portuguese government has granted more autonomy to schools and colleges, more responsibility for those managing schools and for teaching staff and has introduced systematic assessment of those involved as part of a quality control for school management and teaching. The newly created School Council serves as an advisory body to the Ministry of Education. It is made up of the 60 chairmen of the executive councils. It is a representative body and is intended to ensure the effective participation of schools in defining and monitoring educational policy.

The new *Regime for School Autonomy, Administration and Management* places an emphasis on school level accountability for results.

Portugal conducts compulsory national examinations in mathematics and language for students in Year 9 of lower secondary school. The system for *National Assessments by Standardised Testing* has been improved and is being used more widely, as the best way of assessing national curricular developments and school performance.

Recent reforms and support for innovation – Since 2004, reforms have centred on ensuring greater quality and equity throughout the system. The reforms respond to demographic challenges and imbalances in regional development. Reforms have included: extension of the pre-school network, integration of students with special learning needs into the system; lowering rates of school drop-out and school failure; addressing literacy deficits; diversification of schools and encouraging schools to exercise autonomy fully; and the development of instruments for school evaluation and self-evaluation.

Portugal emphasises the need to develop innovative approaches to teaching in order to meet goals for equity and inclusion.

In addition, the system for *National Assessments by Standardised Testing* was improved and is now used more widely as a way to measure the quality of learning outcomes.

New curriculum guidelines for the first cycle of basic education have been introduced. The guidelines emphasise basic skills in Portuguese, mathematics, and environmental studies/science. Teachers have been trained in the use ICT with the aim of integrating it throughout the curriculum.

Exit and entrance examinations – Students participate in a general national examination at the end of their 12th year, in addition to examinations organised within their schools.

Assessment/examination formats – No information.

Evaluation – In 2002, the Portuguese government approved a *National Plan for the Assessment of Schools for Basic and Secondary Education.* There are external as well as internal school evaluations. Schools sign Autonomy Contracts. Since 1997, the Inspector General of Education has evaluated the quality of provision and management in schools. The Office of Educational Evaluation (*Gabinete de Avaliação Educacional-GAVE*), defines, plans, prepares and implements the control of evaluation tools.

SLOVAK REPUBLIC

Governance and accountability – The management and fiscal functions of the education system have been decentralised.

Recent reforms and support for innovation – Education in the Slovak Republic has been shaped by political, social, economic changes following the end of Communist rule in November 1989. The government has developed educational standards and a core curriculum for pre-primary, primary (Grades 1 to 4), and secondary school (Grades 5

to 8 or 9). New emphases include citizen participation in educational management; greater choice in education for parents and students; and the introduction of lifelong learning.

There has also been attention to the need for students to engage in independent and creative work , and some new methods have been introduced – such as Waldorf and Montessori workshops and pilot projects. In general, however, there have been few changes in school practice.

Greater school autonomy has been seen as a way to encourage grassroots initiatives and greater creativity in education and training. Since 1989, some curricula have been developed by teachers, primarily within the framework of "experiments" and with external international assistance.

Exit and entrance examinations – Students are admitted to secondary education upon completion of basic education and provided that they successfully pass the entrance examinations. At the end of secondary school, students participate in the matura examination. The matura give access to higher education.

Assessment/examination formats – Entrance examinations for secondary schools consist of oral and/or written tests. Subjects tested include Slovak language and mathematics, and sometimes foreign language. Students may apply for admission in more than one school.

The matura examination at the end of secondary school consists of written and oral examinations in language and literature in the mother tongue, and an oral examination in mathematics or foreign language, as well as two oral exams in optional subjects chosen by students. Two- and three-year vocational programmes at secondary vocational schools are completed with a final examination which consists of an oral and applied elements. Successful students receive an apprenticeship certificate. This certificate does not provide access to higher education.

Evaluation – The State School Inspection agency develops reports for the Minister of Education. The agency may include proposals for changes in the school system and, take corrective measures on the basis of identified problems.

SWEDEN

Governance and accountability – The central agency is the National Agency for Education. During the 1990s, the school system was decentralised, shifting to an more outcomes-based approach. New teaching and course curricula, outlining standard achievement levels were also introduced.

There are compulsory school national assessments in Swedish/Swedish as a second language, English and mathematics in the ninth year of schooling. There are also voluntary national course assessments in French, German, physics and biology in upper secondary schools. The national tests are used to develop a picture of student achievement, and also provide a basis for evaluation of school performance.

Recent reforms and support for innovation – As part of the process of decentralization, the government introduced a new curriculum across the school system in 1995. The government determines the syllabus and general orientation of the curriculum. Individual teachers decide the contents, teaching methods and materials to be used. The Swedish country report describes predominant modes of instruction as stressing the importance of students "learning-to-learn", and taking responsibility for their own learning. Group projects and interdisciplinary learning are increasingly common. Teachers are also increasingly collaborating and working in teams. Parts of upper secondary education are now conducted in a place of work, for at least 15 weeks. Schools are responsible for developing training opportunities, and for supervising students during the training period.

The curriculum sets out: goals to aim for (desired standards), and goals to be attained (minimum attainment for graduation). Schools set their own timetables. The latest curriculum allows schools to decide at which grade level subjects must be covered, creating room for the introduction of school and student options.

Exit and entrance examinations – In the 1990's, a new grading system and entrance requirements for secondary school were introduced. There are now seventeen three-year programmes (including a new programme in technology added in 2000) at the upper secondary level. Universities may select students on the basis of marks, university

aptitude examinations or other special examinations, and work experience. Applicants may also present special reasons for admission.

Assessment/examination formats – No information.

Evaluation – The National Agency for Education is responsible for the evaluation of the school system. The agency publishes an annual report, which includes data collected by Statistics Sweden and in-depth studies and evaluations (*Comparative figures for the school*). Evaluations aim at developing a deeper understanding of outcomes and relationships in school activities. The National Agency for School Improvement supports local schools in meeting national goals and ensuring quality and equity in education.

SWITZERLAND

Governance and accountability – The co-ordinating body is the Swiss Conference of canton directors of public instruction (*La Conférence Suisse des directeurs cantonaux de l'instruction publique* – CDIP). Switzerland's 26 cantons are decentralised. Each canton has its own structure and aims regarding education.

Recent reforms and support for innovation – Across Switzerland, educational objectives focus on acquisition of competences (knowledge, skills and attitude). The competences required in the world of work are also those used in new approaches to teaching and learning: autonomy, teamwork, responsibility, English and a second national language (of Switzerland's four official languages: French, German, Italian and Romansh). In addition, there is an emphasis in learning by doing – for example, a participating in an activity in a developing country, linguistic exchange, and so on. These projects have had an important influence on schools.

The Swiss country report recognises that lower secondary schools are attempting to balance two objectives that often are in conflict: on the one hand, providing students the time to choose their direction, and on the other hand, preparing them for the selective upper secondary schools. Cantons have taken diverse approaches to balancing these different goals.

While educational aims are set at the level of the canton, those set out in the Geneva canton are illustrative: creating a desire for lifelong learning; helping each student to develop his own personality, creativity, intellectual and physical abilities; preparing students to participate in social, cultural, civic, political and economic life of the country; developing respect for others and spirit of solidarity and co-operation; and, providing students with equality of opportunity.

While there is diversity in the organisation of schools, there are also common approaches. All cantons teach the language of the region, mathematics, a second national langue, natural sciences, geography, history, civic education, art and music, physical education. Teaching methods are varied. Recent reforms have had more of an impact at the primary level than the secondary level.

Exit and entrance examinations – Lower secondary schools are divided into two to four different selective, courses of study. Entrance requirements vary by canton (examination, grades, interview, parents' wishes, and so on). Students may either go on to upper secondary school (for which students must pass an admission examination), or professional school. Upper secondary schools prepare students for university. Students study languages (first, second, and ancient), social sciences, experimental science and mathematics, art, and physical education. They may also participate in transdisciplinary studies. Students must also complete a major project – either independently or with a team.

Assessment/examination formats – Under new rules, students take matriculation examinations in five disciplines. The examinations are written, and may also be taken as oral examinations. Grades from the last year of school are taken into consideration.

Evaluation – In Switzerland, a culture of evaluation is only beginning to develop.

TURKEY

Governance and accountability – The national agency is the Ministry of National Education. The system is centralised. With the development of new curricula, new standards of education and learning and teaching methods have been adopted.

Turkey has developed new standards, and assessments and evaluations along with a new curricula (see below). The UNESCO country report does not provide any detail on the new assessments/evaluations. According to the OECD's *Education at a Glance* (2008), Turkey conducts national assessments of student performance in mathematics, science, language and other subjects for students in Years 6, 7 and 8 of lower secondary school (no data are gathered for upper secondary school).

Recent reforms and support for innovation – The curriculum developed in 2005 emphasises eight core skills, including: critical thinking, creativeness, communication, research and causality, problem solving, information and communication skills, entrepreneurship and language skills. These eight core skills are to integrate basic literacy and mathematics skills. Lifelong learning and social inclusion are also emphasised. The curriculum has moved from an emphasis on the traditional "informative"/transmission approaches to "constructivist"/student-centred teaching methods. New textbooks and teachers' manuals have been prepared to reflect principles of the new curricula. Some new courses were put into place immediately following the introduction of the curriculum, while others were first developed as pilot projects.

Exit and entrance examinations – Students take a matriculation examination. Students applying for university must take a national examination (conducted by the Student Selection and Placement Centre). High-school grade point averages are also considered.

Assessment/examination formats – The university entrance examination is a two-stage examination. The first stage, the Student Selection Examination, consists of verbal and quantitative sections. Candidates scoring between 105 and 120 points have a restricted choice of higher education programmes. Those with a minimum score of 120 may take the second stage of the examination, which includes: natural sciences, mathematics, Turkish, social sciences and a foreign language. Admission to higher education is based on the composite score of the two examinations.

Evaluation – No information.

UNITED KINGDOM

Governance and accountability – The education and training systems of England, Wales and Northern Ireland are broadly similar. The National Curriculum setting out content and standards for learning was introduced into schools in England and Wales in 1988. National assessments in English and mathematics are given to students at 7, 11 and 14 years of age. Students are assessed in science at ages 11 and 14. Results are published for each child, each school and the national averages for comparison purposes. In Wales, the results of National Curriculum assessments are published nationally and for each local authority area. The national agency in England is the Department for Children, Schools and Families. In Wales, the national body is the Ministry of Education and Training, and in Northern Ireland, the national agency is the Department of Education.

The education system in Scotland has always been completely separate with its own laws and practices. There are national tests in language and mathematics for students aged 5 to 14. Assessments are at five levels of difficulty, and take the form of a large bank of units from which teachers may choose test materials to fit their own teaching programmes and contexts. In Scotland, the national body is the Scottish Executive, Education Department.

Recent reforms and support for innovation – England's *Five-year Strategy for Children and Learners* (2004-09) includes reforms of school budget, school building and management issues, the establishment of "school improvement partners" (usually a head teacher from a successful school); universal specialist schools, better specialist schools and "foundation partnerships", allowing groups of schools to work together; and allowing high-performing

specialist schools to serve as training schools or leaders of partnerships.

Schools have some discretion over when to start teaching each key stage. In England, six key skills are embedded in the National Curriculum, including: communication, numeracy, ICT, collaboration, improving one's own learning and performance, and problem-solving. Wales' comprehensive ten-year strategy document (*The Learning Country*) includes plans to develop schools' capacity for innovation, flexibility and responsiveness, and strengthens support for teachers.

The **Northern Ireland** school system has gone through substantial changes since the Education Reform Order, 1989. More recently, the Education Order, 2006, introduced revised statutory curriculum, to be phased in between September 2007 and 2009. The revised curriculum is less detailed, and more oriented toward skill development. In addition to literacy, numeracy and ICT, the curriculum emphasises skills for thinking, problem solving and decision making; self management; working with others; managing information; being creative. Economic awareness and careers education are taught in secondary schools. Statutory assessment arrangements for Key Stages 1 to 3 are to be replaced by standardised annual reports or "pupil profiles".

In Scotland, the Scottish Executive Education Department (SEED) is responsible for the curriculum. The curriculum includes: language, mathematics, environmental studies, expressive arts, religious and moral education, and flexible studies. The new *Curriculum for Excellence* develops a streamlined curriculum for 3 to 18-year-olds along with new approaches to assessment. Following a familiarization phase in 2007, the implementation of the reviewed curriculum was expected to begin in August 2008. The new curriculum is based on seven principles, including: *challenge and enjoyment*; *breadth*; *progression*; *depth*; *personalisation and choice*; *coherence*; and *relevance*. Curricular areas include: expressive arts; health and well-being; language; mathematics; religious and moral education; science; social studies; and technologies. There are also cross-curricular activities. Students may take a short modular courses as an alternative to courses leading to higher and advanced higher level examinations.

Exit and entrance examinations – In England, Wales and Northern Ireland, the General Certificate of Secondary Education (GCSE) is principal examination taken by secondary school pupils at 16 years and older. The GCSE is a single-subject examination. GCE A-level (Advanced) examinations may be taken following after two further years of study. Students typically take courses in two or three related subjects (for example, sciences and the humanities) and are assessed in these areas.

In Scotland, students may take their first National Qualifications examinations, either Standard Grade or the newer Intermediate levels 1 and 2 of the National Qualifications (General and Credit Standard Grade, respectively). Achievements are recorded on the Scottish Qualifications Certificate (SQC, formerly the Scottish Certificate of Education).

Assessment/examination formats – The GCSE exam emphasises coursework, examination and application of knowledge. School-based assessment of coursework by teachers may be used for a significant percentage of the final results.

Evaluation – In 2007, England unified the Office for Standards in Education, Children's Services and Skills (OFSTED), under Her Majesty's Chief Inspector of Education, Children's Services and Skills (HMCI). The new organisation is still known as OFSTED. The Education and Inspections Act 2006 gives local authorities in England new powers to require failing and underperforming schools to federate or take another partner for school improvement. Schools are responsible for self-evaluation and for their own improvement. The central government and from local authorities provide support for school self-evaluation.

The functions of Her Majesty's Inspectorate for Education and Training in Wales are similar to those of OFSTED in England. In Northern Ireland, the Education and Training Inspectorate has sole responsibility for the inspection of education and training. In Scotland, Her Majesty's Inspectorate of Education (HMIE) supports improvements in educational standards and quality. HMIE evaluates pre-school and community learning, schools, colleges, teacher education and local authorities.

UNITED STATES

Governance and accountability – The national agency is the Department of Education. The education system in the United States is decentralised. With few exceptions, the federal government is legally forbidden from intrusion into state and local curriculum decisions, and may only involve itself in school management in a limited set of circumstances (such as enforcement of civil rights). State-level policy makers generally establish graduation requirements, select classroom texts, develop state-level assessments, and provide curriculum guidelines and technical assistance. In some states, laws are prescriptive, and in others, local districts have considerable flexibility. Local districts may add their own requirements.

States were required to adopt content standards in reading and mathematics by the 1994 Elementary and Secondary Education Act (ESEA). The No Child Left Behind (NCLB) Act of 2001, re-authorizing the ESEA, required states to assess: English proficiency (reading, writing and oral skills) of all English language learners beginning in the 2002-03 school year), all students in Grades 3-8 in mathematics and reading or language arts on an annual basis (beginning in the 2005-06 school year). By the 2007-09 school year, science assessments were put into place. They are to be administered at least once between Grades 3-5, Grades 6-9 and Grades 10-12.

Recent reforms and support for innovation – The current wave of reforms is strongly focused on strengthening of standards and accountability, as described above. The reforms were catalyzed by publication of the report, *A Nation at Risk* (1983). The report recommended that states: strengthen graduation requirements, develop more rigorous and measurable standards, require students to spend more time in school, and for systems to invest in improvement of teaching.

NCLB gave state and local education agencies more flexibility in the use of federal funding, and freedom to implement innovations. In exchange, states were required adopt standards in language, mathematics and science (most states had already adopted language and mathematics standards by 1996/97). States may also adopt standards in other content areas as they see fit. Most states have developed *academic standards* describing what students should know and be able to do, *content standards* describing the body of knowledge students should have in mathematics, language arts, science and social studies. They have also developed *performance standards* describing basic, proficient and advanced performance levels. There are ongoing debates as to whether standards are too high, too low, sufficiently clear or relevant.

States must demonstrate that they have adopted a single, statewide accountability system, defining "adequate yearly progress" (AYP) for all public school students. All students must achieve a state-defined "proficient" level by the 2013-14 school year. States must put in place benchmarks and report progress toward goals. AYP must also address the progress of specified subgroups of students (for example, by race, ethnicity, migrant status, special education status, and so on). Interventions to be used in the case of low school performance are specified (for example, financial sanctions or school closure). The NCLB requires states and school districts to give parents easy to read "report cards" on which schools and districts are succeeding and why. NCLB required that all teachers of core academic subjects be "highly qualified" by the end of the 2005-06 school year.

At least 40 of the 50 states also have some form of minimum competency assessments. These assessments serve as a means to influence local curricula, identify problems and initiate remediation. States currently conduct minimum-competency assessments in reading, writing and mathematics for students in Grades 3 or 4 in primary school; Grades 6, 8, or 9 in lower secondary school; and Grades 11 or 12 in upper secondary school.

At the national level, the National Assessment of Educational Progress (NAEP) assesses students' knowledge in reading, writing, science, mathematics and other subjects in Grades 4 and 8, and tracks short and long-term trends. Beginning in the 2002-03 school year, participation in NAEP was made mandatory.

Exit and entrance examinations – The Scholastic Assessment Test (SAT) tests vocabulary and reasoning skills. About 1.1 million students take the test each year. The American College Testing Programme (ACT) assesses social

studies, natural sciences, mathematics and English. About 800 000 students take this test each year.

Assessment/examination formats – No information available in the reports consulted.

Evaluation – There is no central inspectorate or monitoring system.

[13] Country reports from Mexico and Spain were available in Spanish only. Summaries of these two countries are not included in this annex.

EXISTING OECD EDUCATION WORKING PAPERS

No. 1 *Teacher Demand and Supply: Improving Teaching Quality and Addressing Teacher Shortages* (2002), Paulo Santiago.

No. 2 *Teacher Education and the Teaching Career in an Era of Lifelong Learning* (2002), John Coolahan.

No. 3 *Towards an Understanding of the Mechanisms That Link Qualifications and Lifelong Learning* (2003), Friederike Behringer and Mike Coles.

No. 4 *Measuring Educational Productivity in Standards-Based Accountability Systems: Introducing the SES Return on Spending Index* (2005), Martin Hampel.

No. 5 *PISA 2000: Sample Weight Problems in Austria* (2006), Erich Neuwirth.

No. 6 *Funding Systems and their Effects on Higher Education Systems – International Report* (2007), Franz Strehl, Sabine Reisinger and Michael Kalatschan.

No. 7 *On the Edge: Securing a Sustainable Future for Higher Education* (2007), OECD/IMHEHEFCE.

No. 8 *Globalisation and Higher Education* (2007), Simon Margison and Marijk van der Wende.

No. 9 *Understanding the Regional Contriubution of Higher Education Institutions: A Literature Review* (2007), Peter Arbo and Paul Benneworth.

No. 10 *Effects of Tertiary Expansion – Crowding-out Effects and Labour Market Matches for the Higher Educated* (2007), Bo Hansson.

No. 11 *Skilled Voices? Reflections on Political Participation and Education in Austria* (2007), Florian Walter and Sieglinde K. Rosenberger.

No. 12 *Education and Civic Engagement: Review of Research and a Study on Norwegian Youths* (2007), Jon Lauglo and Tormod Oia.

No. 13 *School Accountability, Autonomy, Choice, and the Level of Student Achievement: International Evidence from PISA 2003* (2007), Ludger Wössmann, Elke Lüdemann, Gabriela Schütz and Martin R. West.

No. 14 *School Accountability, Autonomy, Choice, and the Equity of Student Achievement: International Evidence from PISA 2003* (2007), Gabriela Schütz, Martin R. West, Ludger Wössmann.

No. 15 *Assessment of learning outcomes in higher education: a comparative review of selected practices* (2008), Deborah Nusche.

No. 16 *Approaches and Challenges to Capital Funding for Educational Facilities* (2008), Ann Gorey.

No. 17 *Recent Developments in Intellectual Capital Reporting and their Policy Implications* (2008), W. Richard Frederick.

No. 18 *Employers' Perspectives on the Roles of Human Capital Development and Management in Creating Value* (2008), L. Bassi and D. McMurrer.

No. 19 *Job-related Training and Benefits for Individuals: A Review of evidence and explanations* (2008), Bo Hansson.

No. 20 *A Framework for Monitoring Transition Systems* (2008), Rolf van der Velden.

No. 21 *Final Report of the Development of an International Adult Learning Module (OECD AL Module)* (2008), Bo Hansson and Helmut Kuwan.

No. 22 *What Works in Migrant Education? A Review of Evidence and Policy Options* (2009), Deborah Nusche.

RECENT OECD PUBLICATIONS OF RELEVANCE TO THIS WORKING PAPER

OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris.

OECD (2008), *Teaching, Learning and Assessment for Adults: Improving Foundation Skills*, OECD Publishing, Paris.

THE OECD EDUCATION WORKING PAPERS SERIES ON LINE

The OECD Education Working Papers Series may be found at:

- The OECD Directorate for Education website: www.oecd.org/edu/workingpapers
- The OECD's online library, SourceOECD: www.sourceoecd.org
- The Research Papers in Economics (RePEc) website: www.repec.org

If you wish to be informed about the release of new OECD Education working papers, please:

- Go to www.oecd.org
- Click on "My OECD"
- Sign up and create an account with "My OECD"
- Select "Education" as one of your favourite themes
- Choose "OECD Education Working Papers" as one of the newsletters you would like to receive

For further information on the OECD Education Working Papers Series, please write to: edu.contact@oecd.org.