

**DIRECTORATE FOR EDUCATION
EDUCATION POLICY COMMITTEE**

EDU/EDPC(2009)3/REV1
For Official Use

OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes

Design and Implementation Plan for the Review

This project plan incorporates feedback on an earlier version of this document [EDU/EDPC(2009)3], which was discussed at the first meeting of the Group of National Experts on Evaluation and Assessment, held on 1-2 October at the OECD Headquarters in Paris [EDU/EDPC/EA/M(2009)1].

Contact:
Paulo Santiago
Education and Training Policy Division [Tel: +33(0) 1 45 24 84 19; e-mail: paulo.santiago@oecd.org]

JT03271962

TABLE OF CONTENTS

BACKGROUND	2
Introduction	3
Governance of the Review	3
Purpose	4
Context	5
Analytical framework.....	6
KEY ISSUES AND QUESTIONS FOR ANALYSIS.....	10
Issue 1 - Designing a systemic framework for evaluation and assessment.....	10
Issue 2 - Ensuring the effectiveness of evaluation and assessment procedures	11
Issue 3 - Developing competencies for evaluation and for using feedback	12
Issue 4 - Making the best use of evaluation results.....	13
Issue 5 - Implementing evaluation and assessment policies	13
METHODOLOGY	15
Main Components of the Review	15
Collaboration.....	17
Outputs	18
Dissemination.....	18
Participation and Costs.....	18
REFERENCES	19
ANNEX 1: GROUP OF NATIONAL EXPERTS ON EVALUATION AND ASSESSMENT DRAFT MANDATE	20
ANNEX 2: PROVISIONAL TIMELINE FOR THE REVIEW	21
ANNEX 3: NATIONAL CO-ORDINATOR FOR COUNTRY REVIEWS.....	22
ANNEX 4: ILLUSTRATIVE BUDGET.....	23

BACKGROUND

Introduction

1. Performance in schools is increasingly judged on the basis of effective learning outcomes. Information is critical to knowing whether the school system is delivering good performance and providing feedback is the channel through which performance can be improved. Increasingly, countries are developing a range of tools and techniques for evaluation and assessment in school systems as part of their efforts to improve student outcomes.

2. The OECD Review on *Evaluation and Assessment Frameworks for Improving School Outcomes*¹ will provide analysis and policy advice to countries on how these different assessment and evaluation tools can be embedded within a consistent framework to bring about real gains in performance across the school system.

3. This project plan for the Review sets out the analytical framework, issues for analysis, outputs, methodology, timeline and resources required. The Review will seek to make the most of the OECD's greatest strengths – providing a setting where governments can compare policy experiences, seek answers to common problems, and identify and share good practices.

4. This revised project plan incorporates feedback on an earlier version of this document [EDU/EDPC(2009)3], which was discussed at the first meeting of the Group of National Experts on Evaluation and Assessment, held on 1-2 October at the OECD Headquarters in Paris [EDU/EDPC/EA/M(2009)1]. It seeks to respond closely to the needs and expectations of countries.

Governance of the Review

5. The project is overseen by the Group of National Experts on Evaluation and Assessment, which was established as a subsidiary body of the Education Policy Committee in order to guide the methods, timing and principles of the Review and to allow countries to share information and experience on this issue and to keep in touch with emerging findings of the exercise (see Mandate in Annex 1).

6. All OECD Member countries and Observers to the Education Policy Committee may participate in the Group. Countries themselves choose who should represent them at the meetings of the Group of National Experts. However, countries may wish to take into account two factors in selecting their representatives. *First*, each person attending the meeting is assumed to be competent to speak on behalf of their country and this suggests someone responsible for policies on evaluation and assessment within education ministries. *Second*, countries may find it useful to send several people to the meetings, especially as evaluation and assessment involves different parts of the system, for example, agencies conducting quantitative assessments and school inspectorates.

7. Three meetings of the Group are planned during the Review with the following objectives:

- The first meeting discussed the key policy issues and fine-tuned the scope for the Review (1-2 October, 2009).

1. This Review is designed to deliver the Output Result “Policy analysis on evaluation and assessment frameworks to improve school outcomes” for the Education Policy Committee’s Programme of Work and Budget for 2009-2010 [EDU/EDPC(2008)27/REV1 and EDU(2008)6/REV2] and will contribute to the Medium Term Strategy theme of raising teaching and learning effectiveness [EDU/EDPC(2007)13/FINAL].

- The second meeting will discuss policy issues based on evidence collected and share country experience on specific issues (Q2 2010).
- The third meeting will discuss the preliminary synthesis report and share country experience on key themes of the report (Q4 2010).

8. Progress on the Review will be reported back to the Education Policy Committee at regular intervals using the framework established for all outputs. The Committee will also be invited to comment on the draft synthesis report.

9. The timeline for the project is set out in Annex 2.

Purpose

10. The Review is designed to respond to the strong interest in evaluation and assessment issues evident at national and international levels. The overall purpose is to explore how systems of evaluation and assessment can be used to improve the quality, equity and efficiency of school education. The analysis will focus on primary and secondary levels of education.²

11. Evaluation and assessment in school systems are increasingly important in national agendas. There is widespread recognition that evaluation and assessment arrangements are key to both improvement and accountability in school systems. However, countries often face difficulties in implementing evaluation and assessment procedures. These may arise as a result of poor policy design, lack of analysis of unintended consequences, little capacity for educational agents to put procedures into practice, lack of an evaluation culture, or deficient use of evaluation results.

12. The Review is designed to help countries assess:

- How to develop a coherent evaluation and assessment framework within which different elements, each serve their purpose well and together successfully combine improvement and accountability to improve student outcomes.
- How to strengthen the use of evaluation and assessment results to enhance school practices so student outcomes are improved.
- How to facilitate the implementation of evaluation and assessment policies.

13. It will provide a description of design, implementation and use of assessment and evaluation procedures in countries; analyse strengths and weaknesses of different approaches; and provide recommendations for improvement including of how results should be incorporated into policy and practice. In doing this it will:

- Synthesise research-based evidence on the impact of evaluation and assessment strategies and disseminate this knowledge among countries.
- Identify innovative and successful policy initiatives and practices.
- Facilitate exchanges of lessons and experiences among countries.
- Identify policy options for policy makers to consider.

² The scope of the Review does not include early childhood education and care, apprenticeships within vocational education and training, and adult education.

14. The Review will look at the various components of assessment and evaluation frameworks that countries use with the objective of improving the student outcomes produced by schools. These include student assessment, teacher appraisal, school assessment, system evaluation and also other elements such as the evaluation of a programme. The Review will examine the range of evaluation instruments from summative assessment, “value-added” or “growth” models to formative assessment and self-evaluation, and also other kinds of feedback such as public consultations, national education reports and student surveys. The focus of the analysis on the formative assessment of students will be on its integration with the remaining components of the evaluation and assessment framework (*i.e.* the use of results from the formative assessment of students within the broader evaluation and assessment framework rather than how it is undertaken within the classroom). Considerations on formative assessment of students *per se* will draw on CERI’s previous work on the theme.

Context

15. The Review is intended to extend and add value to the existing body of international work on evaluation and assessment policies. It will partly build on the Education Policy Committee’s work on “Accountability Systems and Evaluating Education Performance” (OECD, 2008a; 2008b). The importance of evaluation and assessment issues is also reflected in a wide variety of other OECD outputs.³

16. As countries strive to transform their educational systems to prepare all young people with the knowledge and skills needed to function in rapidly changing societies, some common policy trends can be observed in one form or another in most OECD countries: decentralisation, school autonomy, greater accountability for outcomes and a greater knowledge management capacity. Decentralisation and school autonomy are creating a greater need for school assessment, while greater IT capacity also allows for the development and analysis of large scale assessments as well as individualised assessment approaches. Evaluation and assessment procedures are also now instrumental in defining strategies for improving school practices with the ultimate goal of enhancing student outcomes. These developments are having a strong influence in the way in which policy makers monitor school and student performance and respond to it.

17. Recent analysis carried out for the OECD indicates that student assessment systems appear to be related to student performance as measured by PISA (Wößmann *et al.* 2007a, 2007b). These results suggest that there is a good case for policy-makers, education professionals, parents and children to pay attention to arrangements for evaluation and assessment.

18. The purpose of feedback on educational performance, at several levels, is to provide information on how well systems and their actors are achieving educational goals and objectives. Goals and objectives vary from country to country but they are typically expressed in terms of quality, equity and efficiency of systems themselves, their contribution to the general quality of life, the sharing of social and economic progress, and values that underpin social cohesion.

19. Countries are increasingly using evaluation and assessment as a tool for understanding better how well students are learning, for providing information to parents and society at large about educational performance and for improving school practices. Strong emphasis is being placed on better equipping and encouraging teachers to carry out self-evaluation and formative assessment, on providing the incentives

3. PISA, TALIS, *Education at a Glance, Measuring Improvements in Learning Outcomes*, CELE’s work on evaluating quality in educational facilities, CERI’s work on formative assessment, schooling for tomorrow, market mechanisms in education, linking research and policy, and systemic innovation in education, and policy outputs addressing teacher policy, school leadership, migrant education, vocational education and early childhood education.

and means for school self-appraisal, on encouraging 'value-added' evaluation and on more regular testing of students and national monitoring of the overall system.

Analytical framework

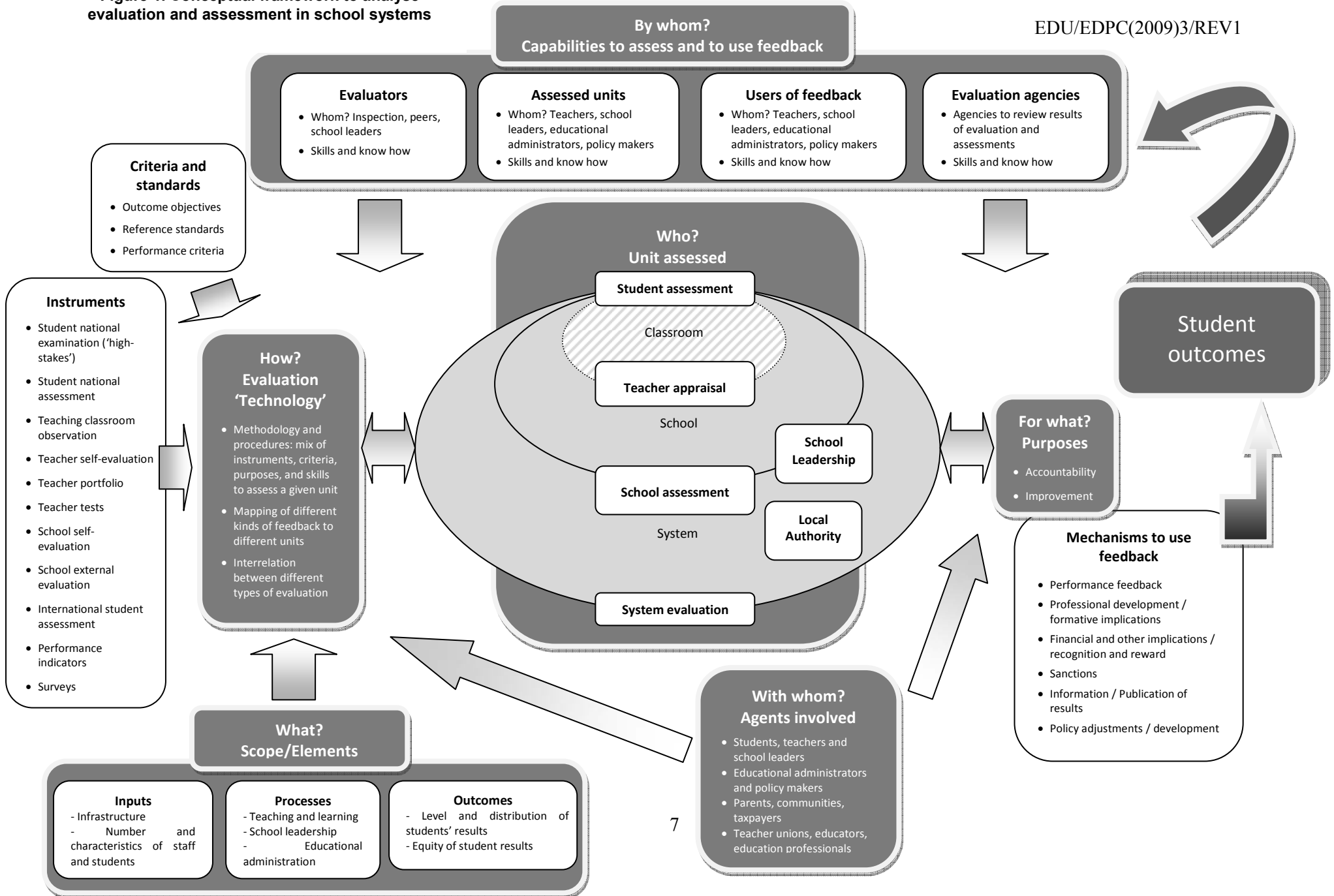
Conceptual framework to analyse evaluation and assessment

20. A key challenge is to understand the complex range of features associated with evaluation and assessment frameworks. Figure 1 provides a conceptual framework summarising the aspects involved and the way they interconnect. The over-arching policy objective is to ensure that evaluation and assessment frameworks contribute to the improvement of student outcomes through enhanced school practices. The conceptual framework has six main interrelated aspects.

- Unit Assessed: *Who?* The subject of the evaluation could include the whole education system or parts of it, the evaluation of given schools, individual teacher appraisal or the assessment of an educational local authority. This central aspect refers to the kind of evaluation process under analysis (e.g. teacher appraisal; school assessment). The nature of the remaining five aspects depends on the unit being evaluated.
- Capabilities to assess and to use feedback: *By Whom?* This aspect concerns the preparation to evaluate, to be evaluated and to use the results of an evaluation as well as the choice of the groups undertaking these functions. It includes issues such as: the choice of the evaluators and the development of the skills to perform an assessment; the preparation to be the subject of an evaluation; the development of competencies to effectively use the results of an evaluation for the improvement of practice; and the design of agencies to review evaluation results with a view to hold agents accountable and to inform policy development.
- Scope/Elements: *What?* Some evaluation processes concentrate on analyses of outcomes such as the level and distribution of students' results. Other evaluation initiatives centre on processes such as the quality of teaching and learning or the effectiveness of school leadership. In addition, inputs such as the quality of the infrastructure can also be the subject of an assessment (nonetheless the Review will not examine evaluation arrangements to assess the quality of school infrastructure).
- Evaluation 'technology': *How?* This aspect refers to the organisation of particular types of evaluation, that is the mix of instruments, criteria and standards, purposes, skills, and scope which are used to undertake a given evaluation or assessment. For instance, a teacher appraisal model may be based on a range of instruments such as self-evaluation, classroom observation and a teacher portfolio; be focussed on the teaching and learning process; be undertaken in relation to reference standards for the teaching profession; have both improvement and accountability purposes; and be based on experienced peers. Hence, this aspect refers to the way different aspects are combined to produce a given evaluation model.
- Purposes: *For what?* This encompasses the objectives of a particular evaluation process and the mechanisms designed to ensure that evaluation results are used in a way such objectives are reached. The objectives of an evaluation process typically consist of improvement and accountability. Examples of mechanisms to use evaluation results include performance feedback, professional development plans, financial and other rewards, publication of results to the general public, and policy adjustments.
- Agents involved: *With whom?* This mostly deals with the political economy of reform aspects of evaluation and assessment procedures. It relates to the involvement of a range of stakeholders such as parents, students, teachers, school leaders, teacher unions, educational administrators and policy makers in the development and implementation of evaluation and assessment processes.

Figure 1. Conceptual framework to analyse evaluation and assessment in school systems

EDU/EDPC(2009)3/REV1



Framework for analysis

21. Evaluation and assessment issues cannot be studied in isolation. Societal, school system, and school-level factors all influence the design of evaluation and assessment policies. These factors influence the design of evaluation and assessment policies in terms of the needs for new policy initiatives, the factors that constrain policy opportunities, and the factors that influence policy implementation, impact and cost. Without an adequate understanding of the range of factors involved, and the ways they influence the impact of evaluation and assessment policies, there is a risk of developing ineffective policy responses. Part of the challenge of the Review will lie in understanding the interactions between factors at these three levels and evaluation and assessment procedures.

22. Figure 2 outlines the framework for analysis for the Review, interlinking the factors influencing evaluation and assessment policies, the key issues for analysis and the main analytical components of the work.

23. Examples of societal factors are:

- Political environment
- Demography and cultural diversity
- Economic conditions, labour market trends
- Role of media and general perceptions of schooling
- Public and private resources for schooling
- Evaluation and assessment policies in the public sector

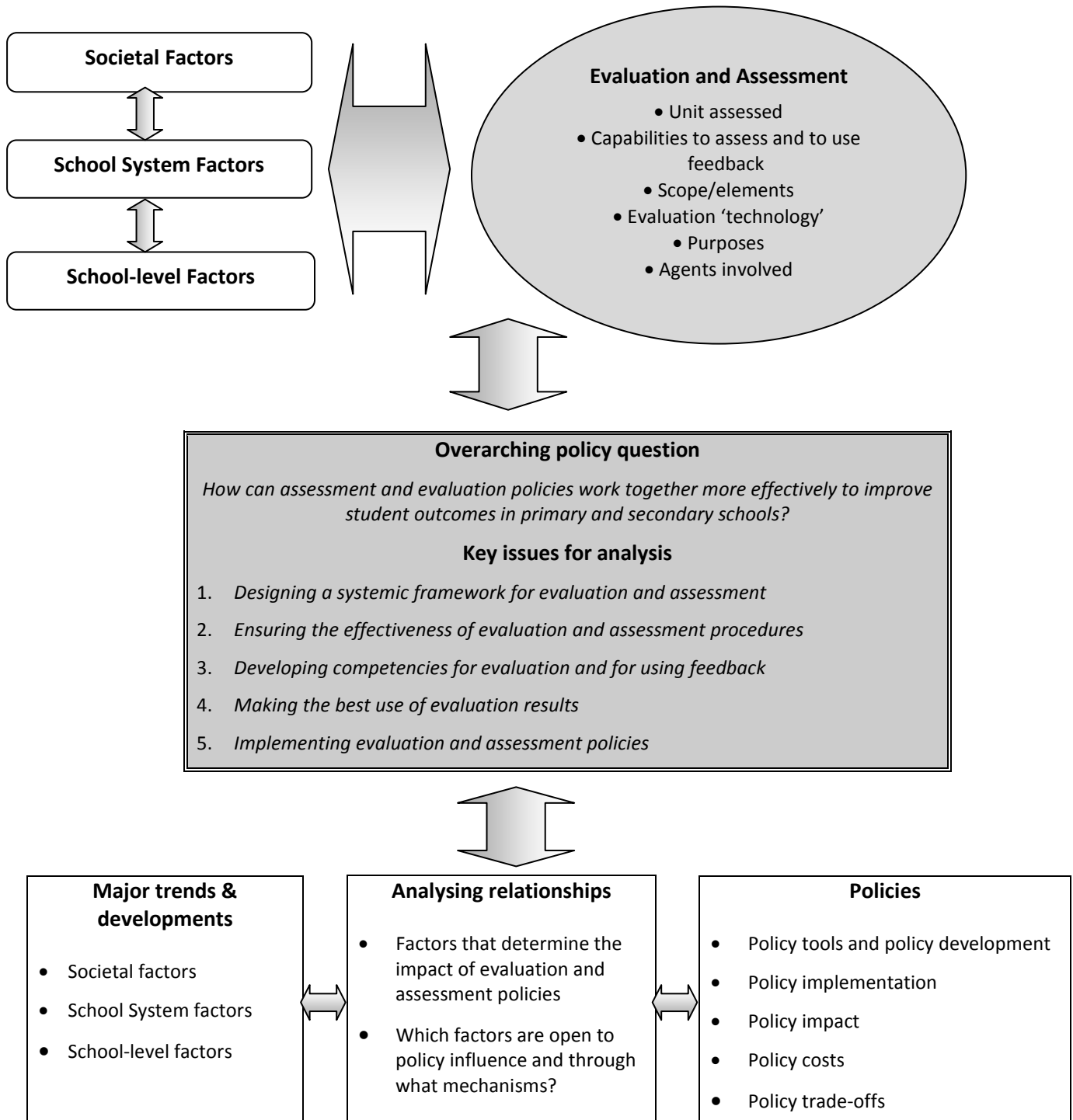
24. Examples of school system factors are:

- Distribution of responsibilities across educational authorities
- School governance and autonomy
- Structure of schooling
- Curriculum, academic standards, learning time
- Market mechanisms / school choice
- Allocation of resources within school system
- Career and reward/incentive structures for teachers and school leaders
- Teacher education and certification policies
- Roles and views of stakeholder groups

25. Examples of school-level factors are:

- Socio-economic context of schools
- Management, leadership and distribution of responsibilities within school
- Professional activities of teachers and other school personnel
- Teaching and learning arrangements
- Community and family involvement
- Learning conditions, support structures

Figure 2. Framework for Analysis



KEY ISSUES AND QUESTIONS FOR ANALYSIS

26. The overarching policy question on which the analysis will be focussed is “How can assessment and evaluation policies work together more effectively to improve student outcomes in primary and secondary schools?”

27. The Review will further concentrate on five key issues for analysis. To ensure the effectiveness of evaluation mechanisms in school systems, the following elements seem necessary:

- that an overall, systemic framework for evaluation and assessment be designed to make sure that its various components complement each other and articulate coherently to achieve given purposes (issue 1);
- that evaluation procedures provide an accurate and reliable measure, thanks to appropriate evaluation techniques which are adequate to the intended purpose of the feedback (issue 2);
- that proper competencies be developed and be provided to the agents in charge of the evaluation process (evaluators and evaluated), of interpreting evaluation results and of elaborating improvement strategies based on these results (issue 3);
- that the channel from evaluation results to drawing conclusions and developing adapted policy response be in place and reactive, so that the best use is made of gathered data (issue 4);
- that consensus be built among all actors around the importance, stakes and implementation options of evaluation, so that cooperation can be achieved and evaluation systems be effective (issue 5).

Issue 1 - Designing a systemic framework for evaluation and assessment

28. Designing a systemic framework for evaluation and assessment entails a range of aspects. First, it needs to be framed in the context of the overall objectives for schooling and the approach to its development depends on a range of established practices in the school system such as the extent of school autonomy, the existence of national curricula and standards, or the culture of evaluation. In particular, it needs to be clear what aspects evaluation and assessment strategies seek to monitor and improve. There is the increasing recognition that the monitoring of student outcomes must extend beyond knowledge skills in key subject areas and include broader learning outcomes, including students’ critical thinking skills, social competencies, engagement with learning and overall well-being. School systems are also concerned not only with levels of performance but also with its distribution (*i.e.* equity aspects) and the progress students and schools are making.

29. Second, the purposes of the framework need to be clearly defined. For instance, evaluation and assessment can be part of the governance structures that balance school autonomy with accountability to parents and society at large. They can also be used as diagnostic tools to generate the improvement of practices and policies in the school system. The tension between improvement and accountability and the way it impacts the design of evaluation strategies will receive particular attention in the analysis. Overall, evaluation and assessment contribute to creating knowledge-rich schooling systems in which educational agents act as partners and have the authority to act, the necessary information to do so, and benefit from support structures to generate improvement.

30. Third, there needs to be a clear understanding of the responsibilities of the different educational actors within the evaluation and assessment framework. Educational authorities at several levels, agencies

in charge of quality assurance, schools, parents and school communities, teachers and students play different roles in ensuring improvement and accountability in the school system.

31. Fourth, there needs to be a reflection on the different components of the framework such as school assessment, teacher appraisal, or standardised national-level student tests to assess students' progress, and ways in which they can be articulated to achieve the purposes of the framework. A particularly important aspect is how the different components have to be interrelated in order to generate complementarities, avoid duplication, and prevent inconsistency of objectives.

32. The analysis will be organised around the following questions:

- How can educational authorities design a coherent evaluation and assessment framework to achieve given purposes?
- How should responsibilities for evaluation and assessment be shared across the educational system?
- How can the multiple evaluation procedures within the evaluation and assessment framework be interrelated to achieve best results?
- How can educational authorities design a strategy for evaluation and assessment that successfully combines the purposes of accountability and improvement?

Issue 2 - Ensuring the effectiveness of evaluation and assessment procedures

33. Effective evaluation is challenging at several levels: accuracy of the measurement, inclusion of all the dimensions of what is meant to be measured, consistency with the final goals of the feedback exercise, adaptation to the needs of those who will use the results (parents, teachers, decision makers,...), cost-effectiveness for the assessment agency and for the unit observed, practical feasibility and adaptation to the unit assessed.

34. Performance measurement calls for a cautious choice of measurement instruments and techniques. For instance, learning is the result of a joint effort, and is conditioned by uncontrolled factors such as the socio-economic background. So measuring the performance of one element that is part of the joint effort is a particularly arduous question, which has been addressed more and more precisely (*e.g.* value-added models for school assessment, OECD 2008c). This illustrates the relevance of the comparative efficiency of various assessment procedures for a given evaluation.

35. Especially if an evaluation is used for accountability, it needs to be designed to induce a minimum of bias in the behaviour of those assessed. In particular, school systems may suffer perverse effects from ill-designed instruments which leave aside some dimensions of what is meant to be measured, inducing actors to over-allocate resources on what affects one performance indicator.

36. Actors at various levels of the school system have different needs for feedback, and evaluation techniques need to be designed accordingly. Students and teachers may need quicker and more frequent feedback, helping determine deficiencies when objectives are not met. Students and parents need extensive information at specific juncture points in time, for orientation matters, showing how curricula and places match better a given student profile. Policy makers may need longer term evaluation to assess a programme or the effectiveness of a subsidiary body, with indications on how resources should be allocated to achieve various goals. Schools need information that helps them improve collaboration between teachers and raise aspirations of students.

37. A multitude of practical arrangements for the process of evaluation itself have been experimented at local, national and international levels. Improving evaluation effectiveness implies choosing the most suited procedure to a given unit being assessed and to a given policy objective (*e.g.* internal *vs.* external process, paper and pencil *vs.* computer-based evaluation). Another relevant issue is the role of technology in improving evaluation and assessment procedures. The Review should catalogue the techniques and discuss their pros and cons depending on the circumstances in which assessment is performed.

38. The analysis will be organised around the following questions:

- How should specific types of evaluation such as school external assessment, teacher appraisal, or system-level evaluation be designed to achieve given purposes? (*i.e.* the mix of instruments, scope, criteria and standards).
- How can unintended effects of evaluation procedures be anticipated and minimised? What assessment procedures are prone to encourage cooperation among the different actors in the school system?
- What practical arrangements and what formats for evaluation and assessment processes are best suited to achieve assessment effectiveness at given levels of the educational system, and in the context of given policy objectives?
- What are the relative costs and benefits of evaluation and assessment procedures?

Issue 3 - Developing competencies for evaluation and for using feedback

39. The effectiveness of evaluation and assessment relies to a great extent on ensuring that both those who undertake evaluation activities as well as those who use their results are in possession of the proper skills and competencies. This is crucial to provide the necessary legitimacy to those responsible for evaluation and assessment. Since evaluation has strong stakes for the units assessed and since school outcomes heavily depend on individual relations and cooperation at the school level, successful feedback mechanisms require particular attention to developing competencies and defining responsibilities in the evaluation process.

40. In addition, competencies for using feedback to improve practice are also vital to ensure that evaluation procedures are effective. Assessment for improvement requires the inclusion of actors such as teachers in the process of school development and improvement. As a result, for instance, it is pertinent to include training for evaluation in initial teacher education alongside the development of research skills. Similarly, the preparation to become a school leader is expected to include educational leadership with some emphasis on feedback mechanisms. Particular groups such as inspectorates are also in a good position to engage in modelling and disseminating good practice in areas such as school assessment and teacher appraisal.

41. It is also critical to ensure the availability of vast information and guidelines about evaluation and assessment procedures to all involved in them.

42. The analysis will be organised around the following questions:

- How can educational authorities develop competencies in the system so evaluation and assessment are effectively used to improve school-driven student outcomes?
- What preparation and experience ensures the legitimacy of those with the responsibility to evaluate?

Issue 4 - Making the best use of evaluation results

43. There are diverging views on how results from evaluation and assessment can and should be used. Some see them primarily as tools to reveal best practices and identify shared problems in order to encourage teachers and schools to improve. Others extend their purpose to support accountability or market-mechanisms in the allocation of resources, *e.g.* by making comparative results of schools publicly available to facilitate parental choice or by having funds follow students.

44. Making the best use of evaluation results raises a number of challenges such as feeding data on performance back to those who deliver educational services, most notably teachers and school principals; establishing rewards, support systems and consequences that flow from them; and developing the channels which ensure that information generated by evaluation and assessment is used for educational policy development. These challenges highlight the importance of knowledge management in any evaluation and assessment framework.

45. It is increasingly recognised that providing parents and the wider public with information on the quality of learning outcomes is a key lever to improve educational performance. But it is also more and more accepted that evaluation and assessment tools should not merely “brand” schools but also promote development and improvement and examine learning outcomes in the context of students’ overall well-being and their socio and economic background.

46. Should all available data be disclosed to all stakeholders (in particular, parents and students) or should one fear possible adverse effects of full disclosure on cooperation and motivation within the school? Another crucial question is how to relate evaluation and assessment results to reward schemes or possible sanctions and to mechanisms for improvement such as professional development. The Review will address these questions by providing a view of the consequences of each option, relating to the characteristics of the education system at stake.

47. The analysis will be organised around the following questions:

- What are effective ways to use the results of evaluation and assessment procedures, including links with reward systems and professional development?
- How can institutional arrangements be strengthened to effect change through feedback?
- What channels better facilitate the use of evaluation and assessment results in the development of educational policy?
- How can effective accountability be articulated with a supportive and productive learning environment?

Issue 5 - Implementing evaluation and assessment policies

48. Adverse effects are particularly prone to occur when consensus has not been reached upon about the stakes of an evaluation, its importance for the performance of the school system, and the practical options to implement the mechanisms of evaluation and feedback. Consensus is all the more precious to reach since local actors may be in the best position to foresee unintended consequences or judge what is feasible in practice, and since the effectiveness of the evaluation heavily depends on their cooperation.

49. Implementation difficulties may arise as a result of a wide range of factors. There might be little experience with and tradition of evaluation or a system may be unprepared to undertake large scale evaluation as a result of the limited professional expertise of those with responsibility to evaluate. Other

obstacles may be a sense of unfairness by those being evaluated, excessive bureaucratic demands on schools, or inadequate dissemination of evaluation results by the media.

50. It is therefore important to analyse the challenges of implementation of evaluation and assessment policies. This includes reconciling the diverging interests of stakeholders, carefully analysing policy alternatives and their likely impact and discussing them with stakeholders to aim towards consensus. The analysis will also need to consider the political and legal dimensions of the implementation of assessment and evaluation strategies. For instance, a relevant issue is how to balance need for knowledge with political imperatives. Further, the analysis will need to identify some common impediments to implementation, with a view to overcome them and identify the conditions conducive to successful policy adoption. In this context it will be important to explore the role of bargaining processes as well as that of incentive structures facilitating compliance with new policies, as a way to ensure policy implementation in the longer term. Other strategies such as pilot projects before wide-scale implementation will also be analysed.

51. The analysis will be organised around the following questions:

- What are the main obstacles/difficulties for implementing evaluation and assessment strategies?
- How can the motivation of the various actors be ensured in order to facilitate cooperation and the success of an evaluation and assessment programme?
- Which policy settings matter most for the successful implementation of assessment and evaluation procedures?
- What approaches can be used to facilitate the implementation of evaluation and assessment procedures?

METHODOLOGY

Main Components of the Review

52. The Review methodology is based on countries working collaboratively with the Secretariat. It involves examining country-specific issues and policy approaches in using assessment and evaluation to improve school-driven student outcomes, and placing these experiences within a broader analytical framework to generate insights and findings relevant to OECD countries as a whole. The collaborative approach provides countries with an opportunity to learn more about themselves by examining their experiences against those of other countries. It is also intended to add to the broader knowledge base by accumulating international evidence on the impact of policy reforms, and the circumstances under which they work best.

53. Despite these potential advantages, comparative work can be problematic. The contexts within which evaluation and assessment arrangements operate can vary markedly across countries depending upon their historical traditions, social structures and economic conditions. Policy initiatives to improve the effectiveness of evaluation and assessment procedures that work well in one national context are not necessarily transferable across national borders. The Review will be sensitive to the role played by contextual factors in influencing the particular policies that are attempted in countries, and the factors that shape their implementation and impact. The questions to guide the Review will be framed so that evaluation and assessment policies can be understood in relation to the values, vision and organisation of schools in different countries as well as the broader economic, social and political contexts.

54. The work will be a combination of desk-based analysis and country reviews. This is complemented with the meetings of the Group of National Experts to discuss progress and share experiences. The work will be organised in three phases:

- An *Analytic Phase*, to draw together evidence-based policy lessons from international data, research and analysis.
- A *Country Review Phase*, to provide policy advice to individual countries tailored to the issues of interest in those countries, on the basis of the international evidence base, combined with evidence obtained by a team of experts visiting the country.
- A *Synthesis Phase*, with the preparation of a final synthesis report to blend analytic and review evidence and provide overall policy conclusions.

Analytic Phase

55. The Analytical Phase uses several means – country background reports, literature reviews and data analyses – to analyse the factors that shape evaluation and assessment in school systems and possible policy responses.

Part 1: Identifying the key questions for analysis and the background information needed from countries

56. A first step in the Review is to identify the key questions for analysis. Recent experience with conducting policy reviews has demonstrated the value of reaching consensus on the key policy question(s) to focus the analysis and keep it on track. An earlier version of this document gave countries the opportunity to provide feedback on the proposed issues for analysis thereby assisting with the shaping of the key questions around which the Review will be focussed. The present version of the document incorporates such feedback and provides the agreed questions for analysis. This allows the Secretariat to

identify the information it needs to collect from countries on their approaches to evaluation and assessment, the challenges of implementation and the evidence on their impact.

Part 2: Reviewing the literature and evidence on the impact of evaluation and assessment procedures

57. The Review will take stock of the existing knowledge base within the OECD and member countries as well as academic research on the relationship between assessment and evaluation procedures and performance of students, teachers and schools. It will look at the quantitative and qualitative evidence available on the different approaches used to evaluate and assess educational practice and performance. This work will include a typology of the existing evaluation and assessment frameworks in educational systems across the OECD. The Secretariat will review the literature and evidence during the course of 2009 and 2010, present it to countries and summarise it in working papers. It will also refine its policy analysis tools to use in analysing individual country practices.

Part 3: Gathering data on countries' policies and practices

58. There are important differences between countries with respect to the extent to which the spectrum of evaluation methods are developed and the extent to which actors at different levels are equipped and inclined to use evaluation results. The Secretariat will collect more detailed information on evaluation and assessment arrangements in countries, which will take into account the key questions for analysis. The gathering of data will also draw on and entail collaboration with the Indicators of National Education Systems (INES), PISA and TALIS to better understand the different assessment frameworks.

59. Every effort will be made to collect this information from as many member and observer countries as possible, in ways that minimise the burden on countries and make it easy to share and add to the public knowledge base. The principal objective is to take stock of current policies and practices in countries and summarise them either as a standalone output or as part of the final synthesis report.

60. Information on countries' policies and practices will be gathered through Country Background Reports (CBRs). These are prepared following guidelines prepared by the Secretariat in response to a common set of issues and questions, and use a common framework to facilitate comparative analysis and maximise the opportunities for countries to learn from each other. The Guidelines for the preparation of the CBR are set out in a dedicated document [EDU/EDPC/EA(2009)1/REV1]. If deemed necessary, this information might be complemented at a later stage of the Review by a questionnaire to collect information on features of evaluation and assessment arrangements in a tabular format proposed by the OECD Secretariat.

Country Review Phase

61. Country reviews both provide value to the individual country and important input into the comparative work. Countries are invited to request a Country Review, which involves an OECD-led review team providing from an international perspective, an analysis of evaluation and assessment policies in the country, and recommendations for policy development and implementation.

62. The scope and focus of each review will be determined by the country in consultation with the Secretariat, depending on the present arrangements and their strengths and challenges, and country priorities (e.g. school assessment, teacher appraisal). By providing an external perspective on evaluation and assessment issues, the Country Reviews are also intended to contribute to national discussions, as well as inform other countries about policy innovations underway.

63. Each country review will include an intensive country visit (typically 4-6 days in length, but possibly longer to account for country's circumstances) to fully understand the country's context, policies

and practices and will include meetings with all major stakeholders. Each country opting for a Country Review is expected to appoint a National Co-ordinator to take responsibility for co-ordinating the review with the OECD. The role and typical profile of a national co-ordinator is outlined in Annex 3. For each country review the Secretariat will deliver a short draft report to the country for comment four months after the review visit has taken place. Country comments are taken into account before the report is finalised and published.

64. In some countries, education policy is, in part or in whole, a regional/provincial/state responsibility. Subject to the approval of national authorities, the Secretariat may carry out a “country review” at a sub-national level rather than the whole country.

65. To ensure quality and delivery of the country review reports, it is necessary to limit the number of country reviews that will be undertaken before moving to the synthesis phase. Given Secretariat capacity constraints, it is likely that only seven country reviews can be undertaken in 2010. If more than seven countries seek country reviews, the Secretariat would ask the Education Policy Committee to consider allowing a second round of countries in 2011-12.

66. To be able to carry out this phase of the work, including assigning appropriately qualified Secretariat staff to the project, it is necessary to establish at an early stage in the project how many countries wish to have such a review. Countries are asked to formally confirm their interest to the Secretariat no later than **October 31, 2009**.

Synthesis Phase

67. This comparative phase involves the Secretariat using the earlier steps to analyse policy options and highlight good practices across countries. This could include the following points:

- Analyse the merits and drawbacks of different approaches and strategies that countries use to assess and evaluate educational performance and practice.
- Analyse how countries can effectively use the results of evaluations to enhance student outcomes through improved school practices.
- Explore policy options available to countries to improve the effectiveness of evaluation and assessment procedures.
- Explore strategies that facilitate the successful implementation of evaluation and assessment strategies.

68. This phase will be carried out during 2010 with feedback by countries. The findings will be published in a short synthesis report.

Collaboration

69. A variety of other OECD activities also address different aspects of evaluation and assessment issues. Therefore the Review will work in close collaboration with these activities, which include the regular work of INES, PISA and TALIS, country-specific projects on policy implementation, work by CELE on evaluating quality in educational facilities, work by CERI on formative assessment, innovative learning environments, market mechanisms, linking research and policy, and systemic innovation in education and policy outputs addressing early childhood education, equity in education and migrant education.

70. The growing attention being paid to evaluation and assessment issues is also evident in the work of other international organisations. The OECD Review will therefore be conducted in close co-operation

with other international organisations to prevent duplication of efforts and develop potential synergies. Social partners will also be involved namely through the contribution of TUAC (Trade Union Advisory Committee to the OECD) and BIAC (Business and Industry Advisory Committee to the OECD).

Outputs

71. The project will produce a number of outputs designed to assist policy making, including:

- A stock-take of current policies and practices in countries;
- Short country reports presenting the results of the country reviews and proposing priority actions for improving the country's evaluation and assessment framework;
- Meetings of national experts to review progress and share experiences;
- A typology of approaches to evaluation and assessment in school systems in as many OECD and observer countries as possible;
- A final synthesis report that will draw out the key lessons for policy makers and policy options available to countries to improve the effectiveness of evaluation and assessment frameworks; and
- 'Spotlight reports', customised summary of the findings of the comparative work that puts a specific country in the spotlight, for those countries which provide a voluntary contribution for the comparative analysis.

Dissemination

72. The Review will have a strong emphasis on dissemination from the outset. The outputs will be disseminated to a wide range of stakeholders – to the Education Policy Committee through OLIS and to the wider general public through the OECD public website. National seminars to discuss country review reports may also take place. A more detailed communication strategy for the full set of outputs will be developed in due course.

Participation and Costs

73. The Review is designed to be financed by a mix of general OECD resources and voluntary contributions. An illustrative budget, based on assumptions about the level of country participation, is set out in Annex 4. The comparative analysis set out in this project plan (project development and analytical phase, synthesis phase and dissemination phase) is estimated to cost 720 K EUR. Of this, 420 K EUR is covered from the Organisation's general budget and the remainder would need to be covered by countries providing voluntary contributions to support this work. To this end, countries are encouraged to support this work through a voluntary contribution of 20 000 EUR each or *in-kind* contributions. The comparative part of this project plan would be covered by 15 countries contributing this amount. Countries providing this voluntary contribution will receive a 'spotlight report', a customised summary of the findings of the comparative work that puts their country in the spotlight. The Secretariat would revise the budget and modify the project plan and outputs to take account of the voluntary contributions actually received.

74. Country reviews are fully funded by the countries concerned. Each country review is estimated to cost an additional 78 K EUR per country, and assumes that there will be a short preparatory mission as well as the policy mission. Countries would also meet the costs of preparing their Country Background Report as well as the costs of participating in meetings of the Group of National Experts. For those countries opting for a Country Review, internal travel costs of the review team as well as potential interpretation costs would need to be covered separately by the country.

REFERENCES

- OECD (2008a), How do OECD Countries Take Stock of Progress and Performance in Education Systems? Evidence and Issues, Paper for discussion at the Education Policy Committee Ad Hoc Workshop “Taking Stock of Education Performance: From Student Testing to System Evaluation” 21-22 October [EDU/EDPC(2008)33], OECD, Paris.
- OECD (2008b), Report on the Outcomes of the Education Policy Committee Ad Hoc Workshop “Taking Stock of Education Performance: From Student Testing to System Evaluation” 21-22 October [EDU/EDPC(2008)46], OECD, Paris.
- OECD (2008c), Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools, OECD, Paris.
- Wößmann, Ludger, Elke Lüdemann, Gabriela Schütz, Martin R. West (2007a), School Accountability, Autonomy, choice and the Level of Student Achievement: International Evidence from PISA 2003, OECD Education Working Paper No. 13. EDU/WKP(2007)8/REV1, OECD, Paris.
- Wößmann, Ludger, Elke Lüdemann, Gabriela Schütz, Martin R. West (2007b), *School Accountability, Autonomy, choice and the Equity of Student Achievement: International Evidence from PISA 2003*, OECD Education Working Paper No. 14. EDU/WKP(2007)9/REV1, OECD, Paris.

**ANNEX 1: GROUP OF NATIONAL EXPERTS ON EVALUATION AND ASSESSMENT
DRAFT MANDATE**

The Group of National Experts on Evaluation and Assessment is a subsidiary body of the Education Policy Committee and is responsible for overseeing work on the policy output "Evaluation and Assessment Frameworks for Improving School Outcomes". In particular it will:

- guide the methods, timing and principles of the policy output "Evaluation and Assessment Frameworks for Improving School Outcomes"; and
- allow OECD countries to share information and experience on this issue and to keep in touch with the emerging findings of the exercise.

Subject to approval, this mandate shall enter into force immediately and expire upon completion of the policy output "Evaluation and Assessment for Improving School Outcomes" on 31/12/2010 unless carried over into the next biennium. The membership and observership of the Group of National Experts will be the same as those of the Education Policy Committee.

ANNEX 2: PROVISIONAL TIMELINE FOR THE REVIEW⁴

October-December 2009

- The first meeting of the Group of National Experts on Evaluation and Assessment convened on 1-2 October, 2009, at OECD Headquarters in Paris.
- The Secretariat continues its analytical work on the impact of evaluation and assessment procedures and initiates the organisation of country visits.
- Countries commence work on their country background reports following Secretariat guidelines.

January-March 2010

- Some countries deliver their country background report to the Secretariat.
- Country visits begin. The schedule of the visits will be determined in consultation with countries.

April-June 2010

- Country visits for countries opting for a Country Review continue. Review teams prepare Country Review reports based on their site visits.
- Remaining countries deliver their country background report to the Secretariat.
- The second meeting of the Group of National Experts on Evaluation and Assessment discusses policy issues based on evidence collected and country experiences on specific issues.

July-September 2010

- Work on the final synthesis report is initiated.
- The Secretariat delivers the country review reports.

October-December 2010

- Work on the final synthesis report continues. Dissemination activities are planned.
- The third meeting of the Group of National Experts on Evaluation and Assessment discusses the preliminary synthesis report and country experiences on key themes of the report.

Early 2011

- The final synthesis report is published as an OECD official publication and dissemination activities are launched.

⁴ The timeline might need to be adjusted to account for the circumstances faced by the Review.

ANNEX 3: NATIONAL CO-ORDINATOR FOR COUNTRY REVIEWS

OECD experience with country reviews has shown that the national coordinator plays a critical role in ensuring the success of the review. For this reason, this annex sets out the role of the national co-ordinator and a typical profile of a national coordinator to assist countries in selecting an appropriate person.

Role

The national co-ordinator would be responsible for: communications with the OECD Secretariat about the Review; communications within the country about the Review; ensuring that country background information is provided to the OECD Secretariat Report on schedule; liaising with the OECD Secretariat about the organisation of the review team visit; attending meetings of the group of national experts; co-ordinating country feedback on draft materials; and assisting with dissemination activities.

Typical profile

The national co-ordinator should have some responsibility for policies on evaluation and assessment within either education ministries or agencies for evaluation and assessment in school systems. The national co-ordinator is also expected to have an in-depth knowledge of the national evaluation and assessment framework and be in a good position to co-ordinate the work across the different agencies with responsibilities for evaluation and assessment. Preferably the national co-ordinator should have a good knowledge of international developments in the area of evaluation and assessment and some familiarity with OECD work.

ANNEX 4: ILLUSTRATIVE BUDGET⁵

	2009	2010	total
Income			
Part I budget	64	356	420
VCs for comparative work 15 countries contributing 20 K EUR each			300
VCs for country reviews 8 countries contributing 78 K EUR each			624
total income			1,344
Expenditure			
Comparative work			
staff costs	210	370	
meetings and translation	36	72	
overhead	11	20	
	257	462	719
Country reviews (assuming 8 countries)			
staff costs		411	
experts		96	
missions		88	
overhead		27	
		622	622
total expenditures			1,341

5. A more detailed breakdown is available from the Secretariat on request.