The INES Compendium

Contributions from the INES Networks and Working Groups

11-13 September 2000
Tokyo, Japan
FOREWORD

1. The fourth General Assembly of the OECD Education Indicators programme is being held from September 11-13 2000 in Tokyo. It will, as on previous occasions, determine the objectives for OECD’s quantitative work in the field of education and training over the next five year mandate 2001-2006. It is expected that this General Assembly will provide the framework for INES work to progress from the identification and measurement of knowledge and skills that are crucial for tomorrow’s economies and societies, through the estimation of the economic and social returns to these skills, towards a better understanding of what the educational, economic and social factors are that contribute to skill development. The General Assembly will also provide the occasion to review important analytical perspectives that INES work could take in its next phase, in particular quantifying progress towards lifelong learning and addressing issues of equity in educational and learning opportunities and outcomes.

2. In 1999 the INES Networks and working groups were invited to consider how they might contribute to the development of a volume, intended as background documentation for participants to the General Assembly. The stated purpose of the volume (« the compendium ») was to synthesise the state of play of developmental work in the INES Networks and working groups, in areas central to future work in INES. It was to describe the current state of international statistics and indicators in the principal domains of interest in the sphere of education, including the progress achieved over the recent past, and to identify the remaining gaps where efforts might usefully be directed. The need for policy relevance was highlighted as a continuing guidepost in this exercise.

3. This volume represents a selection of the outcomes of these efforts. It represents the considered views of networks and working groups regarding achievements over the current mandate, in particular with respect to student outcomes, the measurement of skills and cross-curricular competencies, labour force outcomes of education, the school-to-work transition, indicators of continuing education and training, teachers and the conditions of teaching, decision-making, process indicators on the functioning of schools, and conceptual underpinnings for equity and lifelong learning indicators. The compendium, however, constitutes not just a stock-taking, but also a look to the future. If the progress achieved over the current mandate has been significant (see the annex to this foreword for an inventory of INES activities and outcomes since the beginning of the mandate), much still remains to be done.

4. Many of the chapters presented here present some proposals and preliminary directions for future work. More may come out of the discussions at the General Assembly itself. Current and proposed new activities can usefully be viewed in the context of the main themes of the General Assembly, namely:

   – Measuring the quality of education outcomes
   – Estimating economic and social returns to learning
   – Identifying key determinants of educational success
   – Applying cross-cutting policy perspectives in analytical work.

5. Each theme constitutes one element in an implicit cause-and-effect model of education that includes upstream determinants of quality and downstream impacts of outcomes, the whole against a backdrop of a lifetime view of learning and of a concern for equity. What makes the exercise a difficult
one is that effects, whether they be on educational outcomes or as result of these, tend to be multiply
determined, so that it may be difficult to determine precisely what contributes to what and to what degree.

6. The breadth of ideas and proposals for future work will certainly outweigh the capacity of
member countries and the OECD Secretariat to pursue them all. Comparability of indicators may be weak,
coverage imperfect, measurement error significant. Moreover, the investment needed in some areas to
bring about improvements may be out of proportion to the expected results. Whether it is desirable to
pursue current work or whether resources should to transferred to new areas or areas in which progress has
been limited are questions that need to be considered seriously, in the light of national policy needs. There
will be need to focus attention on areas which are likely to yield dividends and which involve levers over
which policy has some control. If this compendium of INES work helps to clarify what the possible
options are in this regard, it will have fulfilled its intended role.
SUMMARY OF MAIN INES ACTIVITIES AND PRODUCTS SINCE THE THIRD INES GENERAL ASSEMBLY IN 1995

November 1995  The INES Technical Group finalised the UNESCO/OECD/EUROSTAT (UOE) data collection on education statistics which covers access to participation, progression and completion of education, educational finance and education personnel. Since 1996, this instrument has been administered annually in electronic form, replacing various paper—and-pencil data collections that the three organisations had previously administered independently.

December 1995  Publication of *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*.

October 1996  The OECD Education Committee and CERI Governing Board established the INES Steering Group to provide broad policy directions for the OECD’s statistical activities in education and related fields.


November 1996  Completion of the INES Network A survey on educational goals.

April 1997  Adoption of the Network A data strategy for the development of information on student achievement on a regular basis by the Education Committee and CERI Governing Board.

May 1997  Completion of the taxonomy survey of educational programmes by the INES Technical Group which provided the basis for the revision of ISCED.

June 1997  The INES Steering Group established a programme of work for developmental work in the priority domains identified at the 3rd General Assembly in Lahti. As part of this, three new ad-hoc working groups were created: one on lifelong learning, a second on the definition and selection of skills (DeSeCo), and a third on disparities and equity.

June 1997  Launch of the Network B pilot survey on the transition from school to work.

June 1997  The INES Steering Group suspended Network D on attitudes and expectations.

September 1997  Launch of the OECD/UNESCO World Education Indicators programme with 11 countries (Argentina, Brazil, Chile, China, India, Indonesia, Jordan, Malaysia, Pakistan, the Philippines and the Russian Federation).

October 1997  Launch of the OECD Programme for International Student Assessment (PISA).

November 1997  Launch of the Network B data collection on fiscal and private returns to education.

November 1997  Publication of *Literacy Skills for the Knowledge Society: Further Results from the International Adult Literacy Survey*.
November 1997  Publication of the 5\textsuperscript{th} edition of *Education at a Glance* and the 2\textsuperscript{nd} edition of *Education Policy Analysis*.

April 1998  Publication of the 3\textsuperscript{rd} edition of *Education Policy Analysis*.

April 1998  Publication of *Human Capital Investment – An International Comparison*, drawing on Network B indicators on labour market outcomes and returns to education and training.*

November 1998  Publication of the 6\textsuperscript{th} edition of *Education at a Glance*.

December 1998  Report by the INES ad-hoc working group on lifelong learning.

January 1999  Incorporation of the Network B core database on labour market outcomes into the OECD education database. Implementation of ISCED-97 in the Network B surveys.

May 1999  Publication of the *Manual for the Implementation of ISCED-97 in OECD Member countries*. This manual paved the way for considerable improvements in the comparability of the OECD education indicators. Subsequent negotiations between countries, through the Technical Group, led to a better understanding of the similarities and differences of education systems and to an improved and more comparable allocation of national programmes to ISCED-97. The 2000 edition of *Education at a Glance* was the first publication based on ISCED-97.

May 1999  Publication *Measuring Student Knowledge and Skills - A New Framework for Assessment* that presents the conceptual framework for the PISA instruments.

July 1999  Meeting of experts on the development of comparable measures of adult education and training.

August 1999  Public release of the OECD education database on CD-ROM.

September 1999  DeSeCo symposium on the definition and selection of skills.

October 1999  Publication of the 4\textsuperscript{th} edition of *Education Policy Analysis*.

November 1999  Launch of the Second Expenditure Comparability Study.

2000  Data collection for the first PISA survey cycle.

February 2000  Publication of *Investing in Education - Analysis of the World Education Indicators*, the first publication resulting from the OECD/UNESCO World Education Indicators programme.

March 2000  Information meeting of Education Ministers on the education indicators in Copenhagen.

June 2000  Publication of *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*.*

April 2000  Publication of *Measuring Student Knowledge and Skills - the PISA 2000 assessment* that illustrates the assessment instruments of the first PISA survey cycle.

May 2000  Publication of the 7\textsuperscript{th} edition of *Education at a Glance*, based on the revised ISCED. Dissemination supported by a new strategy that involved country press briefings.

June 2000  International workshop on comparative data on education to work transitions.

July 2000  Launch of the Network C survey on upper secondary schools.
August 2000  Launch of the second PISA survey cycle.

September 2000  Public release of the 2nd edition of the OECD education database on CD-ROM and through the OECD on-line dissemination system.

*Note:  The three reports marked by * were not prepared formally under the aegis of INES. However, it forms an important part of OECD’s work on statistics and indicators in education and training and provided material for inclusion in EAG.
# TABLE OF CONTENTS

## FOREWORD

<table>
<thead>
<tr>
<th>Foreword</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
</tr>
</tbody>
</table>

## SECTION I - CONTRIBUTIONS FROM NETWORK A AND DESECO

<table>
<thead>
<tr>
<th>Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>From “Horse Race” to Educational Improvement: The Future of International Education Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
</tr>
<tr>
<td>BY Eugene Owen, Maria Stephens, Jay Moskowitz, Guillermo Gil</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cross-curricular Competencies in PISA Towards a Framework for Assessing Problem-Solving Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
</tr>
<tr>
<td>BY John Dossey, Beno Csapo, Ton de Jong, Eckhard Klieme, Stella Vosniadou</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>From Comprehension to Literacy: Thirty Years of Reading Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
</tr>
<tr>
<td>BY Dominique Lafontaine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Definition and Selection of Key Competencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
</tr>
<tr>
<td>BY Dominique Simone Rychen and Laura Hersh Salganik</td>
</tr>
</tbody>
</table>

## SECTION II - CONTRIBUTIONS FROM NETWORK B

<table>
<thead>
<tr>
<th>Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Labour Market Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>79</td>
</tr>
<tr>
<td>BY Allan Nordin</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indicators on Continuing Education and Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>89</td>
</tr>
<tr>
<td>BY Anna Borkowsky</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Educational Determinants and Other Aspects of the Transition Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
</tr>
<tr>
<td>BY Patrice de Broucker, Michel-Henri Gensbittel, Chirstine Mainguet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New Indicators for Economic Returns to Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>117</td>
</tr>
<tr>
<td>BY Zhongren Jin</td>
</tr>
</tbody>
</table>
SECTION III - CONTRIBUTIONS FROM NETWORK C .................................................. 137

INTRODUCTION
THE LEARNING ENVIRONMENT AND ITS RELATIONSHIP TO OUTCOMES .................139
BY JAAP SCHEERENS

MEASURING PROCESS INDICATORS ON SCHOOL FUNCTIONING
BY MEANS OF SURVEYS .............................................................................................................151
BY JAAP SCHEERENS, CARMO CLIMACO, BODHILD BAASLAND

INDICATORS ON TEACHERS AND THE CONDITIONS OF TEACHING ....................173
BY MARIA A. HENDRIKS AND THOMAS M. SMITH

MEASURING FUNCTIONAL (DE)CENTRALIZATION .........................................................191
BY JOEL SHERMAN, JAAP SCHEERENS

SCHOOL TO WORK TRANSITION .......................................................................................212
BY JITTIE BRANDSMA

SECTION IV - CONTRIBUTION FROM THE TECHNICAL GROUP ............................ 217

INTRODUCTION................................................................................................................... ..........219

TOWARDS A STATISTICAL FRAMEWORK FOR MONITORING PROGRESS
TOWARDS LIFELONG LEARNING .........................................................................................221
BY WALTER HÖRNER

SECTION V - CONTRIBUTION FROM THE AD-HOC GROUP ON EQUITY .............. 245

TOWARDS A SYSTEM OF EQUALITY AND EQUITY INDICATORS...............................247
BY WALO HUTMACHER
SECTION I

CONTRIBUTIONS FROM NETWORK A AND DESECO
INTRODUCTION

1. At the 1995 INES General Assembly, Member countries encouraged the OECD to develop international benchmarks that would allow policy makers to assess the extent to which students approaching the end of compulsory education have acquired knowledge and skills that are essential for full participation in society.

2. In response to this, Network A began to develop a data strategy that would provide regular, policy-driven data on student achievement in a range of relevant subject matter areas and to explore methods of assessing students that would be strong at measuring relevant skills, provide a predictable tool for monitoring educational progress over time, and ensure that comparisons are valid across countries and cultures. The implementation of the data strategy commenced in 1997 and the data collection for the first PISA survey cycle was successfully completed in 2000.

3. With this accomplished, Network A has turned its attention towards the further development of the data strategy and to supporting the ongoing analytical work in INES. In particular, the Network seeks to:

   – Contribute to the long-term development and evaluation of the PISA data strategy and, in particular, to ensure the overall coherence of the data strategy over time and its adaptation to changing needs and conceptual and methodological innovations in the field of student assessment;
   
   – Develop measures for the assessment of cross-curriculum competencies and to contribute to their integration into the data strategy;
   
   – Continue with the development of new types of indicators, particularly in the area of equity of educational outcomes, and develop and maintain high standards of technical quality in the preparation of indicators for Education at a Glance; and
   
   – Identify improved ways of analysing and presenting outcome data so that the collected information will become most useful and relevant to education policy development.

4. Three types of activities are currently underway:

   – First, the Network strives to extend the coverage of PISA subjects. A survey of educational goals that the Network in 1996 revealed the importance which education systems place on ensuring that individuals not only acquire knowledge in varied fields, but that they also develop abilities with which to evaluate, compare, and critique; abilities with which to imagine, hypothesise, discover and invent; and abilities with which to put into practice knowledge and skills. The development of a self-assessment on self-regulated learning, that captured aspects such as motivation, goal orientation, effort and persistence as well as the social context of learning was a first step to widen the scope of PISA. The development of a framework for the assessment of problem-solving skills is currently another important activity in the Network. Other domains, such as the assessment of information technology literacy and communication skills are longer-term goals which the Network began to explore in 1999.

   – Second, the Network started to explore the assessment of target populations beyond the end of compulsory schooling that currently forms the focus of PISA. A particularly important question
that Member countries need to answer is whether and how a survey approach to the assessment of adult skills should be undertaken in the framework of INES.

Third, the Network is seeking to improve the linkages between performance measures, on the one hand, and learning processes and other factors considered relevant for educational success, on the other. As a first step into this direction, the Network is working on the development of a framework for analysis that will help guiding the development of the context questionnaires for future PISA survey cycles.

5. In preparation for the INES General Assembly, Network A prepared a series of issue and background papers, addressing the future development of the PISA data strategy; the development of measures for the assessment of cross-curricular competencies; the development of methodologies for the analysis of complex, hierarchically structured data; and ways for improving the comparability and cross-cultural validity of achievement measures.

6. Three of these papers have been selected for publication in the INES Compendium to illustrate the work of Network A.

– Chapter 1 reviews progress in the development of international achievement surveys during the 1990’s and identifies key challenges that such surveys will face over the years to come.

– Chapter 2 outlines Network A’s framework for the assessment of problem-solving skills as part of the second PISA survey cycle.

– Chapter 3 reviews the steps which successive international studies of reading literacy have taken, from the IEA six subject study in the early 1970’s until PISA, in their strive for ensuring cross-cultural relevance and linguistic and psychometric equivalence in achievement measures.

7. Member countries agreed, in 1997, to focus the first PISA survey cycle on the assessment of important literacy skills, based on a new and innovative concept of literacy that encompasses reading literacy, mathematical literacy, scientific literacy and technological literacy, and that emphasises the ability of students to apply knowledge in a range of situations, backed by a broad understanding of key concepts, rather than the possession of specific subject matter knowledge. Member countries considered an assessment of these literacy skills as particularly relevant for a cross-national comparison of schooling outcomes, sufficiently founded on scientific theory, and sufficiently feasible in the field. While this pragmatic approach is widely accepted, Member countries are increasingly expressing the need to establish a more theory-oriented framework that can ensure that the widening of the scope of future assessments is guided by an understanding of the knowledge, skills and attitudes that are most relevant for societies and economies. To respond to this, the INES Steering Group launched, in 1997 the ad-hoc working group on the definition and selection of competencies (DeSeCo). This group, which is being led by Switzerland, is seeking to:

– work towards a better understanding of the competencies that are important for today’s societies and economies and the societal mechanisms through which they are defined;

– further advance the theoretical underpinning of the measurement of skills and competencies;

– develop reference points for the scientific validation of assessment indicators and for a more accurate and appropriate interpretation of empirical results;

– provide feedback for education policy by helping that relevant criteria for successful evaluation of the educational system are selected on the basis of a theoretically grounded framework.

8. An international symposium was organised by DeSeCo in September 1999 to review progress and to establish the directions for further work in this area. Chapter 4 provides a synthesis of the main issues discussed at this symposium.
FROM “HORSE RACE” TO EDUCATIONAL IMPROVEMENT: 
THE FUTURE OF INTERNATIONAL EDUCATION ASSESSMENTS

Eugene Owen, National Center for Education Statistics, United States
Maria Stephens, American Institutes for Research, United States
Jay Moskowitz, American Institutes for Research, United States
Guillermo Gil, National Institute for Assessment and Evaluation, Spain

Comparative international assessment is rising at this moment because of a growing interest among public administrators responsible for education to have information that can contribute to improving the quality of education. There are great new projects in the process of being established, broader and more stable than those in the past, and which benefit from the prior experience and worldwide infrastructure of researchers and institutions that have been created as a direct result of previous comparative educational assessments.

– Guillermo Gil, Spain

1. The past ten years have clearly demonstrated the usefulness of education assessment data for policy purposes, their importance politically, and the public’s and media’s intrigue and appetite for this information. This successful expansion of comparative assessments has been due, in part, to timing. The major political issues faced in developed countries during the past decade—the economy, education, and youth employment—were paralleled by responses from the education sector to establish higher and more clearly articulated standards, implement systemic reform, and increasingly look to education statistics and assessment to determine how students were progressing.

2. As the twenty-first century dawns, the past will serve as prelude to the future, and it is safe to assume that international comparisons of student achievement will continue well into the new millennium. Reflecting briefly on the history of international assessments in education—and most specifically, on the lessons of the past decade—this chapter will look ahead to the next decade of assessment, describing some of the opportunities afforded by methodological innovations, technology, and expanding policy interests and speculating on some of the challenges and issues that are likely to arise.

A brief look back at international education assessment

3. International assessments of educational achievement started more than three decades ago with the trailblazing work of Torsten Husen and the establishment of the International Association for the Evaluation of Educational Achievement (IEA) in 1959. Its inquiries started with a Feasibility Study that explored the possibility of conducting international assessments that would be sufficiently scientifically rigorous. While this important study demonstrated that international comparative studies were indeed feasible, it also illustrated the difficulties (e.g., methodological and organisational) such studies present.

4. Over the years, IEA implemented many international assessments of education aimed at providing data that would be useful for ascertaining the effectiveness of educational systems (see Exhibit I). Those studies focused largely on the relationship of teaching and the curriculum to student achievement. For a variety of reasons, one of the frequent approaches used for reporting the results from these studies was to rank countries using total scores, or selected sub-scores, on tests presumed to measure
student achievement in various subject areas. This use of country rankings, an outgrowth of countries’ increasing concerns about their economic development and their students’ abilities to compete in a changing economy—known as the “horse race”—became a primary characteristic of assessment, often against the wishes of the researchers conducting the studies.

5. This focus on the “horse race” took hold in the 1980s and the popularity of international education assessment continued to grow, as well. Throughout the 1980s, the IEA brought together researchers from as many as 24 countries to assess students’ performance in subjects such as reading, mathematics, science, civics, and computer literacy, and to describe those factors that contribute to differences in achievement among the countries studied.

6. However, it was not until the 1990s that the education community saw a dramatic increase in countries’ interest in international assessments of education and in the development of methodologies that further validated the reliability of results from comparative studies. In the 1990s, we had the IEA’s Third International Mathematics and Science Study (TIMSS) which was repeated in 1999 (TIMSS-R); the IEA Reading Literacy Study, conducted in 1990-91 in 32 countries; the International Adult Literacy Survey (IALS), administered in 1994 and 1996 in a total of 12 countries; and the IEA Civic Education Study (CivEd), expanding upon previous studies of civic education with the integration of both qualitative and quantitative methodologies.

7. It also was in the 1990s that we saw the first move away from the “horse race.” During this time, planning began for activities that will be implemented in the 2000s, including OECD’s Programme for International Student Assessment (PISA) and the Adult Literacy and Lifeskills (ALL) Survey, each born out of a need for more policy-oriented information about societies’ competencies for the 21st century. These planned studies are characterised by a broad conceptualisation of what knowledge and skills should be assessed and by their goals to move beyond the “horse race” to dissemination of information useful for improving education.

International education assessment in the last decade

8. Perhaps the most marked characteristic of assessment in the 1990s was the dramatic increase in the number and range of countries participating in international assessments. Forty-five countries participated in the TIMSS assessment of eighth-grade students—more than had ever participated in such studies before—with a large number of countries participating in the TIMSS assessments of fourth-grade and end-of-secondary school students, as well. What also was remarkable was that countries from all over the world and from widely ranging levels of economic development participated in TIMSS, supplementing the list of G-7 countries and other European countries that historically participated in such endeavours.

9. Assessment in the 1990s also was characterised by unparalleled advances in statistical methodology and quality control procedures. For instance, TIMSS used modern psychometric techniques such as matrix sampling and Item Response Theory (IRT), which permitted the development of more robust assessments covering a wider range of material than in previous studies and which improved the comparability of results across countries. The quality, depth, and accuracy of results from international assessment also will be enhanced by multi-variate analysis techniques, multi-level modeling based on Hierarchical Linear Modeling, and structural equation modeling with LISREL analysis. Motivated both by a desire to use these new psychometric approaches and by an even stronger need to be included in cross-national assessments and to publish comparative results, countries showed amazing willingness to adopt often “foreign” statistical methods.

10. Comparability of assessments was further validated by the enhancement of quality control procedures and the investiture of external quality control monitors in the international assessment process. Previously, lack of uniform test administration and other difficulties at the national level created problems

1 Formerly, the International Life Skills Survey (ILSS).
in obtaining quality data and sufficient response rates. Over time, with these new procedures, countries (sometimes begrudgingly) persisted in improving the national implementation and administration of international assessments. Also over the years, resource estimation for personnel, materials, and operations improved greatly, such that precise predictions of both international and national costs could be made at a project’s outset.

11. With new players and new procedures also came new forms of reporting. Prior to this last decade, results from educational studies were primarily research-focused narratives. However, the publication of the results from international assessments in OECD’s annual indicator report, Education at a Glance, brought the assessment results to a much wider audience. From 1992 to 2000, OECD drew upon studies from IEA and the International Assessment of Educational Progress (administered by the Educational Testing Service) to develop a series of policy-relevant indicators, which, over the years, included mean performance of countries (the “horse race”), distribution of achievement scores within countries, and the variation in student achievement associated with differences among students and among schools. Although the information contained in Education at a Glance was from previously published sources, Education at a Glance presented the information within a policy context and used an indicator approach to presenting the results.

12. Taking their cue from policy makers’ and educators’ desire for comparative information, the media provided extensive coverage of educational failures and successes. Each new edition of Education at a Glance, for example, garnered coverage in leading newspapers throughout the world. The media kept the story fresh by focusing attention on new aspects of results from year to year. Initially, it only reported on the ranking of countries (again, the “horse race”), but in subsequent years, it explored specific topics in greater depth. For example, the U.S. press often examined differences in education finance and in teacher qualifications and salaries among schools and systems in relation to student outcomes. The results of assessments, made public via the media, became a new and major force in the political debate and public discussion on education. The press (as well as many other groups) in some countries used the indicators to support existing education policies, while the press and others in other countries used the information to document the need for reforms.

13. During this decade, the development of an international system for education indicators supported the development of national indicators systems and the growing practice of national assessment. In fact, a survey conducted by the French Ministry of Education (1994) documented that, in the early 90s, the number of countries conducting national assessments had significantly increased, and that the number of subjects and grade levels being tested and the frequency of testing had increased, as well. For instance, Portugal and Sweden introduced national assessment programmes in this decade, and several programmes introduced or revamped in the 1980s (e.g., those in Spain and the United Kingdom) took further hold in their respective education communities. Also, many countries began compiling and publishing results from international assessments, national assessments, and other sources of education data in indicator reports relevant to their specific national contexts.

Factors contributing to the growth of international education assessments

14. A variety of factors account for the marked expansion of international assessments. The primary reason is simply that the comparative information gained from such assessments came to be of prime importance to policy makers. Aptly codified in the World Declaration on Education for All in 1990, there was a widespread recognition that information on the “actual learning” of students, or outcomes, to supplement traditionally gathered information on the inputs to education, was of vital importance. Too, the increasingly pervasive belief that an economy’s health would be determined by the capacity of its citizens to compete in a global environment underscored for many policy makers that assessments of education would have to include information that compared their students to those in other countries. Hence, international assessments received widespread support.

15. This was further heightened by the adoption in many countries of systems of accountability, or standards. The “standards movement” served as the encapsulating catch phrase for many countries’
attempts to raise the quality of education. Standards provided the narrative descriptions of what educators and policy makers thought students should be able to do, while assessments served as the tool to measure how well the standards were being achieved. International assessments were viewed as a natural extension of the accountability and standards movement in that they provided policy makers with information on how well students were progressing toward the standards set by the highest-performing students in the world.

16. Finally, the inclusion of policy makers in the planning and implementation of assessments further fed their growth. For instance, although researchers played the key role in the administration, development, and execution of TIMSS as they had in earlier IEA studies, for the first time, national education ministries in many countries also became active participants. The role and interest of national ministries is even larger in PISA, as they are the prime impetus and audience for the programme, as well as the setters of its broad policy directives through membership in a Board of Participating Countries.

Impact of international education assessment in the last decade

17. The results of international assessments of education have been put to many uses and have had positive impacts in many countries. In some countries where there is no national assessment, international assessments play an especially important role in providing the only available macro-level data. Moreover, in countries with more decentralised systems, where there is no clear definition of educational objectives or clearly established curriculum at the national level, it is even more important to have as a point of reference other countries that might have similar contexts (Husen and Tuijman, 1994; as cited in Gil, 1999). In other countries, as mentioned previously, international assessments have spawned increased attention to the development of national assessment and indicator systems.
<table>
<thead>
<tr>
<th>Study</th>
<th>Grade</th>
<th>Type</th>
<th>Year</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasibility Study</td>
<td>12</td>
<td>LS</td>
<td>1960</td>
<td>Mathematics, science, reading, comprehension, geography, and non-verbal reasoning.</td>
</tr>
<tr>
<td>First International Mathematics Study (FIMS)</td>
<td>13</td>
<td>LS, US</td>
<td>1964</td>
<td>Mathematics. Questionnaires on student interest in and attitudes toward mathematics.</td>
</tr>
<tr>
<td>Second International Mathematics Study (SIMS)</td>
<td>20</td>
<td>LS, US</td>
<td>1982-83</td>
<td>Mathematics. Stemmed from the interest to study the effects of the curricular changes that occurred from FIMS and to establish trends over time.</td>
</tr>
<tr>
<td>International Assessment of Educational Progress (IAEP 88)</td>
<td>6</td>
<td>LS</td>
<td>1988</td>
<td>Mathematics and science.</td>
</tr>
<tr>
<td>International Assessment of Educational Progress (IAEP 91)</td>
<td>9-20</td>
<td>P, LS</td>
<td>1991</td>
<td>Mathematics and science with an optional component in geography.</td>
</tr>
<tr>
<td>International Adult Literacy Survey</td>
<td>12</td>
<td>A</td>
<td>1990-91</td>
<td>Reading literacy. Student, teacher, and school questionnaires.</td>
</tr>
<tr>
<td>Study Name</td>
<td>Code</td>
<td>Type</td>
<td>Year</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>Third International Mathematics and Science Study – Repeat (TIMSS-R)</td>
<td>41</td>
<td>LS</td>
<td>1999</td>
<td>Mathematics and science. Student, teacher, school, and curriculum questionnaires. To give a trend line from TIMSS, allowing comparison of TIMSS-R 8th grade with TIMSS 4th and 8th grade.</td>
</tr>
<tr>
<td>Civic Education Project</td>
<td>13</td>
<td>LS</td>
<td>1999</td>
<td>Factual knowledge of civics and using civics related knowledge. Survey of attitudes toward institutions, issues, and actions. Student, teacher, and school questionnaires.</td>
</tr>
<tr>
<td>Adult Literacy and Lifeskills (ALL) Survey [Statistics Canada and U.S. National Center for Education Statistics]</td>
<td>(20)</td>
<td>A</td>
<td>2002</td>
<td>Prose literacy, document literacy, numeracy, and analytic reasoning in everyday situations. Questionnaire on information technology use and background questionnaire.</td>
</tr>
<tr>
<td>Progress in Reading Literacy Study (PIRLS)</td>
<td>(16)</td>
<td>P</td>
<td>2002</td>
<td>Reading for literary experience and reading to acquire and use information. Student, teacher, school, parent, and curriculum questionnaires.</td>
</tr>
</tbody>
</table>

Key to populations: P=Primary; LS=Lower Secondary; US=Upper Secondary; A=Adult

Notes: During the past three decades, the IEA also conducted a Pre-Primary Study and surveys on computer use or access. However, since these are not assessments of achievement (rather they are descriptive studies), they are not included here. Parentheses () in the second column indicate the anticipated number of countries to participate.

Source: Compiled from Basic Concepts and Current Reality of the International Comparative Evaluation of Student Outcomes (Gil, 1999) and other (Internet) sources.
18. International assessments have had an impact on the research community, as well—contributing significantly to the improvement of conceptual frameworks for the cognitive areas being tested and to the development of innovative methodologies to assess these areas. New management techniques for collaborative efforts and better dissemination mechanisms for results have been established, as international assessments have progressed. And, in a number of countries, participation in international assessments greatly promotes the training and professional development of local education researchers, while at the same time, the process allows the incorporation of contributions from experts of different backgrounds, education, and origins in the generation of a common framework.

19. The impacts of the information gained from assessments has been felt, as well. Comparative information has been used to arouse and stimulate debate among policy makers and to assist them in determining priorities. As one example, the Flemish community of Belgium drew upon the results of SIALS (an adult literacy survey) to inform the development of national goals for secondary school students.

20. By illuminating relative strengths and weaknesses, international assessments also have been useful in informing policy makers on difficult questions related to the allocation of resources. For instance, as a result of the Second IEA Mathematics Study (SIMS), New Zealand increased the number of mathematics advisors in secondary schools. Similarly, results from the fourth-grade population of TIMSS prompted officials in Sweden to allocate more resources to teacher training for primary school mathematics and science.

21. Test results also have triggered changes in curriculum. In Finland, policy makers report that international assessments have considerable influence on curriculum development; and in the Netherlands, results have spawned the establishment of study centres dedicated to particular curricular areas (e.g., the Dutch language).

22. Finally, policy makers and the education and research communities have found international assessments useful because there are many topics that cannot be studied within the confines of a single education system. Because within most systems, certain variables (e.g., classroom size, number of instructional hours) are relatively uniform, studying the influence of those variables on outcomes is difficult, if not impossible, because one cannot modify them within the system for experimental purposes. Thus, international assessments offer a “world laboratory” where one can compare the effects of multiple variables and their combinations, thereby supporting a variety of education research purposes. One such example of a research agenda within an international assessment programme is found in TIMSS, the assessments for which were complemented by a video study that examined teaching practice in three participating countries (Germany, Japan, and the United States). In the United States, results from the assessments and background questionnaires, coupled with information learned in the video study, were compiled into a toolkit for practitioners on how to interpret the results of TIMSS and how to apply its lesson to their own practice.

23. In short, in the 1990s the international assessment community witnessed a dramatic growth in participation in comparative education studies due in part to countries’ growing concerns about economic competitiveness, their adoption of standards movements, and a general shift in interest to outputs, not just inputs, of education. This growth was supported by advances in statistical methodologies and the inclusion of policy makers in planning and implementation, and was evidenced in the widespread media attention to and uses of the results to review and set education policy. We can only expect that interest in international assessments of education will remain strong and that the education community and the public will continue to pull away, as they began to do in the last decade, from reporting on the “horse race” to using comparative data to support educational improvement. The next section focuses on the opportunities for and challenges to assessment in the next century.
A look forward to international education assessment in the next decade

24. It is clear that international comparative assessment will continue well into the next decade. With PISA, governments now seek to establish an instrument that will regularly assess how far students approaching the end of compulsory schooling have acquired some of the knowledge and skills that are essential for full participation in society. The first PISA assessment has taken place in 2000 and will be followed by similar assessments every three years. Three “domains”: reading literacy, mathematical literacy and scientific literacy form the core of each assessment cycle. In addition, a wider range of cross-curricular competencies are progressively integrated into the assessment cycle as well, starting with a self-assessment on self-regulated learning in 2000 and continuing with an assessment of problem-solving skills in 2003.

25. PISA is also setting new standards in the development and management of international comparative assessment. It brings together scientific expertise from participating countries and is steered jointly by their governments, through the OECD, on the basis of shared, policy-driven interests. Countries are working together to produce methods of assessing students that are valid across countries, that are strong at measuring relevant skills, and that are based on authentic life situations.

26. Within the IEA, which led the repeat of TIMSS just prior to the turn of the century, countries already are discussing a second repeat of this landmark study for the year 2003.

27. Meanwhile, ALLS, CivEd, and PIRLS (IEA’s Progress in Reading Literacy, focusing on 9 year-old students) continue development work, with assessments expected early in the 2000s. While some future developments are fairly easy to speculate upon (as plans are already underway for many activities), others are not as apparent. This section of the chapter describes some of the issues—both exciting and challenging—that the assessment community is likely to encounter in the coming years.

Probable developments in international education assessments in the next decade

28. In the nineties, education reform literature emphasised the setting of high standards, and assessments at both the national and international levels gave us an idea of what benchmarks might be set in the future. An important component of the standards-setting literature, however, was the emphasis on standards for all students. And, as the context for assessment in the 1990s was standards, it is likely that the context for the first decade of the 2000s will be the important goal of equity in education. One possibility for the future is that the education systems will use the information gained from national and international assessments to set performance targets for their students and monitor that all students are progressing towards those targets.

29. It also is foreseeable that future studies will utilise more sophisticated tests that, to the extent that it is financially feasible, take advantage of multiple item types and formats (e.g., constructed response), as we are beginning to see with the reading literacy studies and PISA.

30. The next decade also will likely be the one during which the assessment community tackles important questions—already surfacing—related to the role of technology. Many countries have expressed an interest in utilising the computer as a tool for the delivery of assessment, and in the next years, the assessment community will have to explore the feasibility of doing so in a large-scale, multinational setting with a limited budget. In PIRLS, developers are proposing that one of the test booklets be administered via the computer. One of the most exciting opportunities afforded by technology as a delivery mechanism, which countries are most interested in, is when the technology is an integral part of the design and functioning of the assessment. For instance, the use of computer simulations that are adaptive and interactive—and presumably more authentic and able to gather more precise and meaningful information—is likely to be an area of growing interest in the future. Already from the PISA field trial of a computer-delivered assessment of students’ problem solving abilities in Germany, interesting results are
emerging that indicate that the use of the computer allows the skills that are being tested to be isolated from general intelligence, in a way not found in similar paper-and-pencil examinations.

31. Still another issue related to technology is how students’ skills with technology may be assessed in the future. Information communication technology (ICT) is fast becoming a key component of many curricula, but there is little information about appropriate methods for assessing students’ proficiency with ICT or on assessing ICT’s impacts on their performance overall. In fact, the countries participating in PISA have requested that assessment of ICT skills be an area for exploration and development for the third assessment cycle in 2006, and a group of INES Network A members are exploring this possibility. The mere existence of technology may cause a re-thinking or widening of the skills that are deemed important to measure: for instance, will policy makers want to know more about students’ capacity to extract information from vast sources, their comprehension of computerised presentations of text and their ability to make such presentations, or their judgement on validity of sources of information?

32. Co-ordination among language-, regional, or cultural groups is another trend we are likely to see expanded in the coming years. Already in PISA, several German-speaking countries are co-ordinating their translation efforts and working together to produce a German-language version of the assessment, which then will be adapted for national variations within different German-speaking countries or communities. It also is likely that we will see an increase in regional analyses, or comparisons among countries with similar cultural backgrounds or social contexts. We have seen a first start at this in recent years as several Latin American countries have begun working together to develop a regional indicators system (MERCOSUR). The OECD’s World Education Indicators (WEI) project is assisting countries (linked with one another by their level of economic development) to participate in and to develop a system for education indicators. At the initial stages of inquiry, the Organisation of Ibero-American States, the Arab Gulf Council, and the Organisation of Southeast Asian Ministers of Education each have shown interest in the field of assessment, initiating studies to compare education systems and promoting increased capacity for research and evaluation. Finally, the participation of Brazil, China and the Russian Federation in PISA is showing an impact in their respective regions in subsequent cycles, with more than 20 non-OECD countries applying for participation in the second PISA survey cycle.

33. It also is likely that we will continue to see a rise in targeted research efforts to supplement international assessments. The widespread interest in and utility of the first TIMSS video study and the planned second video study (with an additional 5 countries participating) are indicators that the more in-depth, qualitative information that can be gained through ethnographies, video studies, and case studies will be an important component of international assessments in the future. Such efforts can provide detailed information about teaching practises, quality of life in schools, school leadership, programme offerings, parent and student levels of satisfaction, and levels of parental involvement that cannot adequately be measured with context questionnaires.

34. National options in international assessment are another opportunity for supplementary activities that are likely to grow in importance in the next decade. For example, the problem solving assessment in Germany, referred to earlier in the chapter, is being implemented as a national option in PISA. In Canada, national implementation will over-sample students to allow for regional breakdowns and trend comparison in future cycles. Several countries (e.g., Austria and Ireland) may incorporate a reading speed component, in order to study possible effects of the amount of text in the assessment on results. As a group, PISA countries also have shown interest in a longitudinal component for future cycles of PISA, and in the next cycle, it is expected that countries will have the option of oversampling the population to allow for the future administration of a survey on the post-secondary activities of students. In the near future, countries may very well want to consider the integration of qualitative methodologies or the addition of national options—such as some (notably Canada and Germany) already are doing with PISA—to maximise the utility of international assessments for their own specific contexts.
Issues and challenges for the next decade

35. Although there are many foreseeable opportunities in the next decade of assessment, the assessment community also should be prepared for several issues and challenges that may have to be addressed.

Answering the "Limits of Testing" proponents.

36. Almost since their inception, international assessments of educational achievement have been challenged on a variety of grounds. One particular challenge has been the implicit assumption that the student achievement scores reported in such studies can in fact be attributed to their educational experiences, and thus represent a valid assessment of the comparative effectiveness of national educational systems. Critics of international assessments contend that the very nature (i.e., collaborative) of international assessments produce results that are irrelevant to national interests. This vocal minority challenges, for example, the comparability of student populations being tested and the relevance of the frameworks used to guide test development, as well as whether the tests are measuring the domains they claim to measure and the degree to which cultural differences (e.g., students’ motivation for or comfort with the test) may bias results. Moreover, there is uncertainty in this group about the extent to which a student’s (or a country's) performance can be generalised beyond the items on the test. In some cases, critics call for the development of new methods of assessment; in other cases, the need for international assessment is questioned more broadly.

37. As we enter the next century, a critical mass of developed countries will have been participating in international education assessments for almost two decades and in national assessments of education, as well. The level of financial investment in such activities will have been massive, and critics of assessment, such as those mentioned above, will have been vocal. Assessment proponents will have to assist their governments in searching for economies of scale: in looking for ways to harmonise their international assessment activities with their national assessment activities. In the next decade, countries will have to answer questions about the utility of, sometimes divergent, international and national assessment programmes. They will have to determine the optimal or necessary linkages among assessments (e.g., between subject matters or ages or grade levels), and they also will have to determine how to usefully present and learn from the stories that emerge from these endeavours.

38. For instance, by the end of this decade, we will have trend data and results from several cycles of several international assessments. Because of this, it may be a challenge to convince governments of the continued need and utility of international assessments. Also, there is the concern that we may find that, assessment after assessment, countries’ relative standings are remaining constant—forcing us to more closely examine what may account for this. This will require us to find optimal intervals for assessment and to better explain why such monitoring exercises are indeed valuable.

39. Further, where lessons are learned and policy decisions are made based on assessment results, it may take a long time for those decisions to have an impact. The reflection of those decisions in changed assessment results may take even longer. This may give a golden opportunity to sceptics of international assessment to then hold sway over the public and policy makers, questioning the need for large-scale programmes.

40. To answer them, countries will have to evaluate the purpose of their testing programmes, continually seek appropriate measurement tools, and, most likely, adapt their justifications for international assessment. For instance, a current reason for the interest in testing often is given as the need to ensure that students are able to compete in a global economy and thus contribute to the economic development of a country. Although the relationship between student achievement and indicators of economic development can appear paradoxically, and erroneously, simple—consider the below average performance of U.S. students in a strong economy versus the standard-setting performance of Japanese students in a currently weakened economy—it actually is complex and multi-layered. In this light, the
justifications for international assessment will have to leave behind the economic rhetoric that has characterised it for the past years. It will be especially important for the education and assessment communities to continue ask proactive questions that capture the public’s attention, as the economic stability questions once did.

**Grappling with ideological shifts**

41. Several important shifts related to assessment were described in the last section—namely the move away from focusing exclusively on the curriculum and the increasing orientation toward policy as well as research. As we can already see with the wide support and clear mandate for PISA, there is a shift away from assessing the curriculum to assessing broader concepts, such as literacy in mathematics or the sciences. Many in the education community support this change because they believe it will capture a more general range of what students know, which is more applicable to what students will be required to do in work and life. However, others in the community have and will continue to criticise the shift because test content may then move outside and beyond what students are actually taught in school, thereby putting an education system’s high, or even predictable, results on assessments at risk. Moreover, some may consider it difficult to cope with implementing changes to the education system based on results that are presented outside of a strict curricular framework.

42. The shift from research-oriented to policy-oriented uses of information also is likely to raise issues in the 2000s. As the results of international assessments continue to be information of public interest, the assessment community will have to expend additional effort educating the public and policy makers to understand assessment data and how it can be used. It also will require that the assessment community continues to support and look for supplementary, targeted research efforts, as described in an earlier section, to provide the type of in-depth explanatory information that policy makers and the public crave.

43. One activity that may help the public to understand the results of international assessments (and national assessments, for that matter), and something that may be requested, is public disclosure of test items and answers. As the assessment community asks the public to believe accounts of student performance (positive or negative), to make judgements between critics’ and proponents’ positions on testing, and possibly to support policy changes based on results of assessment, there will be pressure to provide information about test construction and design, as detailed as the items and scoring procedures themselves. Although public disclosure has historically been more a facet of examination versus assessment programmes, the rise in public disclosure in the former and the current tone of discussions among international assessment participants indicate that disclosure of at least some test items is important to many countries that are looking to disseminate information that will be useful at the school and classroom level. However, as assessment programmes have become more sophisticated and trend-oriented (e.g., the repetition of TIMSS and PISA’s cyclical nature), the confidentiality of items that must be used in multiple rounds of assessment—items that were developed laboriously and to fit the assessment frameworks over time—is a singular priority.

**Developing valid measures in non-traditional domains**

44. Another challenge for the assessment community will be to continue the tradition of sound measurement. As there are advances in cognitive theory and more is learned about knowledge acquisition and the development of 21st century skills, so we will look again to our psychometricians to help us find ways to measure such skills. In particular, the assessment of some domains (e.g., communication or teamwork), which appeal to policy makers, educators, and the public as important 21st century competencies, present challenges to finding appropriate, valid means of measuring them.

45. This latter concern has a great deal of relevance for the current development of measures of cross-curricular competencies (CCCs) and other domains in which assessments are being pioneered. In the ALL project, four domains are being tested, two literacy scales, numeracy, and analytic reasoning in
everyday contexts (a major component of problem solving), with an accompanying questionnaire on information technology. Two additional areas—teamwork and practical cognition—were initially explored, but it was determined that additional development work would be required before their international debuts in a household survey such as the ALL project. Test developers, for instance for the experimental domains considered for the ALL project (e.g., teamwork), will use several rounds of pilot study to find valid ways to measure them.

**Looking into and beyond the social context**

46. For years, studies of educational achievement have told us that socio-economic status (SES) is important in explaining student achievement, that low-SES children often are among the lowest performing students and high-SES children are among the highest. This leads to the perception among the public that it often is the uncontrollable variables (like SES) that explain all the differences in student performance. However, while SES is undoubtedly an important explanatory variable, it will be important to move beyond this fact in the future and explore it in more depth to find explanations that may be more policy-malleable. For instance, income-adjusted studies and multi-variate analyses will be important means for policy makers in the future for making adjustments to education systems that are meaningful to all students. Researchers can ask and answer questions such as, “How do achievement levels of low- or high-SES students compare across countries?” or, more probing, “Do some countries have greater (or lesser) variation in achievement among low- or high-SES students?”

**Summary**

47. International assessments have gained notable importance by providing policy makers, researchers, and the education community with valuable information about how well their students are performing. They offer an external and relative point of reference that contributes to the objective evaluation of education systems’ efficiency and effectiveness. Since the groundbreaking work in this field in the 1960s up to the present, we have witnessed tremendous growth in participation in these studies, in the scope of the studies, and in the technical expertise to implement and administer such massive endeavours. However, despite a long list of accomplishments, achieving comparability of populations and domains, exploring new domains that reflect the requirements of present-day society, and presenting results in such a way that scientific rigour and understanding by the general public are married continue to challenge our community as we enter the new millennium.
CROSS-CURRICULAR COMPETENCIES IN PISA TOWARDS A FRAMEWORK FOR ASSESSING PROBLEM-SOLVING SKILLS

Mr. John Dossey, Chair
Mr. Beno Csapo, University of Szeged, Hungary
Mr. Ton de Jong, University Twente, The Netherlands
Mr. Eckhard Klieme, Max Planck Institute for Human Development, Germany
Ms. Stella Vosniadou, National and Capodistrian University of Athens, Greece

Introduction

1. The OECD Programme for International Student Assessment (PISA) seeks to progressively integrate the assessment of cross-curricular competencies into the survey instrument. The assessment of problem-solving skills has been stated as a goal for the second survey cycle.

2. Problem-solving is a central educational objective within every country. Educators and policy makers are especially concerned about students’ competencies for solving problems in real-life settings. These competencies involve recognising a problem, formulating the exact nature of the problem, using this knowledge to plan a strategy for solving the problem, effectively executing the strategy, reflecting on the tentative solution, making adjustments, and communicating the solution to others. The processes of problem-solving, so conceived, are found across the curriculum, in mathematics, in the sciences, in the language arts, in the social sciences as well as in many other areas of schooling. Problem-solving provides a basis for future learning, for effectively participating in society, and for conducting personal activities.

3. The INES Network A has, through an expert committee, explored the development of a framework for the assessment of problem-solving skills. This document is a result of this work and seeks to:

   – provide an overview of the domain of problem-solving;
   – outline what an assessment of problem-solving might contain; and,
   – describe approaches PISA might take in assessing problem-solving and suggests related test specifications.
Overview of the domain of problem-solving

It seems that all cognitive activities are fundamentally problem-solving in nature. The basic argument is that human cognition is always purposeful, directed to achieving goals and to removing obstacles to those goals.

John R. Anderson (1985, p. 199)

Conceptions of problem-solving

4. As stated by John Anderson, problem-solving is an ever-present human activity. Central to the development of a framework for the assessment of problem-solving skills among 15-year-olds in the various countries participating in the PISA assessment is a clear vision of what constitutes problem-solving. Several writers have observed that there is no agreed-upon comprehensive definition of problem-solving (e.g., Frensch & Funke, 1995; O’Neil, 1998). Yet, there is a large body of literature on learning and work that discusses problem-solving, often without explicit definition of the term in context.

5. The nature of problem-solving is often described in the literature via the cognitive processes required. As the proposed PISA assessment of problem-solving is restricted to in real-world or discipline-based contexts, the following process components are of interest:

− Problem representation: This component includes searching for information, structuring it, and integrating it into a mental representation of the problem, taking into consideration the information present in the context of the situation;

− Constructing a solution: This component includes various kinds of reasoning, based on the representation, as well as the planning of actions and other solution steps; and,

− Execution and evaluation of a solution: Solution steps have to be executed and evaluated. The problem solver has to monitor and regulate his/her activities, constantly attending to the context of the problem. In dynamic environments, the problem solver must continuously process external information and feedback.

6. These problem-solving components make use of some very basic psychological mechanisms such as:

− retrieving, considering, and evaluating contextual information, as well as recalling and using general knowledge;

− applying mental tools (e.g., diagrams which may help in representing a problem or envisioning a solution) as well as using cultural tools (pencil and paper, calculator, computer, etc), and culturally provided systems of representations (language, symbol systems, etc);

− using various kinds of inductive and deductive reasoning;

− relating previous experiences and known strategies to new problem situations (e.g., by drawing analogies, by using metaphors); and,

− regulating emotional and motivational factors with cognitive factors.

7. Thus, problem-solving is the combination of many different cognitive and motivational processes that are orchestrated to achieve a certain goal that could not be reached by simply applying a well-known routine, or algorithm. Problem-solving competence is the capability to do this kind of orchestration within a certain range of tasks and situations. Problem-solving assessment aims at identifying the processes used and measuring the quality or the products of problem-solving activities.
8. In studying problem-solving, PISA’s assessments obviously have to concentrate on the range of contexts and tasks observable via large-scale problem-solving assessments. As such assessments must necessarily depend on context- or domain-specific knowledge and strategies, every measure of problem-solving competence will, to some extent, be context- and domain-specific. Therefore, the domains, contexts, and situations in which problem-solving is assessed have to be selected very carefully.

9. However, research on problem-solving in differential psychology has also shown that on the level of (latent) abilities, problem-solving competence, defined operationally by the degree to which a person can successfully solve problems, will always be close to general reasoning.

Directions in defining problem-solving

10. The existing bodies of work on problem-solving do not lend themselves to a simple taxonomy of perspectives on problem-solving. Nevertheless, the literature may be organised roughly into two sets of work: “Academic” and “Applied.” Each of these sets of work may be grouped into further categories, some of which may be likened to “schools” or “traditions.” These categories may be identified with certain researchers and writers as shown in the following table. Given the absence of any accepted classification of work on problem-solving, the table and on the next page and the following elaboration of its contents are necessarily tentative and possibly controversial, and should be viewed in that light.

<table>
<thead>
<tr>
<th>Fields of work*</th>
<th>Associated researchers/writers/organisations*#</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACADEMIC</td>
<td></td>
</tr>
<tr>
<td>Artificial intelligence</td>
<td>Newell &amp; Simon; Hinrichs</td>
</tr>
<tr>
<td>Cognitive psychology (including Complex Problem-solving in disciplines)</td>
<td>Gestalt School; Glaser; Chi; Mayer; Dörner; Funke; Frensch; Sternberg</td>
</tr>
<tr>
<td>Developmental Psychology/Epistemology</td>
<td>Piagetian; Vygotsky</td>
</tr>
<tr>
<td>Situated cognition</td>
<td>Brown, Collins and Duiguid; Scribner; Lave; Rogoff; Greeno</td>
</tr>
<tr>
<td>Applied</td>
<td></td>
</tr>
<tr>
<td>Learning in subject disciplines, e.g., math, physics, reading, writing, history</td>
<td>Schoenfeld; Chi; Stanovich, Flower; Weinberg; Klieme, Voss</td>
</tr>
<tr>
<td>Vocational education, e.g., medicine, electronics, mechanics</td>
<td>Barrows; Lesgold; Gitomer; Lurch</td>
</tr>
<tr>
<td>Generic work skills</td>
<td>Carnevale; SCANS (US); Key Qualifications (Germany); Key Competencies (Australia)</td>
</tr>
<tr>
<td>Business</td>
<td>Arlen; Brightman</td>
</tr>
<tr>
<td>Life skills</td>
<td>Hyman</td>
</tr>
</tbody>
</table>

* These lists offer examples only. No claim is made for their completeness.
# The named researchers would not necessarily choose to be labelled as they are here. For example, Jean Lave describes herself as an anthropologist although her research forms a crucial part of the foundations of situated cognition.

Academic and applied studies of problem-solving

11. “Academic” work here refers to research designed to further understanding of problem-solving. This work includes German Gestalt tradition, the Piagetian (Swiss-French) and Vygotskian (Russian)
developmental perspectives to problem-solving, the German project tradition in problem-solving, and finally work in North America originally conceived via the information processing paradigm and more recently under the paradigm of situated cognition.

12. "Applied" work includes the research on problem-solving and the use of concepts related to problem-solving in a variety of fields ranging from school mathematics to vocational education to business more generally. This work bridges the academic and the applied categories by pursuing an understanding of problem-solving within the context of improving learning in specific disciplines. In doing so, it treats the process of problem-solving itself as relatively unproblematic, proposes standard sequences of problem-solving steps, and explores the conditions for effective use of problem-solving to get work done.

13. This categorisation of problem-solving studies offers a way of viewing the extant work on problem-solving. However, all such categorisations are problematic. Obvious examples of this are the distinction between "cognitive psychology" and "learning in the subject disciplines," since the key researchers identified with these groupings are widely recognised as leaders in advancing knowledge about problem-solving in ways that apply beyond their specific disciplinary interests. This is also true, in some extent, in making a distinction between those working in artificial intelligence and those working in cognitive psychology.

**Academic Approaches**

14. While there are a distinct number of approaches to the study of problem-solving, few differences exist about the nature of problem-solving itself. While the language used to describe these elements and the connections among them differs, most individuals involved in the study of problem-solving are in general agreement with the following definition of the essence of problem-solving, derived from Mayer and Wittrock, (1996):

Problem-solving is cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver.

15. However, there are differences among the schools of work on the various components of problem-solving.

16. The components of problem-solving might be described as shown in the following table:

<table>
<thead>
<tr>
<th>The Problem Situation</th>
<th>(the recognition of a situation needing attention)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Context</td>
<td>(the setting in which the problem-solving takes place)</td>
</tr>
<tr>
<td>The Nature of the Task</td>
<td>(the tools which can be used, the cultural boundaries, time constraints,…)</td>
</tr>
<tr>
<td>The Problem Solver</td>
<td>(the content and procedural knowledge of the person undertaking the problem-solving, his/her ability to monitor progress towards achieving the goal, and his/her familiarity with the problem or similar problems)</td>
</tr>
<tr>
<td>The Problem-solving Process</td>
<td>(the interaction between problem solver, problem, and context)</td>
</tr>
</tbody>
</table>

17. The work in each category acknowledges the existence, more or less explicitly, of all of these components of problem-solving. But each set of work pays particular attention to one or two of these
elements. The following table identifies these differing emphases. Once again, this formulation is tentative and possibly controversial, and should be viewed in that light.

18. The academic fields of work on problem-solving reveal marked distinctions in the emphases placed on the elements of problem-solving. Their starting point is essentially the same, however. Each field is founded upon efforts to understand the nature of human problem-solving.

<table>
<thead>
<tr>
<th>Fields of work</th>
<th>Components of problem-solving emphasised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial intelligence</td>
<td>process; prior knowledge (Schank); situated perspectives (Winograd)</td>
</tr>
<tr>
<td>Cognitive psychology (including Complex Problem-solving)</td>
<td>process; problem solver’s knowledge (expert vs. novice)</td>
</tr>
<tr>
<td>Situated cognition</td>
<td>context; problem solver’s knowledge and the tools provided by the culture</td>
</tr>
<tr>
<td>Learning in subject disciplines, e.g., math, physics, reading, writing, history</td>
<td>process, problem solver (expert vs. novice)</td>
</tr>
<tr>
<td>Vocational education, e.g., medicine, electronics, mechanics,…</td>
<td>process, problem solver (expert vs. novice)</td>
</tr>
<tr>
<td>Generic work skills</td>
<td>problem solver, process</td>
</tr>
<tr>
<td>Business</td>
<td>problem, process</td>
</tr>
<tr>
<td>Life skills</td>
<td>? (maybe context)*</td>
</tr>
</tbody>
</table>

* The question mark here indicates lack of sufficient information so far to warrant an entry.

Artificial intelligence

19. Work on artificial intelligence focuses on the analysis of the processes people use to solve problems and the development of intelligent programmes that can emulate these processes, even to the extent of using their programmed heuristics to solve problems that were not part of their initial programming. While some of this work focuses on the characteristics of different kinds of problems and use is made of studies of expert vs. novice problem solvers to identify the problem-solving strategies that separate experts from novices, the main focus is on the problem-solving process. That is, the interaction between the problem solver and the problem.

20. This research suggests that problem-solving is a search through a problem in an attempt to close the gap between an existing state and a desired goal state, where the gap constitutes the problem. The problem solver develops a mental representation of the problem and adopts strategies to move from the existing to the goal state. Such strategies may be strong or weak and more or less strategic, depending on the knowledge demands of the problem, the knowledge of the problem solver and the problem solver’s experience with solving similar problems. Search strategies (or heuristics) identified through this research include:

- “trial-and-error”;
– adopting a backwards-and-forwards strategy of moving forward until a barrier is encountered, then backtracking to a point at which forward movement is again possible, adopting a modified strategy, and so on; and

– means-ends analysis, leading to a strategy such as adoption of a goal recursion strategy in which the problem solver develops a representation of the problem and the goal state that will allow him/her to establish and achieve intermediate goals while keeping the ultimate goal in mind.

Cognitive psychological research

21. The major interest of mainstream North American cognitive psychological research has been to uncover the processes by which people think and learn (Bransford, Brown, & Cocking, 1999). The role problem-solving has played in this research is that it has provided a vehicle for examining cognition. As such, the problems selected for use in cognitive psychological research studies have been selected more for their capacity to help reveal thinking processes than for their capacity to illustrate the kinds of problems people solve on an everyday basis. A classic example of such a problem is the “Tower of Hanoi” problem used by Newell and Simon. Such studies have helped to identify heuristics (strategies, “rules of thumb”) that problem solvers use to construct mental models of a problem situation and the goal to be achieved. They also deal with methods to identify how solvers find their way through the related problem space (the thinking space encompassing a problem, the goal state, and the pathway(s) between the problem and the goals(s)), in order to arrive at a solution. A further common feature of this research has been the study of different kinds of problem solvers as they tackle certain problems.

22. Other research in cognitive psychology has focused on the differences between experts and novices in specific content areas. The purpose of this research has been to uncover differences in their approaches with a view to establishing the implications for learning; that is, what novices in a given field need to learn in order to respond to problems in ways that are similar to how experts respond. Such research overlaps with the “applied” research on learning in the subject disciplines (e.g., mathematics, physics) and in vocational education (e.g., electronics, medicine) in comparing the problem-solving of experts and novices in these knowledge-rich domains. The results have been used to inform research on learning in these disciplines and in vocational occupations. The move towards concentrating problem-solving research in specific subject domains resulted from the lack of findings in early studies identifying approaches to problem-solving that would generalise to wide realms of application. As with the work on artificial intelligence, this research suggested that problem-solving heuristics that generalise across domains are usually weak and do not provide much assistance in the solution of real-world problems that demand strong, domain-specific heuristics.

23. Expert problem solvers, because of their superior knowledge and experience, demonstrate greater capacity than novices in a number of areas related to problem-solving. They are better at constructing a mental representation of the problem space, organising knowledge and relevant information into structures that will facilitate solving the problem; and in selecting and adapting more strategic and efficient heuristics for arriving at the goal state (Chi, 1989; Siegler, 1989).

24. Included in this work are the findings of those researchers associated with the school of Complex Problem-solving. Their orientation arose from different motivations than the North American focus on learning. Like the North Americans, the Europeans realised that the results from research on solving simple problems did not generalise to more complex and life-like problems. But rather than move, as the North American research did, towards domain-specific problem-solving, the European research focused on exploring the nature of complex problems.

25. From this point, two main approaches developed in Europe — the work initiated by Broadbent in Great Britain and the work initiated by Dörner in Germany. The two approaches share an emphasis on relatively complex, semantically rich tasks that are similar to real-life problems. Both approaches make extensive use of computerised tasks. In fact, the kinds of tasks used in this research would be almost impossible to construct without the technical capacity provided by computer programmes. The tradition
initiated by Broadbent, however, emphasises the distinction between the cognitive problem-solving processes that operate under awareness compared with those that operate outside of awareness. The tradition initiated by Dörner, on the other hand, focuses on the interplay among the cognitive, motivational, and social components of problem-solving. The problems designed to support this research utilise very complex scenarios with as many as 2000 interconnected variables. The most widely known example of such a problem is the “LOHNAUSEN” problem designed by Dörner. This problem involves a scenario in which the assessment subjects take on the role of being “a good mayor” of a simulated town.

26. According to Frensch and Funke (1995), Complex Problem-solving occurs in individuals to overcome barriers between a given state and a desired goal state by means of behavioural and/or cognitive, multi-step activities. The given state, goal state, and barriers between given state and goal state are complex, change dynamically during problem-solving, and are intratransparent. The exact properties of the given state, goal state, and barriers are unknown to the solver at the outset. Complex Problem-solving implies the efficient interaction between a solver and the situational requirements of the task, and involves a solver’s cognitive, emotional, and social abilities and knowledge.

27. Complex Problem-solving focuses on how people deal with novel and complex tasks. This emphasis on task novelty further distinguishes Complex Problem-solving from the North American mainstream tradition. Considerable work has been devoted to identifying factors that affect the complexity of a problem. These include:

- the number of variables in the problem and their interconnectivity;
- the dynamic aspects of problem situations; that is, the extent to which the conditions affecting the situation are subject to change over time or as a consequence of changes in related variables;
- the intra-transparency or opaqueness of the situation; that is the extent to which the problem and its characteristics can be clearly discerned;
- the number of goals and the need to select priorities and balance possibly competing or contradictory goals. (Funke, 1998)

28. Over time, two strategies have been adopted for studying how people interact with complex, novel problems. One has been to use naturalistic scenarios that embody everyday problems and to try to identify individual differences in the ways subjects respond to the problems. The other approach has been to create well-defined tasks with known characteristics (e.g., in terms of variables and their interconnectivity) and to systematically manipulate features of the task environment to test how people acquire and use knowledge in interacting with these tasks (Buchner, 1995).

Situated cognition

29. Situated cognition offers an alternative orientation towards problem-solving. It emphasises the relationship between the problem solver and the setting in which the problem-solving takes place. Drawing on studies of cognition by ordinary people in everyday situations, this work argues that knowledge is situated, being in part a product of the activity, context, and culture in which it is developed and used. Such a perspective has significant implications for learning. One aspect of this relates to the kinds of activities selected to support students’ learning. Authentic tasks are advocated over “school” activities. Another is the concept of cognitive apprenticeship in which students are enculturated into the ways of thinking in a given domain through a process of situated modelling and coaching. A third is the importance of collaborative approaches to learning. It is argued that learning in a group setting is important, not just because groups offer a convenient way to accumulate the individual knowledge of their members, but because social interaction is an essential part of the development of knowledge. (Brown, Collins, & Duguid, 1989).
Applied approaches

Learning in subject disciplines and vocational education

30. The emphasis in the applied fields of research on learning in subject disciplines and vocational education mirrors closely the emphasis on the problem-solving process described above with reference to the North American tradition in cognitive psychology. In fact, in large part, this work owes its existence to the failure of earlier work to identify general problem-solving strategies that would apply effectively across subject domains.

Business and work skills

31. Work on problem-solving with business applications includes both empirical research and training programmes and other personnel management advice. These reflect an emphasis on the nature of the problem and the problem-solving process. Other dimensions of these problems include factors such as the operating level (on a continuum from day-to-day problems that must be solved but do not usually have long-term effects to strategic problems that are critical issues with long-term effects) and problem awareness (a continuum from crisis problems requiring reactive problem-solving to opportunity problems that provide opportunities for proactive problem-solving) (Brightman, 1980; Barton, 1999).

32. Work in the business field often identifies standard sequences of problem-solving steps, such as: identify the need; plan the project; collect facts; analyse data; develop alternatives; present recommendation; and implement the decision (King, 1981). There are also findings in this field of study that would support an emphasis on the situational aspects of the problem. Meacham and Emont (1989) argue that problem-solving is essentially a social activity, suggesting some connections to situated cognition.

Identifying components of problem-solving

33. The literature identifies and describes a range of characteristics of each of the components of problem-solving. Some of these characteristics overlap (e.g., problems may be described as being knowledge-rich or knowledge-lean and may be described as being domain-specific or non-domain specific). Some of these characteristics may produce interactions among those components (e.g., the knowledge demands of a problem may interact with the knowledge possessed by the problem solver).

34. Research findings from academic cognitive psychology and Complex Problem-solving studies, and to some extent from applied business problem-solving research, yield information about the characteristics of each of the elements of problem-solving. These results may be an artefact of the generally fine-grained nature of these studies. However, the search for information at this level of detail is not complete, and it is possible that other fields may yet yield relevant material.

35. The purpose of unpacking these components and characteristics is to:
   – highlight the complex nature of each component; and,
   – illustrate the variety of dimensions of problem-solving that need to be considered in the design of an assessment instrument.

36. The listing of problem characteristics here is not comprehensive, nor is the discussion of each characteristic complete. A study of problem-solving needs to be sensitive to the wide variety of characteristics that can be associated with problems. Problems can, by themselves, have several dimensions, such as the degree to which they are open- or close-ended. The actual conditions under which the students are asked to solve the problems and the resources available to them during that time constitute
another set of considerations. One also has to consider the previous experiences of the problem solver—
to see to what degree have they experienced similar problems or questions. A fourth area of considerations
is the degree to which specific problems may require unique strategies or not. Additional characteristics
may be added or removed as the work on this framework progresses.

Problem dimensions

Degree of problem definition

37. Problems can be viewed as “well-defined” or “ill-defined.” Problems vary both in terms of the
completeness of problem specificity (thus, how well the problem is understood) and the certainty with
which the correct or optimal solution can be recognised (Arlin, 1989). Problems also vary according to
the extent to which a problem provides values for its “parts.” By considering the known, partially known,
and unknown information it is possible to differentiate among well-defined and ill-defined problems, as
well as to develop solution paths. Well-defined problems have relatively specific values while ill-
defined problems have one or more unknown values.

38. Well-defined problems relate to repetitive and routine tasks which are well defined and can be
solved by standardised/automated procedures (Brightman, 1986). These problems are more like “puzzles”
(Luszcz, 1989) that have one specific goal or answer and are solved using explicit rules and algorithms.
Because of their explicitness, the problem space for well-structured problems is a closed system.

39. Ill-defined problems, on the other hand, are less specific, harder to grasp, may have multiple
goals and solutions. Their solutions are more difficult to attain because the problem space for ill-defined
problems is an open system (Luszcz, 1989). Ill-defined problems are novel, elusive, often ambiguous
“out-of-focus messes” (Brightman, 1986). These types of problems are solved using judgement,
creativity, problem-solving processes, and heuristics.

Degree of domain-specificity

40. The degree of domain specificity refers to the extent to which a problem might be inserted into
the specific practices of a certain area of expertise. Domain specificity deals with the degree to which the
problem, because of the knowledge required, must be categorised in one or a small number of academic
domains. For example, a problem dealing with the solution of a Diophantine equation would fall with in
the domain of elementary number theory in mathematics. In the literature, this characteristic is also
referenced by terms such as knowledge-rich vs. knowledge-lean; high vs. low knowledge of relevant
subject matter.

Nature of and relationship among variables

41. The nature of and relationship among variables deals with the extent to which the values of the
variables involved in a problem are fixed, static, both over time and in relation to the other variables. That
is, does the relationship between variables and the actual values of the variables themselves remain
constant, static, or shift, dynamically, as problem-solving occurs.

42. Another major characteristic involving variables is whether the problem is univariate or requires
multiple variables. Research has shown that the complexity of problem-solving is positively correlated
with the number of variables involved. Multivariate problems can be further categorised by whether the
values of one variable are dependent on those of another variable, or by whether the values of the
variables, or specified subsets of the variables, interact with one another in a variety of ways.
Problem intensity

43. Arlin (1989) defines problem intensity as the motivational attraction of the problem where the problem solver is sufficiently engaged that it is worthwhile to attempt a solution. Brightman points out that problem intensity can affect the problem-solving processes in business situations. High intensity, or “crisis,” problems are forest fires that cannot be avoided and require immediate action, while opportunity problems exploit possibilities for action and are more long-term endeavours. Both crisis and opportunity problems are important, motivating, and must be solved but the immediate nature of crisis problems leads to reactive types of actions and problem-solving while opportunity problems are solved in a proactive manner.

Degree of realism

44. This problem characteristic applies to presented (assessment, instruction, research) problems. Problem temporality, as this characteristic is sometimes called, is the perception of whether or not the problem is one that may be encountered by the problem solver (Arlin, 1989). If the problem presented is one that is unlikely to be experienced, what is the motivation for expending effort toward a solution? If the problem is perceived as one that may be encountered in the present or near future then there might be a higher motivational level associated with an attempt to reach a solution.

Nature of solution

45. Problems vary in their expected, or actual, solutions. Some problems, especially those that can be modelled by an equation, have a unique, or deterministic, solutions. Other problems, for example problems involving design and construction issues, have a multitude of possible solutions. The degree to which a problem has a very narrow solution space or a very expansive solution space affects the manner in which students approach and handle problems.

Required response mode of problem

46. Another characteristic of problems is whether they require the problem solver to select a solution or to create or fabricate a solution. In many cases these are now referred to as multiple choice, matching, or true false for the selection options and regular student constructed or extended student constructed problems for those requiring student work.

The context or setting in which the problem is to be solved

47. The administration of a problem-solving assessment brings with it additional considerations which must be factored into the design of the assessment and into the interpretation of the resulting data. Several of these issues are discussed in the following.

Context

48. Central to problem-solving tasks are the situations themselves and the conditions they bring to the problem solver. To what degree does the problem involve context with specific situational requirements? How many different disciplinary domains are called on for a solution to be developed? What requirements are made in terms of context-specific processing requirements?
Task requirements

49. What logical demands does the task place on the problem solver? What requirements are there for communicating the results? Is the process required single or multi-faceted? This particularly relates to the degree of information processing required by the problem solver. How complex is the issue at hand?

Time constraints

50. This consideration deals with the degree to which the problem-solving assessment a timed assessment. Do students have to work under a time constraint? Will their problem-solving processes curtailed prior to showing all of the information they know and being able to construct all of the answers they could have produced? This is akin to the question of speediness in a classical assessment design.

Access to additional resources

51. In contemporary assessments, students often have access to the technology they commonly use in regular learning settings. This may involve hand calculators, computers, or other forms of electronic information retrieval tools. It may involve access to non-human resources such as the library in a school, current newspapers, or community resources. This is contrasted with assessment administrations where students work at a desk with pencil-and-paper in the absence of any supporting resources. This also raises the issue of cultural tools and cross-cultural differences in the use of different cultural tools. There are different forms of calculation instruments and different cultural reactions to different test response formats.

Social interaction

52. Some assessments provide for a measure of the affect of working alone, as compared to working as part of a group of problem-solvers. The Pacesetter Mathematics assessment of the College Board provides both individual and group assessments as part of its culminating assessment process. Vygotsky has paid a lot of attention to differences in problem-solving ability when subjects work alone versus when they work with others or with a teacher. People differ in their ability to profit from instruction. Some researchers have tried to make this issue relevant to current methods of assessment (Brown, ).

The problem solver

53. Two of the key characteristics that must be considered in any assessment are the background and knowledge of the problem solver. These and other dimensions related to the problem solver are detailed below.

Knowledge of relevant subject matter

54. This dimension deals with the degree to which the problem solver might be considered an expert or novice in the field of the problem posed. What knowledge does the individual have in the specific domain if the problem is domain specific? To what degree is the problem solver’s knowledge connected across disciplines and to what degree can the problem solver shift among multiple representations?

High strategic/procedural knowledge

55. This second dimension deals with the toolkit of heuristics and algorithms the problem solver has relative to the problem posed. A problem solver with knowledge of explicit strategies and algorithms
related to the problem posed is much better off than a problem solver armed only with general strategies and disconnected knowledge of related procedures.

Familiarity with the problem

56. The degree that a problem solver is familiar with the problem is a third dimension. However, it is not only the familiarity with the problem, but also knowledge of the subject domain to which the problem belongs. What appears to one person as a novel and seemingly impenetrable problem may be little but a routine task to someone else. These differences may relate to individual differences in knowledge of the subject matter or to problem familiarity developed through numerous encounters with a variety of problems, or both. One of the characteristics that separates experts from novices is the ability of the former to recognize patterns of information relevant to a problem and quickly identify those patterns likely to lead to a satisfactory solution. For example, see the studies of chess experts and novices. These abilities derive not only from knowledge of the relevant subject matter but also from extensive experience with similar kinds of problems.

57. How well do the contents of the problem match with the “stuff” of the problem solver’s everyday life (Arlin, 1989)? This is an important point when discussing cross-cultural issues and when attempting to transform contrived, knowledge restricted, or research-based problems into real-life situations.

58. A couple of concerns are immediately apparent. The first is that if a problem becomes too familiar, the problem solver may view it as a well-known, often encountered problem and therefore employ a “rule-of-thumb” solution. Thus, the problem becomes a non-problem since the solution is obvious or trivial and there is no opportunity to assess the problem solver’s analytical reasoning abilities or problem-solving.

59. A second concern is that what may be familiar to one person is not necessarily familiar to another. In particular, what the problem designer might believe to be familiar content may be unfamiliar to the solver. This is an obvious concern in cross-cultural settings but it is also a critical element when dealing with age and social differences.

Motivation

60. A fourth consideration is the degree of motivation shown by the problem solver in confronting the problem. Students facing assessments that drop into their classrooms from the sky and have little consequences for them may be less motivated to show their full problem-solving powers than a student who knows that the results of the assessment have meaningful consequences. Motivation may rise and fall due to the interaction of a number of variables in a student’s life. Disregarding the degree of potential motivation and effort on the part of the student can lead to drawing wrong conclusions.

61. Also tied to this area are differences in students’ achievement motivation. Some are driven by intrinsic factors to succeed on assessments. Others need extrinsic factors to motivate them to achieve. When assessments appear, as the PISA assessments, for which there are no external rewards or individual scores to be reported for either the student or back to the school, personal motivations, or motivations of students from an entire ethnic group or culture, could be affected.

Assessing problem-solving in PISA

62. The foregoing material has provided a review of the extant knowledge concerning problem-solving from both an academic and applied viewpoint, it has reviewed some of the issues central to considering factors in problem-solving, and it has delved into aspects of what an assessment of problem-solving would have to consider in starting to measure students’ work in a cross-curricular setting. Central
to all of these issues is the topic of analytical reasoning—the ability of students’ to reason, and reason well, in inductive, deductive, and critical/complex settings. It also deals with students’ abilities to reason in settings where goals have been established for them and situations in which they must set their own goals and targets as they consider the problem at hand. It is these issues that should define the core of a problem-solving assessment in PISA.

63. Most researchers involved in the study of problem-solving, via one conception or another, generally agree that the general essence of problem-solving is as follows:

   Problem-solving is cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver (Mayer and Wittrock, 1996).

64. Problem-solving in the school curriculum is differentiated from such general problem-solving in that goals are often set for students, whereas in the real world, the goal may or may not be evident at the outset. In the PISA assessment of problem-solving both forms of problem-solving will be studied, those with externally established goals and those with goals, or sub-goals, set by individual students as their study of a problem unfolds. The paper refers to the former class of problems as goal-directed problems and the latter class as open-ended problems.

65. PISA assessments also call on students to display their problem-solving skills in a variety of settings. Each of the assessments, reading, mathematical literacy, and scientific literacy has problem-solving components listed in its framework. In reading, problem-solving is assessed through students’ attempts to bring meaning to passages, to develop interpretations, and to determine voice. In mathematics, problem-solving is assessed through students’ posing and solving of problems. Particular emphasis is given to students’ choice of solution strategy, correctness of work and purported solution, and reflection on their work. In scientific literacy, problem-solving is assessed through students’ identification of relevant knowledge and related questions and through their drawing and evaluating of conclusions (PISA Assessment, 2000).

66. The PISA assessment of problem-solving moves beyond these frameworks in establishing an additional set of measures, measures of the students’ performance in problem situations where the:

   – content moves from familiar material to unfamiliar settings,

   – context moves from school-based topics to real-world applications, and

   – complexity moves from simple and complex translation activities to situations calling for the multiple applications of processes and heuristics.

67. Such problems call on students to make executive decisions about the paths of work to follow and to determine the relative efficacy of different courses of action as they work to arrive at a solution. In particular, they involve the students’ capabilities to engage in analytic and analogical reasoning settings. These allow measures to be developed to describe the efficiency of students’ work, the productivity or idea generation aspects of their problem-solving approaches, and their command of the central processes of problem representation and the use of problem-solving heuristics.

68. The question of what problem-solving competencies an individual needs for life is more closely related to the solving of problems in a cross-curricular format than it is to solving problems in the single domain of a specific school subject. Everyday real-life problems call on individuals to merge knowledge and strategies from a variety of fields to reach some resolution. These problems call for individuals to move among different, but sometimes related, representations and to exhibit some degree of flexibility in the way in which they retrieve and apply their knowledge. The problems in the PISA assessment of problem solving engages students in contexts involving problems they have not seen in their school work or other examinations and, thus, provides a measure of their transfer of problem-solving and reasoning competencies to new settings.
This paper examines the essence of such problem-solving through both the frame of a cross-curricular assessment of problem-solving competencies and as seen through assessing problem-solving within the extant portions of the PISA assessment using released items from the first cycle. These examples are provided to draw the distinction between the nature of items used and the information gained from the cross-cultural assessment of problem-solving and the current PISA assessments of problem-solving in the areas of reading, mathematical literacy, and scientific literacy.

The assessment of problem-solving as a cross-curricular competency calls for the development of ways to assess a student’s quality of knowledge and the processes through which that knowledge is applied in non-routine problem-solving settings. The goals of this assessment are to find how the student:

- uses conceptual understanding and procedural knowledge in non-routine settings;
- applies reasoning to understand a given problem context and related information;
- employs strategies and representations to link assumptions to desired goals;
- formulates questions to understand the task at hand;
- uses a variety of forms of reasoning in new settings;
- uses knowledge from one situation to reason in another; and
- evaluates his/her current work and makes adjustments prior to citing he/she has a solution.

Such foci allow for the analysis of students’ ability to generate possibilities, to define and search problem spaces, to generate possibilities, to validate courses of inquiry, and to both communicate and reflect on their problem-solving activities. Such opportunities are rarely followed in subject-based inquiries of student problem-solving, as they focus on students’ abilities to acquire the correct answer or to apply the appropriate procedures.

In the PISA assessment of problem-solving, the major feature of interest is the analytical reasoning that a student applies in addressing new and non-routine problems. Content specific problem strategies are developed within the school subject fields and are assessed there. What is specific to cross-curricular problem-solving is the analytical reasoning competencies that students develop throughout their schooling and which they are able to apply in new and non-routine situations which ask them to integrate their knowledge from individual fields of study. Problems measuring such competence should involve both a new context and levels of expectation of student performance than what would be found in a content-based area. Specific examples will be provided later. With analytical reasoning at the core of the cross-curricular competencies in problem-solving, it will be important to separate these reasoning behaviours from those normally associated with problem-solving in content areas that are more discipline-based.

The assessment of students’ cross-curricular competencies in problem-solving call for students to reason in situations where the problems:

- are novel, but presented in meaningful situations;
- are based around real-life contexts, distinguished from those seen in instructional settings;
- involve the application of integrated curricular content from mathematics, science, civics, history, economics, sociology, or other area; and,
- require students to apply analytical, including analogical, reasoning skills.

The important feature is that the situation posed must call on the student to integrate their knowledge and understanding in new and unique ways.
75. In doing so, the assessment must provide opportunities to see students’ reflective, self-regulating, problem-solving behaviour as much as possible, to monitor and evaluate the strategic approaches they employ, and to note the formal models and representations they employ in their work.

76. On the other hand, an assessment students’ disciplinary problem-solving competencies in the mathematics framework would attempt to examine many of the same features. However, the focus would be more on specific strategies developed for use in situations involving quantities, spatial knowledge, chance, or data. The contexts involved would:

- be more familiar from a curricular standpoint;
- employ more familiar problem-solving approaches;
- be placed in novel, but related to familiar, settings; and,
- be restricted to mathematical situations.

77. The focus would be on connecting and integrating knowledge within mathematics as a discipline and in developing new mathematical knowledge through thinking, generalisation, and insight as defined in the PISA framework for mathematical literacy (OECD, 1999).

78. Each discipline has specific problem-solving strategies that students are expected to master and apply to a wide variety of exercises within the discipline. It is analytical reasoning that ties these cognitive approaches together with their new representations and interpretations. The student who is successful in employing inductive, deductive, and critical reasoning in cross-curricular settings is the one who is able to note the given information, abstract the structure out from the context, reformulate it, and act on it to produce new knowledge or structures.

79. Analytical reasoning strategies go beyond students’ opportunity to learn logic and reasoning in the classroom. They are the ultimate problem-solving tools that transcend disciplinary boundaries. Items selected for the cross-curricular assessment must allow for students to show their reasoning. While problems are classified with respect to being goal-directed or open-ended, it is the way students approach these problems that is of interest. While problems are also classified with respect to whether they require inductive approaches, deductive approaches, or more mixed critical thinking strategies, it is the nature and quality of student thinking that is of primary interest.

80. Analytical reasoning was selected as the core of the PISA problem-solving assessment efforts because such reasoning is the heart of the problem-solving process. The nature by which one represents problems, selects strategies for attacking them, transforms their contents to equivalent formats, and solves them involves a carefully welding together of strategies and knowledge developed with individual disciplines. The levels of use of inductive, deductive, and critical reasoning skills that separate the expert problem solver from a novice are well known, but these have rarely been examined in school settings in such a large-scale assessment. The inductive and deductive strategies are used to manipulate and investigate certain concepts. The critical thinking strategies operate at a higher level in examining concepts. However, they are also used in testing statements purporting to explain various issues or to provide persuasive explanations.

81. This reasoning is often reserved for study by advanced or ‘gifted’ students. Most students actually “catch” it from their teachers, rather than have it formally taught to them. However, when students must reason in non-routine real-world contexts involving social, political, ethical, or personal issues and, at the same time, apply knowledge from different disciplines (including mathematical and scientific knowledge) it is analytical reasoning that provides both the direction of thinking and provides the glue that ties the work together. The reasoning runs from control of simple operational skills with propositional logic to larger units of these processes. Such larger units involve samples of inductive reasoning, deductive reasoning, and critical thinking that employ rules of inference, hypothesis testing strategies, and checks for common misconceptions and misapplications of logical reasoning. Such
applications of logic are applied in both goal-directed and more open problem-solving situations to problems from cross-curricular work involving combinatorial, proportional, probabilistic, and spatial aspects.

82. Analytical reasoning starts in the early grades as students begin to classify objects and make comparisons between objects as they develop language. Later the production of the child’s language is guided by applications of this logic as students begin to structure their thoughts using connectives such as “and,” “or,” and “not.” Gradually the logic of “If..., then...” statements begins to formally emerge. With them come understanding of the more formal uses of reasoning schema known as modus ponens, modus tollens, law of syllogism, analogical reasoning, use of the contrapositive, and other approaches to structuring and evaluating arguments.

83. This growth of analytical reasoning tools and skills is accompanied by a greater metacognitive awareness of the uses of analytical reasoning by students and their teachers. This growth is shown by the broader realm of representations that students use in characterising problems. Students must develop both ways of structuring their thoughts and actions within disciplinary based work and, at the same time, begin to monitor their actions in doing so. This shift in the level of metacognitive awareness of reasoning processes and representational skills in school parallels the development of students’ ability to use these processes across disciplinary walls. The more students become aware of their analytical reasoning and its application, the more able to avoid misconceptions, prejudices, and mistakes arising from content issues.

84. Problem-solving in cross-curricular settings calls for skills that focus one’s competencies in complex settings drawing on varied aspects of one’s total knowledge base. Such problems are more real world in that they draw on two or more disciplines. They are embedded in rich contextual settings, settings that often involve emotional aspects for those involved in the problem-solving. These emotional aspects add to the difficulty levels, as well as to the complexity levels, in such problems. Such problems often involve aspects that can trigger specific reasoning or conceptual miscues that plague students’ growth as they try to apply their knowledge in wider domains.

85. To ascertain the true scope of students problem-solving competencies in cross-curricular settings, PISA must assess students’ competencies in settings that focus on the solution of complex tasks similar to those that may occur in real life. Such tasks are necessary in order to be able to test the possible effects of schooling on the development of analytical reasoning skills. These items must include tasks of greater complexity, greater novelty, and contexts where prior knowledge may stand in the way of drawing valid conclusions.

86. Recommendations for change in education world wide have assumed that the knowledge acquired in school can improve students’ everyday thinking by making students more capable of applying known schemata in novel situations, preparing them to deal with more complex situations, and immersing them in situations that require restructuring existing knowledge (as in some gestalt problem – solving or as in some of the conceptual change literature) to form new conceptions or ways of approaching problems. The proposed PISA framework outlines a way of assessing such goals.

Possible framework for the assessment of problem-solving in PISA

87. The framework for assessing problem-solving in a cross-curricular setting consists of a matrix with columns representing the directiveness of the problem posed and the rows representing the type of reasoning required to solve the problem:

<table>
<thead>
<tr>
<th></th>
<th>Goal directed</th>
<th>Open ended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inductive reasoning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deductive reasoning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical reasoning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
88. The goal directiveness is determined on the basis of setting a very deterministic situation for the students and directing to find that solution. An open-ended problem is one where students have more freedom to consider possible outcomes and select a strategy for comparing and contrasting among them.

89. Items measuring goal-directed inductive reasoning would be items that propose a specific goal, but only provide students with information about cases from which they can abstract a pattern and propose a generalisation as a solution based on the cases. An open-ended inductive reasoning problem might be an item that provides students with data and asks them to determine a particular outcome. The problem may specify a specific line of attack or method for analysing the data as well.

90. Items measuring goal-directed deductive thinking might have students consider a law, rule, or generalisation and a set of evidence and then build an argument based on the evidence and the nature of the principle given. Students would be given a specific goal and generally be either given the generalisation or directed toward it by one or another means of specification.

91. A goal-directed critical-reasoning problem might present students a series of data sources and a reading or a listing of information. Students would have to combine various aspects of inductive and deductive reasoning, along with organisation of intermediate findings to structure a final specific outcome as solution. Such problems require a number of logical decisions and generally involve a reformulation of information and multiple representations of given information in moving to the solution.

92. Items measuring open-ended inductive reasoning might present students with a listing of cases and ask the student to consider the cases and discuss what possible related outcomes might be. In particular, such problems might have an embedded pattern or generalisation in the data, but it would remain for students to discern the generalisation and develop it.

93. Items measuring open-ended deductive reasoning might provide students with a folder of data and a listing of laws relative to the situation and ask students to develop the materials in the folder. Such an item might present students with a dilemma and leave students to argue one side or another. The products might be assessed with respect to the quality of the argument that was constructed and the logical links evident in the structure of the students' reasoning.

94. An open-ended critical-reasoning problem might present students with a complex set of data and directions to potential outcomes. Students would be left with the task of bringing order out of chaos and developing some structure and organisation, including some generalisations, to develop understanding of the situation and knowledge of the principles that might be embedded in the data. As in previous cases, the organisation and discussion would be based on the structure developed through a combination of inductive and deductive reasoning on the part of the student.

95. Items meeting any of these descriptions can be embedded in more complex project-like tasks (such as those of Klieme). One project, or theme based assessment block, can include a number of items, possibly of different formats, that test a variety of the cells of the model proposed. Some may be inductive, others deductive. Some may measure goal directed aspects of topics within the theme, while others allow the students to impose their own structures on the situation. As one examines any of these problems relative to the problems found in the curriculum area based assessments, they would see that the content has shifted from familiar to unfamiliar, the context from school based to real world, and the complexity from relatively straightforward to novel and challenging.

**Goal-directed versus open-ended problem-solving**

96. The contrasting of goal-directed with open-ended problem-solving would see differences in the items along the dimensions of:

- the end-state is specified and the problem-space is bounded,
– the rubric is considerably tighter in expected outcome format,
– the means-ends relationships are more obvious and specified,
– the possibility of ordering being specified, and
– the evaluation strategy for problem being more evident.

97. The analysis of student work in open-ended settings may reflect students being able to set goal as a metagoal, that is, to what degree can they reflect on the setting and maintenance of a self-set and self-regulated goal in problem-solving.

Inductive problem-solving

98. In assessing the quality of student reasoning in inductive settings, the assessment should provide students with ample opportunities to show their command of argument through similarity and analogy, through pattern recognition, and through rule induction. In the latter case, one would be looking to see whether the student looked for confirmatory and exclusionary evidence. Evidence should be collected to see the degree to which students have some flexibility in applying the simple operational skills associated with inductive problem-solving methods. Situations here might include comparison buying or trip planning.

Deductive problem-solving

99. In assessing the nature and quality of student reasoning in deductive settings, the assessment should provide students with ample opportunities to show their command of argument through applications of reasoning schemes built on prepositional logic, especially those where the logic must be applied to contextual based information. In particular, are students capable of handing the standard argument structures and reasoning through syllogisms and analogies. Situations here might include interpretation of zoning requirements, structuring of proofs, and restructuring of common language into more formal arguments.

Critical/complex reasoning

100. In assessing the quality of student competencies in critical reasoning, the assessment should provide students with complex logical decision making situations. These are situations involving larger units of processes. They, most likely, are drawn from the social sciences, but involving mathematical or scientific aspects. To what degree can students critically evaluate information that they are given (checking for falsification, using hypothesis testing, etc.), exhibit aspects of proving ability (using processes to establish the validity of general statements, relate the relevance of the information to tasks at hand, etc.) The information load in these problems is heavy and multiple comparisons or contrasts are required. These situations require careful organisation and reordering of information, as well as careful interpretations to avoid forming misconceptions. Situations here might include issues of health, risk, environment, cost-benefit, or population.

Complex problem solving

101. The PISA problem-solving assessments should investigate analytical reasoning in a variety of situations, contexts, and combinations of disciplinary requirements so that the novelty, complexity and counter-intuitiveness of all facets can be manipulate and examined with care.

36
Future international options

102. In addition to the core portions of the cross-curricular problem-solving assessment, countries participating the PISA assessment should consider the possibilities of two international options being considered for future cross-curricular problem-solving assessments. These options involve assessments of collaborative problem-solving and computer-delivery project assessment designed along the lines of the work of Klieme and his co-workers.

Collaborative problem-solving

103. The collaborative problem-solving option could consist of a separate block of items which students would complete in groups of three. Items in such blocks could build off of items in one of the other blocks of the regular cross-curricular assessment. This would allow for a comparison of students’ work in individual settings with their work in group settings. Such assessment blocks would have to allow time for idea generation and formulation and for the development of group roles on the part of the students involved.

104. The *Pacesetter* programmes of the College Board such have working models of such assessments of group problem-solving. Many recommendations for student competence in problem-solving and in education in general have called for the development of problem-solving competencies in an environment that values social learning. If so, it must also be assessed. Given the relation of this to country specific goals for students and the ease with which it can be pulled off in one country versus another, this is left to be developed as an international option within the PISA cross cultural problem-solving assessment.

Computer based delivery

105. This alternative would allow for the administration of computer delivered project work along the lines described by Klieme (paper to Network A). Such assessments provide a vivid display of students’ problem-solving competencies in a dynamic environment. They also provide for an examination of the ways in which students order and conduct their work in complex settings in a way that no paper-and-pencil based assessment can provide.

*Features of a problem-solving assessment*

106. The assessment of problem-solving in PISA’s must constantly be focused back on PISA’s definition of problem-solving. That is problem-solving that builds on school education programmes, but extends the problem-solving competencies and knowledge developed in school to consider rich and novel contexts that take the student beyond situations they have experienced in the classroom. These situations should also challenge students to address problems that extend, in either goal-directed or open-ended ways, to settings that involve information and concepts from at least two different disciplines. This is the type of learning that students should carry into their lives and chosen fields of endeavour. Students who can solve problems involving a number of issues will be well prepared to address the problem-solving situations they will encounter in life.

107. Other issues are important in measuring students’ problem-solving skills and developing assessments that do so. Some of these are discussed in the following sections.

*Accessibility/equity*

108. The test should be accessible to all students participating in educational programmes in the participating countries. Making an item accessible means that the item can be understood and addressed
by 15-year-old students regardless of the curriculum in which they are enrolled. Items should be developed in a fashion that presents them in a representational mode (graph, table, words, symbols, pictures,...) that is easily interpretable by all students.

109. Further, it is assumed that care will be taken to see that other standard forms of bias are avoided in the design and construction of the test items. For example, excessive technical vocabulary, difficult reading level/vocabulary, and items calling for specific personal life experiences should be avoided.

Information sources

110. Items calling for specific information sources, i.e. specific data, should have that information provided as a part of the assessment. The provision of such information could be done within the item or as an accompanying data-bank book. Part of problem-solving is knowing where and how to look up information at appropriate times. If such information is required, it must be provided as part of the assessment package for all students, either within the blocks of items themselves, or in a data book that all students would have as part of their assessment package.

Technology

111. An assessment of problem-solving is not an assessment of students’ ability to perform calculations. As a result, all students participating in the PISA problem-solving assessment should be allowed to use any hand calculators they routinely use in their regular classroom learning environments. The decision of whether to use calculators should rest with the individual student based on their chosen approaches to the items. No item should be constructed so that its solution is solely dependent on calculator usage or of a length that students not using a calculator would be disadvantaged in performing any calculations required.

Manipulatives/formula sheets

112. Items requiring any manipulatives (geometric pieces, counters, spinners,...), rulers, or protractors, maps, formula booklets, or other materials to be handled in the course of solving a problem must be designed so that those materials can be provided to each student during the assessment sessions. Such materials should be designed to be inserted in the assessment booklets for student accessibility. No special materials should be required that demand extra handling by teachers or others administering the assessment.

Assessment administration conditions

113. Where possible, the assessments should be administered in students’ regular classroom settings. Such administrations help in promoting an academic atmosphere for the assessment and help assure that regular materials (calculators, pencil-sharpeners,...) are all available as they would be on a regular school day. Further, the assessment should be designed to be administered in a single school day in two settings, separated by a short break. This will alleviate the necessity of scheduling make-up administrations for students missing one of the administration periods should the assessment be spread over two or more days.
**Items and rubrics**

**Principles for item selection**

114. Throughout the design process and development of the final blocks of items, it is important that, to the extent possible, all items should be:

- reflective of the cross-cultural attributes discussed earlier;
- broad, rather than narrow, in scope and reflect real-life contexts;
- free of double-negatives, excessive technical vocabulary, and other
- accessible to 15-year-olds in terms of mathematics content and difficulty; and,
- reasonable in light of the views of the general public.

**Item types and examples**

115. In previous large scale assessments of problem-solving the majority of items used have been multiple-choice, true-false, or short response items. These items were used in the name of reliability and because of their low-costs to score and the ease with which such forms could be administered. However, to adequately ascertain a student’s ability to reason, problem solve, and communicate the results of such activities, more extensive records of a student’s work are needed. Hence it is proposed that a variety of item formats will be employed in developing the PISA problem-solving assessment.

**Multiple-choice**

116. Multiple-choice items are appropriate for quickly and inexpensively determining whether students have mastered certain skills, knowledge, or information gathering abilities. Well-designed items can measure student knowledge well beyond simple conceptual and procedural knowledge. They can be designed to reach beyond the ability of students to “plug-in” alternatives or eliminate choices to determine the correct answer. However, these items are somewhat limited in their ability to ascertain the breadth and depth of a student’s knowledge for many contexts.

**Short student constructed responses**

117. Short student constructed response items are items designed to require students to construct their own responses to items. This allows for examiners to ascertain what students can produce from their own understanding of the item and to display the heuristics they have used in approaching the item. SCR items either require students to give brief answers representing either a numerical result or the correct name or classification for a group of objects, draw an example of a given concept, or, perhaps, write a brief explanation for a given result. In general, these items should:

- Require about 2 to 3 minutes of work;
- Ask the student to show integration of information or concepts, along with the way in which these lead to a solution to the problem proposed;
- Tap multiple areas of understanding and require their connection in the response displayed by the student;
– Require the student to justify their answer;
– Use where the situation requires multiple steps to a solution and has several different components;
– State explicitly what the student needs to do in responding.

Extended student constructed responses

118. Extended Student Constructed Response Items are items that require students to consider a situation that demands more than a numerical response or a short verbal communication. These items require students to carefully consider a situation within, or across, content areas within a discipline(s); understand what is required to solve the problem; choose a plan of attack; carry out the plan; and interpret the solution in terms of the original situation. These items require students to provide more complete evidence of their work or to show that they have used more complex thought processes in solving a problem. In either case they are expected to clearly communicate their decision-making processes in the context of the problem (e.g. through writing, pictures, diagrams, or well-ordered steps).

Rubrics

119. Holistic rubrics for evaluating the student responses to items should be constructed about a general framework that values the stages of problem-solving laid out by Polya in his famous book on problem-solving (Polya, 1945). Such rubrics would be noted by giving special attention to student work achieving the various levels of restating the problem and noting assumptions, designing a method of approaching the problem, solving the problem, and reflecting on the solution—perhaps testing it or modifying it for special cases.

120. For regular student constructed response items, the rubric should provide for rating levels of:
   * No response.
   0. Incorrect or irrelevant remarks.
   1. The response contains evidence of an understanding of the problem at a conceptual level evidenced by the logical approach taken. However, on the whole, the response is not well developed. Although there may be serious logical errors or flaws in the reasoning, the response does contain some correct work. Examples provided are incorrect or inappropriate.
   2. The response demonstrates a complete understanding of the problem, is correct, and the methods of solution are clear, appropriate, and fully developed. The response is logically sound, clearly written, and contains no errors beyond ones that may be a result of miscopying from elsewhere in the student’s work. Examples are well chosen and fully developed.

121. For extended student constructed response items
   * No response.
   0. Incorrect or irrelevant remarks.
   1. The response indicates a minimal understanding of the problem posed but does not suggest a feasible approach to a solution. Although there may or may not be some correct work signifying a logical approach, the response is incomplete, contains major errors of reasoning, or reveals other serious flaws. Examples are absent.
   2. The response contains evidence of an understanding of the problem at a conceptual level evidenced by the logical approach taken. However, on the whole, the response is not well
developed. Although there may be serious logical errors or flaws in the reasoning, the response does contain some correct work. Examples provided are incorrect or inappropriate.

3. The response demonstrates a clear understanding of the problem and provides an acceptable approach. The response is generally well developed and coherent but contains minor weaknesses in the development. Examples are provided, but not fully developed.

4. The response demonstrates a complete understanding of the problem, is correct, and the methods of solution are clear, appropriate, and fully developed. The response is logically sound, clearly written, and contains no errors beyond ones that may be a result of miscopying from elsewhere in the student’s work. Examples are well chosen and fully developed.
CROSS-CURRICULAR COMPETENCIES IN PISA
TOWARDS A FRAMEWORK FOR ASSESSING PROBLEM-SOLVING SKILLS

FROM COMPREHENSION TO LITERACY:
THIRTY YEARS OF READING ASSESSMENT

Dominique Lafontaine, University of Liège, Belgium

Introduction

1. In 1973, in the introduction of a book dedicated to the reading comprehension component of the IEA Six Subject Study, Robert Thorndike pointed out that it is both simpler and more complex to elaborate appropriate assessment material for reading than for other subject matters. “In the field of reading, there seems to be much more consensus as to the objectives of the instruction. (...) There would be general agreement that children should be able to get meaning efficiently from written material of various styles and content areas. On the other hand, the preparation of genuinely equivalent tests in reading, where the essence of the task very intimately involves the language of a particular country, would seem to present very serious difficulties” (Thorndike, 1973, p. 14). In a few words, this quotation lays out the issues that international reading assessments progressively sought to address since then. In comparison with other disciplines (such as mathematics or science), the links to the curriculum remain rather loose in reading, while cultural and linguistic connotations prevail. This chapter shows that, even though comparative assessment has advanced considerably, its key challenge remains unchanged: to achieve comparability by coming closer to true equivalence. Before coming back to this important issue, the paper examines major developments that have taken place during the three decades that elapsed between the IEA Six Subject Study and contemporary studies such as PISA, PIRLS and ILSS3. Features of the various studies examined here are summarised in the table at the end of this chapter.

---

3 In 1971, the International Association for the Evaluation of Educational Achievement (IEA) launched an extensive programme of comparative studies in six different subjects (reading comprehension, literature, civics, science, and English and French as foreign languages). The OECD’s Programme for International Student Assessment (PISA) is organised in several cycles and assesses reading literacy, mathematics literacy and science literacy. The Progress in International Reading Literacy Study (PIRLS) is an IEA study that assesses the reading literacy of nine year old-students. The International Life Skills Survey (ILSS) follows after the International Adult Literacy Study (IALS) and the Second International Adult Literacy Study (SIALS); it covers adult literacy and is supported by Statistics Canada and the U.S. National Center for Education Statistics.
From studying reading and literature towards assessing literacy in a broader sense

2. In the 1971 IEA Six Subject Study, two separate studies on reading skills were included: a study on reading comprehension on the one hand and a study on the comprehension and interpretation of literary texts (literature education) on the other. While 10-year-olds were exclusively assessed on reading comprehension, 14-year-olds were assessed on both comprehension and literature. The dividing line between those two domains fits the outline of those categories to which later studies refer to as aspects (PISA) or processes (PIRLS) in their frameworks: the study on comprehension focused on cognitive aspects while the literature study gave more emphasis to aesthetic and affective aspects and to the reflective or critical dimensions of reading. The IEA study of reading comprehension was founded on the implicit model of a reader who gets meaning from the text and restores it correctly, while the IEA literary study reflected the position of the “response to literature” movement, which defines response “as the ongoing interaction between the individual and the work, an interaction that may continue long after the individual has finished reading” (Purves, 1973, p. 36). Basically, the domains covered by those two studies together are not very different in spirit from what PISA or PIRLS will assess in the year 2000. However, from the formal angle, a distinction was made between two study subjects back in 1971.

3. With regard to the actual scope of the reading domain assessed, one could group the IEA Six Subject studies taken jointly, PISA and PIRLS on the one hand, because they assess a broader range of aspects, and group the IEA Reading Literacy Study (RLS) and the International Adult Literacy Study (IALS) on the other, since the latter focus on a narrower definition of reading.

4. The IEA RLS in 1991 and IALS in 1994 were the first international reading assessments to make extensive reference to the concept of “literacy.” While the IEA Six Subject Study was predominantly underpinned by references to school and curriculum, the IEA RLS study was open to the concept of functional literacy (with intentional reference to the issue of illiteracy) and adopted a definition of “reading literacy” that aimed to reconcile societal stakes with the more traditional values of school and/or leisure reading: “Reading literacy is the ability to understand and use those written language forms that are required by society and/or valued by the individual” (Elley, 1994, p.5). This new trend of resorting to the concept of literacy coincided with the massive emergence of documents (non-continuous texts) as stimuli in the assessment tasks (e.g., maps, graphs, tables and forms—the types of text one needs to be able to read in order to cope with our society and in order to not be socially excluded), which did not occur in the 1971 assessment. However—and this may or may not be a consequence of making reference to the concept of “literacy,” which sometimes carries the notion of basic skills—the IEA RLS ends with comprehension and use of reading, without considering the more critical or reflexive aspects of reading.

5. The definition of literacy (only “literacy,” not “reading literacy” in this case) adopted in IALS was quite close to that of the IEA RLS, despite the fact that the focus was much rather on functional aspects of reading than on comprehension, which came as no surprise for a study targeted at adults. Literacy in IALS was defined as: “Using printed and written information to function in society, to achieve one’s goals, and to develop one’s knowledge and potential” (Murray, Kirsch and Jenkins, 1998, p. 17). The test developers intended to break with the simplistic vision of literacy, which traditionally distinguishes those who possess reading skills (the so-called “literate”) from those who do not (the so called “illiterates”). They stress the multidimensional aspect of reading by defining three separate domains of literacy: prose literacy, document literacy and quantitative literacy.

---

4 “The ability to read is denied to one in four of the earth’s adult population. (...) Furthermore in an increasingly complex, information-ridden world, demands on literacy continues to rise in all nations. Illiterates are at serious disadvantage in most countries today.” (Elley, Schleicher, Wagemacher, 1994, p.1).

5 The definitions adopted by the successive reading assessment surveys are presented in Table 2 at the end of this chapter with a view to facilitating their comparison.
6. PISA initiated a new tradition in the field of international reading assessments which was also adopted by PIRLS. Prior to the design and development of the test stimuli and items themselves, a detailed framework was drawn up, including comprehensive explanations relevant to theoretical references, the domain organisation, the methods of assessment, reporting scales and the interpretative scheme. These frameworks were developed by expert panels and submitted for review to the national authorities (OECD, 1999; IEA/PIRLS, 1999). It is likely that previous studies drew up that type of working document for internal use. However, the open discourse of such documents in the scientific community has undeniably led to increased efforts in planning and clarification. Current and future studies will be far more theoretically driven than the previous ones.

7. The definitions of reading literacy developed by PISA and in similar ways adopted by PIRLS clearly reflect higher awareness of recent developments in basic research on reading (e.g., referring to an interactive concept of reading). In PISA, for instance, reading literacy is "understanding, using and reflecting on written texts, in order to achieve one’s goals, to develop one’s knowledge and potential and to participate in society" (OECD, 1999, p. 20). Besides comprehension and use, the focus is on the reader’s capacity to reflect on written information as an important part of constructing meaning from text. In PISA, reading is envisaged as an interactive process between the text, the reader and the context: the reader, in order to achieve his or her reading purposes, implements strategies and draws on previous knowledge to construct meaning and to respond to the text. In the same line of thought, PIRLS suggested the following definition in its provisional draft framework of November, 1999: "Reading literacy is the ability to understand and use those written language forms required by society and valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers and for enjoyment" (PIRLS/IEA, 1999, p. 2). Continuity with the 1991 IEA RLS is evident, although one noticeable difference stands out. The reference to those theories that regard reading as an interactive process of constructing meaning is assertive and will be given concrete expression in the assessment of more elaborate processes, (e.g., “to draw and justify complex inferences and interpretations” or “to examine and evaluate content, language, and textual elements”). These ambitious definitions are far from the narrow, functional concept of literacy. Both recent assessments give definitions of "reading literacy" that, in fact, tally with what the experts and persons in charge of IEA in 1971 ascribed to "literary education or culture". It may be useful to point out that cultural literacy was not assessed for all students at the time, as is the case in PISA or in PIRLS today. The assessment of literature (in the 1971 IEA study) concerned a more or less selective sample of students in academic tracks and more generally speaking, still attending school (at that time, higher proportions of students had already left school at the age of fifteen). At present however, the highest requirements apply to all of the students in PISA or PIRLS, underpinned by a two-fold reference: a theoretical reference which sees comprehension as the outcome of a construction process and of an interaction between the text and the reader’s previous knowledge; and a reference to the requirements of society, which are constantly rising.

Domain organisation and characteristics of the reading tasks

8. As mentioned, documentation on the conceptual development process by functional expert groups or “steering committees” has not always been as widely distributed as it is nowadays. As a matter of fact, for the earliest assessments one often needs to refer to the introductions of reports in order to retrace the main lines of domain organisation and task characteristics. PIRLS is in its development stage and the information contained in the Draft Framework is still subject to revision. However, it seems unlikely that possible amendments would significantly affect comments relevant to this study. Therefore, we chose to include this study in our review of reading assessments. The definitions adopted by PISA and PIRLS are closer to what some would refer to as “cultural, advanced or high literacy” (Harris and Hodges, 1995). The fact that the frameworks of the different studies were released in a variety of ways and the uneven involvement of the author of the present review should prompt one to a certain degree of cautiousness. One cannot rule out, that in her attempt to achieve a historical reconstitution of the surveys’ preparatory phases, she might have overlooked some elements.
Types of text

9. The major feature in terms of domain organisation across the various studies is the reference to types of text or to a variety of written material. In the 1971 assessment of reading comprehension, Thorndike mentioned that “written material of various styles and content areas” was used (e.g., expository texts, literary texts, prose or poetry). However, at that time this suggestion was no more than a safeguard aimed at ensuring some diversity in the stimuli. In fact, no specific balance was required a priori for the construction of the assessment tasks. In 1991, in both the IEA RLS and IALS, the first element of organisation were the “domains” that relate to broad categories of written material: narrative prose, expository prose and documents for the IEA RLS; prose literacy, document literacy and quantitative literacy for IALS. In PISA, and later in PIRLS, the reference made to the type and format of written material remains essential, although it may at times follow a different itinerary. In PISA, a first distinction is made between continuous and non-continuous texts, further broken down into specific typologies (following to theoretical work on text typology). A proportion of units and/or items is established beforehand for every type of text listed, and the test developers were asked to follow the given ratio in as far as possible. While there is no reference to text typology in PIRLS, one of the two main lines of its domain organisation (i.e., the “purposes of reading”) outlines similar divisions to those based on the type of text. Under “reading for literary experience” and “reading to acquire and use information,” one could identify the main types of text listed above. Furthermore, the authors specify that “the PIRLS assessment of reading literacy will focus on the two types of texts that reflect the purposes of reading most often engaged in by students in and out of school” (IEA/PIRLS, 1999, p. 10). Like in PISA, a previously determined proportion of tasks will relate to those reading purposes. Besides, most surveys since the IEA RLS have reported or considered reporting the results on separate scales according to the types of text. This applied to the IEA RLS, IALS and PIRLS. Only PISA is exploring a different strategy which places the main emphasis on the reading processes rather than on the type of the stimulus material.

Skills, abilities, processes or aspects

10. We have seen that the first main line of domain organisation basically focuses on the nature of the texts and stimuli (the “text” pole in the text-reader-context threesome). Hence, it appears quite logical that the second main line would pertain to the nature of the abilities assessed. What aspects, what processes and what abilities will the assessment’s coverage include? While the terminology and concepts may have varied over time, all assessments consider a more or less explicit domain organisation based on the reader’s skills.

11. As mentioned earlier, in 1971 IEA made this distinction by organising two separate studies for reading comprehension and literary interpretation. However, their principles of organisation differ. For the study of reading comprehension, the author of the main report (Thorndike, 1973) states that “it was decided to focus on the cognitive content of the passage and to forego most efforts to get any appraisal of style, feeling, tone or literary techniques” (p. 19). Once this principle was set, the discussion on which skills to assess evolved towards the options of assessing: reading comprehension, reading speed, and knowledge of the vocabulary. The existence of processes in reading comprehension is not mentioned in the rationale of the assessment, even though a careful scrutiny of the items enables one to identify that these mobilise processes or aspects that differ in nature and in complexity (e.g., some reading tasks require identifying the main idea of a passage or drawing inferences). This assessment is illustrative of the state of advancement of research in reading theory at the time, bearing the imprint of a behaviourist concept. That is, it was centred on prerequisites and skills associated with reading comprehension (vocabulary and speed), while comprehension itself was still a sort of black box. Indeed, that study precedes the tremendous development of cognitive models in reading from the 1980s (Barr et al., 1990), which mainly undertook to find out what actually is inside the black box. In the late sixties, the widely diversified items intuitively captured a variety of processes, but the theoretical references needed to plan to capture this diversity a priori did not exist at the time.
12. Conversely, the literary study was founded on a very elaborate rationale. The design of the test was based on a cross-classification grid which intersected four categories of content (literary texts, contextual information, terminology and theory of literature, and information related to mythology) and ten different behaviours ranging from “applying knowledge of specific literary texts” or “developing a coherent preference pattern for...” to “responding to the text...” or “applying a cultural reference to...” At the time, the theoretical framework of literary criticism was well structured, and Rosenblatt’s first research on the transactional theory of literary work, a founding reference in that field, dates back to the thirties (Rosenblatt, 1938).

13. The 1991 IEA RLS was a decisive turning point in reading assessment. According to Elley’s report (1994), it appears that the issue of skills (or mental processes) had not been discussed initially. It did not stand out in the development stages of the test or during the process of defining its domains. The issue was only discussed when it came to developing and describing the final assessment material. The way in which those mental processes were determined still seems to owe a great deal to intuition and empirical approaches. No direct reference to a theory was presented. The Steering Committee classified the items by processes and submitted this classification to the review of 27 experts from 12 different countries. The processes differ, depending on the type of text. For narrative and expository texts, a distinction was made between “verbatim” response and “paraphrase”, between questions about the main idea and inference questions. For documents, two different aspects were defined: locating information on the one hand, and locating and processing information on the other. In comparison with the 1971 studies, intuition had gained in accuracy, even though it does not appear to actually have guided the test developers in designing the assessment. Arguably, one could consider that the need to classify the items by mental processes clearly arose when the option of presenting international results on IRT (Item Response Theory) scales came about.

14. With IALS, the link between theoretical models and assessment was firmly established, and it must be pointed out that this link was reciprocal. On the one hand, the assessments were underpinned by the findings of theoretical research and, on the other, data collected within the framework of the assessment supported further basic research (e.g., factors that account for the difficulty in reading tasks). Kirsch and Mosenthal’s work based on YALS (Young Adult Literacy Study) and the IEA RLS, in particular with a view to modeling factors that account for the relative difficulty in reading tasks, was eventually reinvested in IALS and in other studies. In this way, although the PISA evaluation programme covers a broader domain of literacy than IALS did, the relation of IALS to PISA is evident from the theoretical angle. The processes’ variables, which Kirsch et al. isolated as highly relevant predictors of task difficulty (type of match, plausibility of distractors and type of information), actually became “micro-aspects” in PISA, yet it is clear that the same set of references is at play.

15. The PISA survey illustrates the outcome of this long history, which began with an ex post facto empirical classification and wound up in developing a framework that defines at a highly detailed level...
what domains will be assessed. PISA provides for five macro-aspects\textsuperscript{14} to be assessed: broad understanding of a text, retrieving information, developing an interpretation, reflecting on the content of a text and reflecting on the form of a text. For each of these aspects, the proportion of items was determined before the tests were designed. At the same time, each item was classified according to the macro-aspect and sometimes according to the micro-aspect it refers to. Those two criteria determine what is referred to as the “question intent.” A given item is, for instance, classified as a “retrieving information: synonymous match” item. The need to specify each item in this way and to provide for a given numerical balance between the different categories is closely linked with the issue of IRT scaling, in particular with the prospective number and nature of interpretative scales on which results will be reported. This is a safeguard to prevent that at a later stage one would find out that a key dimension—for which one might wish to report on a specific scale—is not represented in a sufficient number of items or that the items covering that particular dimension do not vary in range of difficulty. Therefore, one can argue that the emergence of IRT scaling forced a transition from intuition or empirical approach to a more theoretical approach and the technically more complex quest of an appropriate balance.

**Design of the assessment and reporting scales**

16. The implementation of IRT analyses (Rasch, 1980) in large-scale comparative assessments can undoubtedly be regarded as one of the major breakthroughs in the last twenty years. As described earlier, the recourse to those scales certainly contributed to the introduction of a stage of theoretical conceptualisation prior to the design of the assessments. Moreover, the fact that such techniques exist makes it possible to resort to complex matrix sampling (with rotating test booklets). More items can be used; more abilities can be assessed without increasing the burden upon examinees. The various items are distributed in booklets that have a common section (e.g., 25 items out of 60) and a variable section. The anchor test technique makes it possible to compare the scores of students who did not respond to entirely identical sets of items.

17. The principle of IRT scaling models makes it possible to characterise the difficulty of an item and the level of proficiency of an individual on a common scale. Assuming that an individual obtains a score with value “n” on a given scale, means that this individual has a probability (which is higher than a threshold usually set at 50 or 80 percent) of performing a task correctly if its level of difficulty is lower than “n.” Conversely, the same respondent’s probability of correctly performing a task for which the level of difficulty is higher than “n” will lie below the set threshold. As Kirsch, Mosenthal and Jungeblut (1998) put it, “item response theory (IRT) scaling provides a way to focus both on people and tasks together” (p. 105). IRT scaling makes it possible to go beyond the strictly normative perspective, which merely consists in ranking individuals (or countries, in international assessments) from the highest to the lowest performance. Before the introduction of IRT scaling, there was only room for a qualitative analysis once the ranking was over. Since IRT models are used, the information can be refined and becomes increasingly diagnostic. Not only does a country become aware of whether its performance ranks higher or lower than that of another one, it can know what proficiency levels are reached by its students. In a certain way, these models succeed rather elegantly in combining a normative approach (which remains at the core of any comparative study) with the diagnostic refinement of a criterion-based approach, since the latter cannot be implemented as such within the framework of an international assessment.

**Question format**

18. Question format was certainly one of the issues that proved to be of great concern to expert groups or groups of national representatives in international assessments: open-ended questions or

\textsuperscript{14} They cannot really be defined as “processes” as such, since several processes of a different nature can be mobilised to perform a single reading task. For example, if the point is to identify the main idea of a text, this main idea may be quoted in the text and the respondent will just need to locate it while, in other cases, it may be necessary to draw a simple or even a complex inference.
multiple-choice items, closed questions or constructed response items? To summarise this issue in an oversimplified way, one could say that the advantage of closed questions and multiple choice items is that they can be rated very quickly, are quite reliable and cost-effective. On the other hand, they are inadequate for the evaluation of a number of aspects (critical reflection, notably). Open-ended questions and constructed response items are judged to be more authentic, they are indispensable in order to assess certain aspects, but they also require working out standard marking guides, which are necessarily expensive to implement. For the latter, comparability of results can turn out to be less reliable (cf. inter-rater reliability). Besides, if an item calls for an elaborate response, the assessment of reading proficiency partly tends to merge with the writing skills of the student. Let us examine how the successive assessments eventually settled this highly controversial issue.

19. In the 1971 IEA study, all efforts were concentrated on demonstrating that such a comparative survey was indeed feasible and therefore the choice of item types remained rather conservative: “the reading comprehension test was to be of the conventional type in which a passage is presented to the pupil together with multiple-choice questions based on that passage” (Thorndike, 1973, p.20). For the literature study, alternative formats were used for the field trial, which included both multiple choice items and open-ended questions, particularly with a view to measuring the impact of question format and improving distractors. Since the results of the field trial suggested that “multiple-choice items per se do not measure anything different from what is measured by open-ended questions on the same topic” (Purves, 1973, p. 67), the main test eventually confined itself to using multiple choice items only. This choice, motivated by caution, was consistent with the reading model that inspired those evaluations. Indeed, for every question asked, there was only one correct answer. This was unambiguously echoed by the way the instructions to the students were worded and therefore, from the theoretical standpoint, nothing spoke against proposing only multiple choice items.

20. In 1991, the IEA RLS repeated those choices with very few deviations. Some national representatives asked for more open-ended questions (constructed response items), but the Steering Committee chose to proceed in a conservative manner. To support their position, the Steering Committee in part referred to a study (Elley and Mangubhai, 1992) in which apparently similar results were obtained with two different question formats. As a consequence, the assessments consisted of a majority of multiple choice items (from 90 per cent for population A to 75 per cent for population B). Ten to twenty percent of the items were open-ended questions to which the answer consisted in a single word, a figure or a few words (i.e., items that do not give rise to any rating difficulties). Two constructed response items were included as an international option at each level, and the results of this option were analysed in Kapinus and Atash (1995).

21. With IALS, a very different approach was taken: all the items were open-ended questions. This choice was, at first sight, motivated by psychological considerations much rather than psychometric or theoretical ones: “all of the literacy tasks were open-ended rather than multiple-choice, because it was thought that adults would be more interested in performing such tasks” (Murray, Kirsch and Jenkins, 1994, p. 19).

22. In PISA and PIRLS, multiple choice items, simple open-ended questions (i.e., one rater is enough, or simple marking) and open-ended questions with constructed responses (e.g., several raters are required, or multiple marking) are used. However, the distribution of item types differs between the two studies—in PIRLS approximately one-third of the items are multiple choice, whereas in PISA about half are multiple choice. In both cases, the choice does not solely rest on psychological reasons or “face validity” (i.e., to make the tasks more “authentic”). It is also guided by theoretical considerations. Some of

---

15 In fact, this seemingly technical question raises the far more fundamental issue of the link with school practices and the possible effect of an assessment that uses methods that do not conform to the line of the school or the line that decision-makers in education would like to see the teachers follow.

16 Of course, everything depends on the criteria by which one chooses to measure similarity. Other studies (Kapinus and Atash, 1995) resulted in diverging or, at any rate, more qualified answers to the same question.
the aspects or processes assessed assume that the reader will bring his or her previous knowledge into relation with the text. In other words, several “correct” responses can be envisaged and it is very difficult to cast that type of question, which draws on the reader’s creativity to some extent, into the mold of multiple choice questions. However, it would be wrong to assume that constructed response items would exclusively assess the more complex processes or aspects and the dimensions of question format and processes would merge. That assumption, indeed, would give rise to serious difficulties of interpretation.

23. The transition from a test in which multiple choice questions prevailed to the need for the respondent to actually write out his or her answers clearly modifies the perspective. Issue of motivation of respondents as well as issues relating to the context of the test administration arise (i.e., the cultural context in the broad sense of the term and the specific context of a given classroom). Research may be necessary to shed more light on those issues which were not resolved by previous relevant studies. The PISA field trial data analysis shows, for instance, that there is an interaction between question format and gender. Items that require a long response are easier for girls and short answer and multiple choice items appear easier for boys. Moreover, this finding is stronger for science than for the other domains.

24. When a high proportion of constructed response items becomes one of the cornerstones of a large-scale evaluation this has several obvious effects: high costs, the need for strict monitoring of inter-rater reliability both within countries and across countries (this was implemented after the test for IALS and is being considered for the PISA main test), and the challenge of working out standardised marking guides. The latter should both include generic rating principles and a sufficient number of example responses, and avoid placing raters in situations of cognitive overload. Another way explored nowadays is double-digit coding, which was already considered in TIMSS (the IEA Third International Mathematics and Science Study) and which will also be implemented for coding the mathematics items in PISA. The principle of double-digit coding is to try to capture, by means of a second coding operation, the path followed by the student, the existence of representations, of partial or erroneous concepts, and of efficient or inadequate approaches. This extends beyond grading the quality/correctness of the response with a higher or lower mark. To take up that challenge would make even richer and more relevant the information that one could draw from the assessment, especially for teachers.

Cultural and linguistic equivalence and comparability: At stake in the past, at stake in the future

25. Even more than in other subject matters, the issue of the comparability of reading assessments recurs like a leitmotiv in literacy assessments. In reading, more than in other subject matter areas, task and passage difficulty prove to be linked to the written material, which can possibly be altered during the course of a translation process. The students’ uneven cultural familiarity with certain types of text or content also is a controversial issue. With regard to those crucial questions, it is quite informative to take a closer look at the devices set up to optimise the comparability of data, possibly with a view to drawing useful lessons from those procedures.

26. In all comparative reading assessments since 1971, participating countries were invited to contribute to the process of building up test material of diversified linguistic and cultural backgrounds based on the principle that test material should be as “universal” as possible in order to avoid biases due to gender, ethnic or linguistic peculiarities. In all the international assessments, the national project managers and their relevant expert committees also were invited to react more or less systematically to the stimuli proposed for the test and to judge whether these could or could not give rise to any particular problems in their national contexts. However, it is difficult to evaluate if this call for “national” contributions actually had a noticeable outcome. There would not be enough data available to assess this. In the IEA RLS, 20 countries submitted material for the assessment, although we do not know how many of the units proposed by different countries were finally retained. In IALS, approximately half of the units originated from North America and the other half from the European participants. In PISA, more countries are involved in the study and so the diversity of origins of texts is greater. However, texts that originate from English-speaking countries are still in the majority [52 per cent of the reading items originated in English (contributed mainly by Australia, Canada, Ireland, the United Kingdom and the United States), 13 per cent in French, 10 per cent of the items originated from IALS, 10 per cent from Swedish, 4 per cent from
German and Swedish each, and the remaining items originated from other languages. Whatever one’s position may be on this issue, the predominance of English speaking materials definitely remains an objective component.

27. Besides, in every case except PISA, proposed texts were to be submitted in English, even if the original language was not English. (In PISA, the units could officially be submitted in English or French and the consortium also accepted units in Dutch, German and other languages). If, for example, a Swedish text is retained and presented to a French-speaking student, it will have undergone two consecutive translations before the student reads it. In the absence of additional control or supplementing verification procedures, one could legitimately challenge the quality of the text, which that student sees, as well as its equivalence with regard to the initial text. The mechanism designed by ILSS (International Life Skills Survey) in order to ensure a greater diversity in the origins of the texts and documents is quite innovative in several regards. Several networks of countries, grouped according to their linguistic or cultural kinship, are invited to interact and make arrangements with the aim to produce a significant amount of test material, items and marking guides in their language of origin (e.g., French, German, Spanish, Portuguese). The material will only be translated and submitted for arbitration by the international centre (Educational Testing Service) after this transnational consultation.

28. Beyond the question of cultural diversity, what control procedures have been set up to ensure optimal equivalence of the reading material? We shall restrict ourselves to a brief outline of the major characteristics of these procedures; because a comprehensive description would be too detailed for our purposes.

29. In 1971, Thorndike’s report on the reading comprehension study did not show too much concern about procedures. The Steering Committee was rather optimistic and relied on common sense, stating that “it was felt that with some care, the test could be maintained as nearly enough the same task from one language to another to make the cross-national comparisons interesting and fruitful” (Thorndike, 1973, p. 15).

30. The experts in charge of the study on literature education were less pragmatic and considered several stages of quality control for the translation. The people in charge at the national level appointed professional translators to verify the translation of passages, items and attitude scales and to perform a back translation. The translations and back translations were then returned to the chairman, Alan Purves, who had them verified by his colleagues from the language departments of the University of Illinois. Purves claimed, “in several cases, minor revisions were made between the pre-test and the final forms, but in general, the translations were accurate and literarily appropriate.” (Purves, 1973, p. 75).

31. The IEA RLS supplied a set of guidelines for the translation of test material, which had to be undertaken by the national project managers. This document recommended that two independent, bilingual individuals make two translations and that a back translation be forwarded together with a report on the translation procedure. Post hoc statistical checks were carried out to see if there were any systematic differences owing to different versions. (Researchers checked the correlation of the mean item difficulty index of each country with the international mean item difficulty index and the mean difficulty index in the English-speaking countries, since English was the source language for that test.)

32. The IALS study made another important step forward. Like the IEA RLS, IALS provided for a set of translation guidelines but, in addition, “each country’s test booklets were carefully reviewed for errors of adaptation. Countries were required to correct all the errors found.” (Murray, Kirsch and Jenkins, 1994, p.77). However, the procedure was often not properly implemented and the countries did not always feel obliged to correct the errors pointed out to them, often believing that they “knew better.” In the end, the translation and a number of methodological aspects wound up to be the reported cause of France withdrawing from the study altogether after having administered the tests. This event triggered an intense reflection process around all aspects relevant to ensuring data comparability, both on behalf of IALS and among its critics, and it can be said that this was of great benefit for the international assessments to come.
33. If one takes a closer look at the PISA procedures\(^\text{17}\) for example, the verification of translation equivalence has acquired the dimension of one of the project’s great strengths, with high quality standards. As usually required for official OECD documents, the PISA assessment material is available in two different source languages\(^\text{18}\) and it is recommended that the countries use both versions in the translation process (double translation). Professional translators who are trained using precise guidelines verify the double translation. A resource expert arbitrates cases of dispute. All national adaptations are closely verified. Finally, a Cultural Review Panel arbitrates open issues related to equivalence. In a nutshell, the sieve is much finer, the verification net is drawn tighter and one has every reason to believe that, as far as translation equivalence is concerned, nothing will ever be like it used to be. In the same perspective, a document that was recently issued by the IEA, the *Technical Standards for IEA Studies* (Martin, Rust, Adams, 1991), sets very strict quality control standards for the translations: double translation followed by a “reconciling” or merging procedure, quality control of the translations to be performed by the international project centre, strict observance of well-defined adaptation rules and, above all, verification as to the extent to which the countries actually follow all recommendations. However, notwithstanding the rigour of the mechanism set up for PISA, reluctant observations are still leveled at several aspects of the new procedures, leading us to surmise that the issue has not yet been resolved to the full satisfaction of all parties. The verification of translation equivalence and task or item difficulty equivalence from one language to another will remain a major, time-consuming concern for international centres in charge of reading assessments.

34. Objections also are raised against a main premise on which international comparisons in reading assessments implicitly rest. It is surmised that assessments measure the same latent features everywhere, no matter what the language, the culture, or the country may be; and that the processes of reading literacy are universal and can therefore be accessed through a comparative analysis. However, not everybody sees it that way. A European project currently examines the possibility of evaluating the skills and abilities of 14-15 year old-students by means of administering tests in their original language, which would not be translated but judged by experts to feature equivalent task difficulty levels. The key challenge in this context obviously lies in that last step: will it be easier or harder to agree on sets of objective criteria that account for text/item difficulty than to ensure the production of equivalent translations? At any rate, the results of that study will undoubtedly give rise to most fruitful discussions.

**Main challenges for the future**

35. This paper on thirty years of reading assessment has highlighted some of the main lines along which reading assessments have developed\(^\text{19}\).

36. In thirty years many of the challenges of such assessments have remained. The objective to achieve comparability in the data collected and to produce performance indicators which are accepted by policy-makers and educators has not yet been completely achieved. Resistance against comparative assessment is still strong in some countries. Nevertheless, the responses and technical solutions to address these challenges have significantly advanced since the seventies:

- The conceptions of what reading comprehension and reading literacy actually involves have expanded considerably, in parallel with progress achieved in cognitive research on reading processes\(^\text{20}\).

---

\(^{17}\) As described in the *Report on the Implementation of the PISA Translation Procedures*, DEELSA/PISA/BPC(99)16, drawn up for the BPC meeting (Board of Participating Countries) held on October 4 and 5, 1999.

\(^{18}\) Several experts verify the comparability of the two source versions.

\(^{19}\) The main characteristics of the surveys examined are presented in a synoptic table at the end of this paper.

\(^{20}\) The relevant definition series are presented in a table at the end of this paper in order to ease their comparison.
– The process of building assessment instruments on a strong theoretical foundation has considerably developed under the two-fold influence of reading research (including cognitive aspects, text typologies, research on task difficulty predictors) and psychometrics (e.g., IRT scaling). The empirical approach followed in the early studies has gradually given way to a well-structured and planned approach in which different theories, constraints and criteria are considered and balanced before the instruments are developed.

– The use of more complex assessment instruments, in which different groups of students receive different sets of items which are then anchored through psychometric procedures, has allowed to significantly increase the coverage of the reading assessments.

– The emergence of scaling procedures based on item response theory has revolutionised methods of presenting and interpreting survey results, including the reporting of results on sub-scales and the establishment proficiency levels.

– Issues relating to the cultural relevance and equivalence of assessments are receiving increasing attention and, as a result, higher standards and quality assurance procedures are sought to address these issues.

37. These different aspects of reading assessments are closely interrelated. They do not develop independently, and it is of course difficult to predict what new orientations are likely to guide reading literacy assessments in the coming years. Clearly, the emergence of theoretical and psychometric models is a powerful factor of change. However, as Kirsch put it in his introduction to the technical report on IALS, to ensure the quality of all the assessment’s aspects remains a crucial need and generates progress by itself: “in retrospect, more attention should have been paid to other [editor’s note: other than psychometric] design elements (...); it is now necessary to refine the other data quality-related aspects of the study in future survey cycles.” (Kirsch, p. 15 and p. 22).

38. In this respect, the qualitative improvement achieved by PISA in the field of quality control of translations will have to be maintained and even intensified. Notwithstanding the efforts made, it appears that some countries are not fully reassured by the procedures implemented or are not fully supportive of these. Also the methods suggested by ILSS (i.e., networking with countries that feature similar languages or cultures and developing more diversified test material, based on a somewhat broader consensus than a purely national one) appears to be quite promising.

39. Furthermore, the increasing prevalence of open-ended items with constructed responses leads to the need to improve marking procedures. In this perspective, the PISA Field Trial, which took place in 1999, contained a high proportion of constructed response items, in compliance with the wish expressed by the OECD Board of Participating Countries. A control procedure was implemented for assessing inter-rater reliability which showed mixed results. In spite of the considerable efforts that have been made to produce reliable marking guides and to train raters in the countries, improvements could still be made: “The coding process is more or less reliable and the countries are more or less reliable” (ACER, 1999, p. 180). In response to these weaknesses, the PISA consortium introduced multiple marking and cross-country coding procedures. “The field trial showed the methodological need to include a reliability study during the main study. It also showed between-country variability, and suggesting that during the main study, a cross-country reliability study will be necessary.” (ACER, 1999, p. 180).

40. Finally, at the dawn of the third millennium, the need to incorporate or associate an assessment of proficiency in reading electronic texts, documents and messages into the scope of reading literacy can no longer be left aside. PISA took the well-considered decision of postponing that aspect until later cycles. PIRLS envisages undertaking a research study and testing a subsample of students using an electronic medium (i.e., by presenting one of the PIRLS test booklets in electronic format) and assessing students’ ability to use the Internet. Network A is currently investigating the possibility of developing an assessment of students’ information and communication technology literacy within the framework of upcoming PISA cycles. The question arises as to whether aspects such as reading texts presented in electronic format, surfing on the Internet, or joint processing of information packages presented as text and images, will become an integral part of reading literacy, of computer literacy or of further literacies that still have to be
discussed. The future will tell but, no matter what the outcome may be, relevant discussions will undoubtedly give rise to an extensive restructuring process of the whole concept of reading literacy, in a world where the electronic medium gradually takes over from printed material.

This chapter was prepared with the financial support of the New Zealand Ministry of Education.
Exhibit I: Synthesis of the Main Features of International Reading/Literacy Surveys

<table>
<thead>
<tr>
<th>Studies/Commissioners</th>
<th>Scope</th>
<th>Population</th>
<th>Domains</th>
<th>Skills/Processes</th>
<th>Scales</th>
<th>Design</th>
<th>Question format</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEA Six Subject Study 1971</td>
<td>Reading Comprehension</td>
<td>10-11 years 14-15 years Last grade of Upper Secondary Education</td>
<td>No domains specified (&quot;variety of contents and styles&quot;) All tests are continuous (narrative and expository)</td>
<td>Reading comprehension/reading speed/vocabulary</td>
<td>…</td>
<td>No rotation</td>
<td>100% multiple choice items</td>
</tr>
<tr>
<td>IEA RLS (1991)</td>
<td>Reading Literacy</td>
<td>9-10 years 14-15 years</td>
<td>Narrative prose/Expository prose/Documents</td>
<td>Verbatim/Paraphrase/Main idea/Inference/Locate/Locate and process</td>
<td>3 scales: narrative/expository/documents</td>
<td>No rotation</td>
<td>75% to 90% multiple choice items + some &quot;closed&quot; open ended</td>
</tr>
<tr>
<td>IALS (1994) OECD and Statistics Canada</td>
<td>Literacy</td>
<td>Adults</td>
<td>Prose/Document/Quantitative</td>
<td>Retrieve information/Broad understanding/Develop an interpretation/Reflecting on the content/Reflecting on the form</td>
<td>Not yet determined</td>
<td>Complex matrix sampling</td>
<td>100% open ended</td>
</tr>
<tr>
<td>PISA (2000)/OECD</td>
<td>Reading Literacy</td>
<td>15 years</td>
<td>Continuous/non-continuous texts</td>
<td>Focus and retrieve explicitly stated information/Make straightforward inferences/Draw and justify complex inferences and interpretations/Examine and evaluate content, language and textual elements</td>
<td>2 scales: purposes of reading (literacy vs. information purpose)</td>
<td>Complex matrix sampling</td>
<td>55% multiple choice items, 45% open-ended questions with constructed response</td>
</tr>
<tr>
<td>PIRLS (2001)/IEA</td>
<td>Reading Literacy</td>
<td>9-10 years</td>
<td>Purposes of reading (reading for literacy purpose vs. reading to acquire and use information purpose)</td>
<td></td>
<td></td>
<td>Complex matrix sampling</td>
<td>70% constructed response items, 30% multiple choice items</td>
</tr>
</tbody>
</table>
## Exhibit II: The Definitions of Reading or Literacy in the International Surveys

<table>
<thead>
<tr>
<th>Surveys</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEA (1971)</td>
<td>Reading comprehension, without a more precise definition.</td>
</tr>
<tr>
<td>IEARL (1991)</td>
<td>“Reading literacy is the ability to understand and use those written language forms that are required by society and/or valued by the individual.”</td>
</tr>
<tr>
<td>IALS (1994)</td>
<td>“Using printed and written information to function in society, to achieve one’s goals, and to develop one’s knowledge and potential.”</td>
</tr>
<tr>
<td>PISA (2000)</td>
<td>“Understanding, using and reflecting on written texts, in order to achieve one’s goals, to develop one’s knowledge and potential and to participate in society.”</td>
</tr>
<tr>
<td>PIRLS (2001)</td>
<td>“Reading literacy is the ability to understand and use those written language forms required by society and valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers and for enjoyment.”</td>
</tr>
</tbody>
</table>
References

1. International Reading Comprehension or Reading Literacy Assessments


2. Other references


DEFINITION AND SELECTION OF KEY COMPETENCIES

A contribution of the OECD Program Definition and Selection of Competencies: Theoretical and Conceptual Foundations

Dominique Simone Rychen, Swiss Federal Statistical Office, Neuchâtel, Switzerland
Laura Hersh Salganik, Education Statistics Services Institute, American Institutes for Research
Washington, D.C., United States

Acknowledgements
This contribution reports on work that has been undertaken by the DeSeCo Program (Defining and Selecting Competencies: Theoretical and Conceptual Foundations) under the leadership of Heinz Gilomen of the Swiss Federal Statistical Office (SFSO). The authors would like to express their thanks to the many individuals who generously have contributed ideas and support to the DeSeCo program since its inception in 1997.

Background

1. The importance of knowledge, skills, and competencies to individuals and society is widely accepted among policymakers in OECD countries. At least at the discourse level, a well-educated, knowledgeable, highly qualified citizenry is seen as playing an eminent role in facing the challenges of the present and the future in an increasingly interdependent, changing, and conflictual world.

2. To date, the major impetus in OECD countries for efforts in the area of key competencies has come from the business sector and from employers. From a purely economic viewpoint, competencies of individuals are seen as important because they contribute to:

   – boosting productivity and market competitiveness;
   – developing an adaptive and qualified labor force; and
   – creating an environment for innovation in a world dominated by global competition.

3. From a broader social perspective, knowledge, skills, and competencies are important because they contribute to

   – increasing individual understanding of public policy issues and participation in democratic processes and institutions;
   – social cohesion and justice; and
   – strengthening human rights and autonomy as counterweights to increasing global inequality of opportunities and individual marginalization.
4. Partly due to a recognition that both human and social capital are important factors for the functioning of society and the economy, there has been in recent years an increased interest in information about knowledge, skills and competencies. Consequently, new indicators of education outcomes beyond the traditional measures (such as number of years of education or highest degree earned) are seen as relevant and important. This includes indicators not only of what individuals know and can do in school subjects but also of their capability to respond to the demands of everyday life.

5. If, in the 1980s, OECD countries adopted a pragmatic approach using available outcome-related data on education systems, in recent years they have moved towards a more systematic, long-term data collection strategy for education indicators programs. At the same time, OECD constituents have realized that additional theoretical and conceptual inputs would benefit the development of such programs and the interpretation of data. An analysis of previous competency-related large-scale studies within the OECD\textsuperscript{21} confirmed the absence of an explicit, overarching conceptual framework to guide the work.

6. The OECD Program, \textit{Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)}, was initiated to work towards filling this gap. Under the leadership of the Swiss Federal Statistical Office (SFSO) and with the support of the United States Department of Education National Center for Education Statistics (NCES), DeSeCo seeks, through an interdisciplinary approach in collaboration with ongoing OECD assessment programs, to

- advance the theoretical foundation of key competencies;
- provide a reference point for indicator development and interpretation of empirical results;
- encourage an iterative process between theoretical and empirical work; and
- respond, eventually, to information needs of policymakers.

\textbf{Guiding Research Questions}

7. The DeSeCo Program\textsuperscript{22} is concerned with a number of wide-ranging issues relating to competency determination. The following list provides a few examples of the kinds of research questions developed to guide the program’s activities:

1. What is meant by the notions of competence, key competencies, skills, etc.? How can these terms be conceptualized and described?

2. Which notions concerning the nature of human beings and society should serve as a starting point for the definition and selection of key competencies? Why do we need competencies? What are the premises for the adequate functioning in society and the attainment of so-called successful life from the perspective of both the individual and society?

3. How can the perspectives of different academic disciplines contribute to the understanding and construction of a set of key competencies? What are the theoretical models, concepts and arguments put forward by these various disciplines?

4. Which key competencies are necessary for understanding and acting in different fields of life – including economic, political, social, and family domains, public and private interpersonal relations, individual personal development, etc.? How can these competencies be described and theoretically justified, and what empirical evidence does available research provide of


\textsuperscript{22} Swiss Federal Statistical Office (1999), \textit{Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)}, September, 1999, Neuchâtel
their importance? Do key competencies operate independently, or should they be viewed as an interdependent set or constellation of competencies? In either case, how do the identified key competencies relate to each other?

5. To what extent are key competencies immutable with reference to social, economic, and cultural conditions? To what extent are they then generally valid, or to what extent does their importance vary by country or region according to particular socio-economic and cultural contexts?

6. To what extent is it possible to identify key competencies independent of age, gender, status, professional activity, etc.? Are certain competencies particularly important in the various phases of life, and if so, which ones? Do we need the same basic, universal key competencies when we are young, join the workforce, establish a family, advance in our professional or political career and retire?

7. What does this scientific discussion contribute to policy-making and practice? Do the perspectives and concepts developed match those resulting from experience in policy-making and practice? Or do representatives from different fields of social activity have completely different criteria and visions for what is important to them? Is it at all possible for a dialogue between theory and practice to develop?

8. What is the role of policy and practice in defining, selecting and describing skills as "key competencies"? What are the political, social, and economic factors that influence the definition and selection processes of key competencies in different socio-economic and cultural environments, and how is this influence exercised?

9. What is the role of social institutions in transmitting competencies to the population? What is the role of education, with its specific institutions and processes – schools and teaching – in the development of key competencies? What is the role of other potential sources for acquiring competencies, such as friends, parents, the working environment, media, religious and cultural organizations? To what extent can the transmission patterns be altered? In other words, to what extent can the transmission of key competencies be controlled by policies?

10. What is the relevance of the emerging ideas on key competencies for the development and interpretation of indicators designed to measure competencies among the population? How could these issues be addressed and developed by future research? What are the potential approaches for conceptualizing abstract competencies at a more concrete level?

8. As represented by Figure 1 below, the underlying hypothesis for this work is that any particular set of key competencies is the result of multiple factors. At the most fundamental level, the underlying vision of the world, including assumptions about society and individuals, affects the identification of key competencies. Different theoretical and conceptual perspectives also have a profound effect on which competencies are identified as key as well as how the problem is approached. Factors such as culture, gender, age, and social status influence the forms that abstract key competencies take in specific contexts. Finally, the definition and selection of key competencies is a result not only of scientific analysis but also of a political negotiation process and consensus formation. Each of these aspects should be taken into account when considering how key competencies are transmitted and developed, and when constructing and interpreting indicators.
Activities

9. The initial undertaking of the DeSeCo program was an analysis of previous competence-related work in the OECD context. The second activity was an analysis of existing theoretical and conceptual approaches to concepts of competence. The third activity, undertaken in 1999, was the production of five expert reports, each detailing a proposed set of competencies from a different theoretical background and academic discipline. The authors were to expound fully on their conceptualization of competence and


25 The authors are:
Monique Canto-Sperber, Centre National de la Recherche Scientifique, France and Jean-Pierre Dupuy Ecole Polytechnique, Centre de Recherche en Epistémologie Appliquée, France, representing a philosophical perspective;
Jack Goody, St. John’s College, University of Cambridge, United Kingdom, representing an anthropological perspective;
Helen Haste, University of Bath, United Kingdom, representing a psychological perspective;
Frank Levy, Massachusetts Institute of Technology, United States and Richard J. Murnane, Harvard University, United States, representing an economic perspective; and
Philippe Perrenoud, University of Geneva, Switzerland, representing a sociological perspective.
(cf. 5 DeSeCo Expert reports downloadable at www.deseco.admin.ch)
justify their approach theoretically, taking into account any available empirical research providing evidence for the importance of the proposed set of competencies. A commenting process followed, wherein the expert reports were distributed among the authors, other academics, and participating leading representatives from various fields (e.g., economic, social, educational, and cultural sectors). In October 1999, an international symposium was held in Neuchâtel, Switzerland, fostering an exchange of ideas between the authors of the scientific reports and other scientists, representatives of leading economic and social institutions, and policymakers. The symposium was the culmination of the program to date, as it was an opportunity to reflect on the DeSeCo Program’s aforementioned, first three main activities.

10. At this stage, DeSeCo is still focused on developing and refining a conceptual and theoretical approach to key competencies that draws on multiple disciplinary perspectives. In its future activities, it will continue the development work by focusing on

   a) the identification and description of key competencies with regard to important specific fields of activities,
   b) the mechanisms that operate nationally in defining and selecting key competencies, and
   c) the role of policy and practice in defining competencies.

11. These are all crucial steps towards determining the potential for developing a conceptualization of key competencies that can be used to guide development of indicators.

12. What follows is an attempt to delineate several common and relevant features among the different proposed approaches to defining and selecting key competencies. It represents a provisional, but significant step towards developing a common conceptual frame of reference for identifying key competencies.

**Normative assumptions – Visions of society and individuals**

13. Any conceptual or theoretical foundation for defining and selecting key competencies is inevitably influenced by conceptions about individuals and society and by what is valued in society and in life under particular socio-economic and political conditions. Today, we are confronted with important and complex challenges such as rapid social and technological changes, economic and cultural globalization, increasing uniformity, and at the same time, increasing social diversity, instability of norms, large-scale value changes, substantial global inequality of opportunities, increasing marginality of certain segments of the population, and ecological pressures. Given that there are many different social and

---

26 The commentators included:
Carlo Callieri, Confindustria, Italy;
Jacques Delors and Alexandra Draxler, Task Force on Education for the Twenty-first Century, UNESCO;
Jean-Patrick Farrugia, Le Mouvement des Entreprises de France (MEDEF), France;
Bob Harris, Education International;
Robert Kegan, Harvard University, United States;
George Psacharopoulos, University of Athens, Athens, Greece (formerly with the World Bank);
Cecilia Ridgeway, Stanford University, United States;
Laurell Ritchie, Canadian Auto Workers, Canada;
M. Boediono, Ministry of Education and Culture, Indonesia; and
Leonardo Vanella, Centro de Estudios e Investigacion del Desarrollo Infanto Juvenil, Argentina.
(cf. Comments on the DeSeCo Expert Opinions downloadable at www.deseco.admin.ch)

27 The analysis presented in this contribution draws heavily on the commissioned reports and comments.
individual responses to these changes, it is important to ask: to what kind of world do we aspire? Defining key competencies raises questions such as what for? in support of which objectives? according to which criteria?

**Key competencies are consistent with the principles of human rights and democratic values**

14. A number of international texts and agreements form a solid basis for describing the universal principles governing a modern, democratic society. This literature emphasizes, among other things, democratic values: respect for the law and for the rights of others; the importance of education for imparting knowledge, skills, and competencies to the population; and learning as a lifelong endeavor.

15. Defining key competencies in a manner that is consistent with these principles is a political and ethical challenge for countries and societies. For example, key competencies should be determined and developed in a way that is consistent with the values of autonomy and liberty. Yet at the same time, they should also be consistent with ideals of social justice so that they contribute to egalitarianism (rather than fostering the interests of an elite). The commitment to democratic values and human rights is reflected in the conceptualization of key competencies as being learned rather than innate, and as necessary for and attainable by everyone.

**Key competencies give individuals the capacity for a good, successful life**

16. Key competencies should reflect a notion of what constitutes a good and successful life for individuals beyond the satisfaction of elementary personal needs. Consistent with any major moral theory, a good and successful life includes close relationships with other people, an understanding of oneself and one’s world, autonomous interaction with one’s physical and social environment, and a sense of accomplishment and enjoyment. This broad and rich conception of a successful life represents an alternative to approaches that consider key competencies predominantly from the perspective of productivity and competitiveness.

**Key competencies are not incompatible with social and individual diversity.**

17. The concept of key competencies assumes that individuals and societies share some basic characteristics beneath the variety of approaches to life, life styles, and customs. For instance, the importance of social influences and the capacity for autonomous action are elements of the human condition, not dependent on any society or culture. Yet, social and individual diversity is an undeniable fact. Thus, it is important to construct key competencies at a relatively abstract level, recognizing that their development and application may take many forms depending on social and individual factors.

**The Concept of competence**

18. Generally speaking, concepts are socially constructed notions that facilitate the understanding of reality while also constructing it. Notions such as “key competencies” and “core skills” have become very fashionable in social policy discourse. However, these terms often have very vague meanings. Therefore, their clarification was considered a necessary prerequisite for defining and selecting key competencies. However, based on a recent, commissioned report, we recognize that in social sciences there is no unitary use of the concept of competence, no broadly accepted definition or unifying theory. In fact, the

---

28 e.g. UN Declaration of Human Rights
30 Weinert (1999), op. cit.
meaning of such terms varies largely depending on the scientific perspective and ideological viewpoints involved and on the underlying objectives associated with their use, both at scientific and political levels. As such, DeSeCo adopts a pragmatic conceptual approach, limiting the use of the concept with criteria which are more or less explicit, plausible, and scientifically acceptable.

Competencies are broader than knowledge and skills

19. DeSeCo focuses on a functional approach, which places complex demands facing individuals at the forefront of the concept of competence. According to this viewpoint, competencies are structured around demands and tasks. Fulfilling complex demands and tasks requires not only knowledge and skills but also involves strategies and routines needed to apply the knowledge and skills, as well as appropriate emotions and attitudes, and effective management of these components. Thus, the notion of competencies encompasses cognitive but also motivational, ethical, social, and behavioral components. It combines stable traits, learning outcomes (e.g., knowledge and skills), belief-value systems, habits, and other psychological features. In this view, basic reading, writing and calculating are skills that are critical components of numerous competencies.

20. While the concept of competence refers to the ability to meet demands of a high degree of complexity, and implies complex action systems, the term knowledge applies to facts or ideas acquired by study, investigation, observation, or experience and refers to a body of information that is understood. The term skill is used to designate the ability to use one's knowledge with relative ease to perform relatively simple tasks. We recognize that the line between competence and skill is somewhat blurry, but the conceptual difference between these terms is real.

Competencies are learned

21. Acquiring competencies is viewed as an on-going, lifelong, learning process. This process occurs in multiple settings. The settings and social institutions relevant for the development of competencies besides school are family, peers, work, political life, religious life, cultural life, etc. The conception of competencies as learned contrasts with one in which competencies are considered innate, inborn characteristics.

22. Competency learning is not only a matter of personal effort. The development of competencies assumes a favorable social and ecological environment, which includes but goes beyond the satisfaction of basic needs (food, housing, health, etc.). It is also dependent on the quantity and quality of learning opportunities. Therefore, the structure of the economy and social institutions plays an important role in the development of competencies.

Identifying key competencies

23. This section briefly summarizes the different viewpoints of the expert authors who were asked to identify key competencies, proposes three generic key competencies, and puts forth four conceptual elements as relevant for defining and selecting key competencies.

---

31 This perspective contrasts with one that is focused on an internal mental structure for cognitive abilities and skills.

32 Within the DeSeCo program, the focus is on the individual rather than the collective concept of competence which focuses on the prerequisites needed by a group to meet complex demands. This choice is pragmatic and implies no judgment e of the two distinct theoretical-conceptual approaches.
Five viewpoints

24. The sets of key competencies proposed by the authors reflect different discipline-oriented viewpoints, different models for understanding the roles of the individual and society and their relationship, and different theoretical emphases regarding the tension between status quo/continuity and change/innovation.

25. **Monique Canto-Sperber** and **Jean-Pierre Dupuy** are philosophers who establish a set of values that are useful for defining a good life and are consistent with major moral theories: accomplishment, choosing one’s own course through life, understanding oneself and one’s world, enjoyment, and deep personal relationships. In addition, they establish the premise that the human mind cannot be reduced to a set of abstract rules or algorithms that describe the workings of a machine. Canto-Sperber and Dupuy proceed to identify five clusters of competencies needed for a good life: coping with complexity (recognizing patterns); perceptive competencies (discriminating between relevant and irrelevant features); normative competencies (choosing the appropriate means to reach a given end, appreciating various possibilities offered, making moral judgements and applying them); cooperative competencies (cooperating with others, trusting others, taking the role of the other); and narrative competencies (making sense of what happens in life to oneself and others, describing the world and one’s own real and desirable place in it). These competencies can be construed as dimensions of a five-dimensional space, with sub-competencies and skills pertaining to several if not all of the five areas of key competencies.

26. The anthropologist **Jack Goody** rejects engaging in a decontextualized discussion of key competencies on grounds that theory must always be considered in the context of practice. Recognizing that there may be some very general qualities required by modern life, Goody focuses on the intractability of specifying key competencies at a level that can span cultures, social contexts, and individuals and at a level that would also be useful for developing measures. He also cautions against limiting the work to developed countries because it is bound to be used in a larger context and have a negative, homogenizing effect.

27. The work of **Helen Haste**, a psychologist, is based on the view that humans are adaptive, social beings whose competencies both derive from these attributes and allow them to meet the demands of particular historical periods and social contexts. She emphasizes the need to look at individuals in a cultural, social, and linguistic context, and proposes creative interaction with and management of the tension between innovation and continuity as an overarching meta-competency. From this, she identifies five key competencies: to adaptively assimilate changing technologies into social practice; to deal with ambiguity and diversity; to find and sustain community links; to manage motivation and emotion; and to focus on morality, responsibility, and citizenship. She describes the competent individual as one that “is self-sufficient, able to focus attention and plan, with a future orientation, is adaptable to change, has a sense of responsibility, has a belief that one can have an effect, and is capable of commitment”.

28. **Frank Levy and Richard Murnane** are economists who see recent social changes related to technology and globalization as transforming the competencies needed in the workplace. They use relevant economic theory and empirical results—as well as their own research about the hiring practices of high-wage firms—to address the question of the competencies workers need to succeed in the labor market. These competencies include reading and mathematical skills (not only for their instrumental use but as the basis for lifelong learning); oral and written communication abilities; skills to work productively.

---

33 Canto-Sperber and Dupuy (1999), op. cit.
in different social groups; emotional intelligence and related abilities to co-operate well with other people; and familiarity with information technology. The competencies are identified in a particular field and for a specific group (e.g., employees in the labor market), but are seen as relevant across social fields and groups.

29. For sociologist Philippe Perrenoud\(^ {37} \), the central question to be addressed is what competencies are needed by everyone to freely exercise his or her autonomy in multiple social fields. He uses the term social field in the sense developed by Pierre Bourdieu—a structured set of social positions dynamically organized around a given set of social interests, with its own norms, power relations and capital governing social interactions\(^ {38} \). In this formulation, competencies apply across all social fields and are not specific to any one field. The key competencies Perrenoud proposes are: being able to identify, evaluate, and defend one’s resources, rights and limits; to form and conduct projects and develop strategies, individually and collectively; to analyze situations and relationships; to co-operate, act in synergy and share leadership; to build and operate democratic organizations and systems of collective action; to manage and resolve conflicts; to understand, apply, and elaborate rules; and to construct negotiated orders beyond cultural differences.

30. To summarize, the experts structure their approaches to identifying key competencies around different organizing ideas, central questions, and conceptual frameworks. With the exception of Goody, they each propose a set of key competencies needed by the individual for what the authors define as success, reflecting their overall approaches. Levy and Murnane, as economists, detail those competencies needed for success in today’s labor market, while the sociologist Perrenoud identifies key competencies as those needed for the autonomous actor to successfully defend his or her rights within society. The philosophers, Canto-Sperber and Dupuy, identify the competencies needed by the individual for a good, meaningful life. Haste bases her competencies on psychological models of how individuals interact with their environment.

**Three generic key competencies**

31. A number of common considerations and arguments transcend the heterogeneity of these discipline-oriented approaches. We have identified three generic key competencies, which encompass many of the key competencies identified in the five expert reports. These are:

– acting autonomously and reflectively;
– using tools interactively; and
– joining and functioning in socially heterogeneous groups.

**Acting autonomously and reflectively**

32. Autonomy of individuals is a central feature of modernity, democracy, and individualism. In particular in Western societies, the value system promotes autonomy as an aspiration and basis for individual identity. Acting autonomously means that individuals can assert their own rights and interests, think and act for themselves, initiate interactions with their physical and social environment, form and conduct projects, and develop strategies to attain goals.

---


33. Reflective exercise of autonomy requires an awareness and an understanding of one's environment: how it functions and how one fits into it. This encompasses a familiarity with the knowledge, values, rules, rites, codes, language, law, institutions, etc. specific to the field—i.e. knowing the rules of the game. Further, it requires incorporating this awareness into how one plays the game. The fact that people live by internalized social norms and in the context of relationships to others is not incompatible with autonomy. Scrutinizing and reflecting on these norms and relationships is part of individual growth and the maturation of identity.

34. It is inevitable that in some environments, the possibility for autonomous action is quite restricted. In these cases, the challenge for society is to empower actors to become as autonomous as possible within these limits.

*Using tools interactively*

35. “Tool” is used here in the broadest sense of the term. It encompasses physical entities, language and knowledge, laws, etc. that are relevant to meeting many important everyday and professional demands of modern society. Tools are seen as instrumental for an active dialogue between individuals and their environment. Underlying this is the idea that tools shape how we make sense of our world, how we become competent in our interactions with it, and how we deal with transformation and change. Using a tool, in this sense, implies not only having the tool and being able to use it effectively, but also understanding how the tool affects the way one interacts with the environment. For example, competencies associated with information technology include awareness of the new forms of interaction made possible by computers (such as chat rooms, email exchange, and other types of virtual communication) and the ability to adapt to them, as well as the technical skills required to manipulate computers. With interactive tool using, new cognitions and new social practices become possible.

*Joining and functioning in socially heterogeneous groups*

36. The third generic key competency is related to human dependence on others for material and psychological ends. Relationships are necessary for a sense of self, identity, and social meaning. In light of the demands of contemporary democratic societies, individuals must have the capacity to form, join, and function effectively and democratically in multiple, complex, and socially heterogeneous groups.

37. Joining and functioning in heterogeneous groups consists of a number of components. One important component is perceiving and understanding the distinctive position of the other. Other components are negotiating conflicting interests in order to find mutually acceptable solutions, operating democratically in groups, constructing negotiated orders over and above cultural differences, and the will to develop joint strategies, etc. This requires balancing commitment to the group and its norms with the capacity for autonomous action.

38. Future efforts in this area should consider how these generic key competencies are interrelated. This calls for both theoretical and empirical work.

*Four conceptual elements of key competencies*

39. Consistent with DeSeCo’s objective of advancing the theoretical foundations of key competencies, our work focuses on discerning conceptual or theoretical elements potentially leading to the definition and selection of relevant competencies, rather than on proposing a new list of key competencies. What follows are four analytical elements, which seem to be relevant for a multidisciplinary approach to

---

key competency in an international context. Together they form a starting point for conceptualizing and describing key competencies from a theoretical viewpoint.

1. **Key competencies are multifunctional**

   Consistent with the functional approach (see above), which views competencies as structured around meeting complex demands, the concept of key competence is invoked only to designate a competence that is needed to meet a range of different and important demands of daily, professional and social life. Key competencies are needed to achieve different important goals and to solve multiple problems in various contexts.

2. **Key competencies are transversal across social fields**

   The term “transversality” is used with a particular meaning. While in the educational field, transversal competencies are viewed as spanning various scientific disciplines or educational subjects (cross-curricular competencies); we refer here to competencies that transverse various sectors of human existence. Thus, key competencies are relevant for effective participation not only in school and the labor market, but also in the political process, social networks and interpersonal relations including family life, and most generally, for developing a sense of personal well-being.

3. **Key competencies refer to a higher order of mental complexity.**

   When constructing key competencies, it is important to make explicit the level of mental complexity involved. Key competencies assume a mental autonomy, which involves an active and reflective approach to life. They call not only for abstract thinking and self-reflection, but also for distancing oneself from the socializing process and even from one’s own values, to be the originator of one’s own positions. To give a few examples, at work this means being self-initiating, self-correcting, and self-evaluating rather than dependent on others to frame the problems, initiate adjustments, or determine whether things are going acceptably well. As a citizen of a diverse society, this means resisting the tendency to make “right” or “true” that which is merely familiar, and “wrong” or “false” that which is only strange, examining at and evaluating the values and beliefs of one’s psychological and cultural heritage rather than being captive of those values and beliefs. In school, this implies critical thinking and being a self-directed learner, taking initiative, setting one’s own goals and standards, using experts, institutions, and other resources to pursue these goals.

4. This active and reflective approach is based on an evolutionary model of human development in which individuals can incorporate higher levels of complexity into their thinking and actions.

5. A variety of research suggests that people have the potential to gradually reach higher levels of mental complexity throughout the course of their lifespan.

---

40 Philippe Perrenoud (1999), op. cit.

4. Key competencies are multidimensional

46. It is also useful in the conceptualization of key competencies to consider them as being composed of multiple dimensions, representing mental processes. They are composed of “know-how”, analytical, critical and communication skills as well as common sense.

47. These five dimensions are:

- Recognizing and analyzing patterns, establishing analogies between experienced situations and new ones (coping with complexity).
- Perceiving situations, discriminating between relevant and irrelevant features (perceptive dimension).
- Choosing appropriate means in order to reach given ends, appreciating various possibilities offered, making judgments and applying them (normative dimension).
- Developing social-orientation, trusting other people, listening and understanding others’ positions (cooperative dimension).
- Making sense of what happens in life to oneself and others, seeing and describing the world and one’s real and desirable place in it (narrative dimension).

48. The three generic key competencies identified earlier in this paper, together with the four conceptual elements, are proposed as a potentially productive avenue for exploring and studying concrete forms of key competencies as manifested in actions, behavior and choices of individuals and groups in different social fields (such as personal, social, economic, political, and cultural life), at different stages in life, and in different cultural contexts.

Implications for policy-making and future research

49. Since the development and interpretation of policy-wise indicators is a goal of this work, we conclude with a brief discussion of policy issues relevant to the topic of key competencies and indicator development.

50. Given the complexity of the demands and the challenges of the modern world, lifelong learning has become a leading theme in education, economic and social discourse. In addition, the development of competencies occurs in a wide range of social institutions where formal and informal learning takes place. Although schools will continue to play a crucial role, other institutions include the family, voluntary associations, religious organizations, cultural activities, recreational activities, and the workplace. Thus, consideration of policies related to competencies raises questions regarding the objectives of education and of the role of different institutional settings in lifelong learning. Policy related to each of these institutions should consider their potential for fostering the development of competencies. Social partnerships should also be taken into consideration.

51. In this context, many questions arise. What is the overall level of competence in the population? What is the competency profile of different groups of people? Where should resources for competency development be targeted? Are existing policies designed to support competency development effective? What is the relationship between competencies and social, political, and economic outcomes? Indicators in these areas will improve the design and monitoring of policies about competencies.

52. Yet, to obtain such information, considerable effort is needed in the areas of scientific inquiry and political discussion. We need a better understanding of how different social fields operate and how

---

42 Canto-Sperber and Dupuy, op. cit.
this affects demands on individuals and society. Further, we need a better understanding of the relationship between general social conditions (such as globalization and new technology) and the demands on individuals and society. An iterative process between theoretical and empirical work (including the development of measurement methodologies) is key to future success in this field. Recognizing that defining and selecting key competencies is ultimately the result of a political negotiation process, in which researchers can only be partners among others, we encourage within the DeSeCo Program the dialogue between scholars, representatives of the economic world and the social sphere, as well as policy makers. As we continue to develop our ideas about key competencies, we need to work towards a conceptualization that is both theoretically grounded and relevant to the policy world.
SECTION II

CONTRIBUTIONS FROM NETWORK B
INTRODUCTION

1. The mandate of Network B from its beginnings has been to develop indicators related to the role of education and training outside of the formal educational system. The aims of education have always been viewed more broadly than the imparting of certain basic elements of knowledge and skills to youth. Indeed the fact that governments historically have become heavily involved in the financing and organisation of education is an implicit recognition of the nature of education as a «public good», which contributes to individual well-being, but also to the creation of a national identity, to participation in democratic institutions and processes and to national economic development.

2. Although these are underlying objectives of educational systems, the extent to which they are successful in attaining these objectives cannot always be assessed within the confines of educational systems themselves. The latter can measure how well youth have mastered basic literacy and numeracy skills or whether certain lessons of history, principles of democratic life, an understanding of the place of man in the natural environment, etc. have been mastered, but not whether these are having the expected downstream effects on the functioning of society and the economy. There is a certain element of faith, or perhaps of common sense here, in the sense that one does expect the outcomes to bear some relation to the underlying objectives, if the education system is a “working” one. Still international comparisons do provide the means to assess how well education contributes, to a greater or lesser degree, to the furthering of the broader objectives outlined above.

3. A second feature of the mandate of Network B rests on the observation that not all learning takes place within the formal educational system, nor does it end with the completion of initial education. In certain respects, this is something of a truism, for the reason that human beings are by nature learning organisms. However, in an environment where new forms of codified knowledge and new technologies are being regularly introduced, to which a significant segment of the work force will not have been exposed during initial education, the ability and the opportunity to learn throughout adult life becomes a critical one, if workers and indeed, the general population, are to be able to adapt. Hence the work in developing indicators that map how education systems, enterprises and individuals go about ensuring continuing learning throughout adult life.

4. One main pillar of the work of this network has been to develop indicators which attempt to link education and training to broader outcomes, and in particular labour market incomes. The work on the school-to-work transition is a special case of this, which focuses on the immediate interface between education and the labour market, where the knowledge and skills imparted during initial education encounter the needs of the labour market.

5. Thus far, the broader work on labour market outcomes (see Nordin in this section) has focused on measuring results in the labour market as a function of educational attainment. Although the results here have not been especially surprising (labour market performance generally improving with educational attainment), they have shown considerable differences in the payoffs to additional education in different countries. Labour market outcomes, however, are heavily influenced by the overall economic situation in countries as well as by particular labour market institutions, over which education systems have little control. For example, collective bargaining systems may induce a certain amount of wage compression, which tends to reduce the payoff to additional education.

6. Labour market outcomes are of course not the only outcomes of education. New indicators for economic returns to education (see Jin in this section) describes several others, in particular differences by
educational attainment with respect to receipt of social transfers and to payment of income taxes. These are helpful in showing more clearly social / fiscal benefits to investment in education by governments. However, these continue to be monetary benefits and do not yet address what are believed to be numerous non-market benefits of education, for example, better health, increased participation in public life and voluntary activities, increased civic-mindedness, etc., which are perhaps avenues, albeit difficult ones, for future development work.

7. Certainly, one issue for the future is how much further to go in linking outcomes to factors over which educational systems have some control or which can act as signals to persons facing a choice among various educational options. Among these education variables, for example, are the field of study and the orientation of the education pursued (general versus vocational). Because it is difficult to assess the likely direction of outcomes over the longer term, however, a focus on these for persons who have recently entered the labour market may be a judicious strategy.

8. The Network subgroup on the school-to-work transition has done some significant ground-clearing work with respect to documenting differences across countries with respect to the transition process. To a certain extent, some of the turbulence and problems of the transition process are inevitable, because they reflect the encounter of inexperienced labour market entrants and of employers with little knowledge of the capacities of applicants. Here too the issue is to address what education and training systems can do upstream to facilitate the transition process, particularly for those whose problems in integrating the work force are likely to be persistent, rather than transitory, which is the case for most young people. Educational determinants and other aspects of the transition process (see de Broucker in this section) describes past achievements in this area, as well as mapping out a road for future work.

9. The measurement of continuing education and training has traditionally been a difficult one, because much of it occurs outside the formal educational system administered by government and has received less attention. The fact that investments in training by enterprises are not considered investments for accounting purposes means that in many cases, record-keeping is less than adequate and obtaining an accurate picture of both participation and spending difficult. Moreover, the character of much learning by adults tends to be informal, although attention tends to be focused on those forms which have clearly identifiable costs, of which on-the-job-training is perhaps the least structured form. Indicators on Continuing Education and Training (see Borkowsky in this Section) outlines some reflections and plans of the Network subgroup that is addressing this difficult area.

10. The work of Network B has up to now consisted of exploiting existing data sources for what they are able to yield on outcomes of education and on participation outside of the formal system. The future is likely to be more a challenging one, because of the need for attention to broader outcomes, a closer link to malleable factors that affect outcomes and the difficulties associated with measuring continuing education and training.
LABOUR MARKET OUTCOMES

Allan Nordin, Statistics Sweden, Stockholm

Introduction

1. Education and training activities are aimed at meeting a variety of purposes. In broad terms one could say that these purposes include preparing the individuals for everyday life and promoting personal growth, general knowledge, critical thinking and democracy. A more specific aim of education and training is to prepare the individual for working life. This chapter focuses on the labour market aspects and mainly on the outcomes of education and training in terms of employment and earnings. Also indicators providing back-ground information on educational attainment are dealt with in the chapter.

2. Some of the key policy questions in relation to education and labour market are:
   - How do those educated at various levels of the education system fare in the labour market? What kinds of education and training matter in bringing about labour market participation and sustained employment and further learning opportunities?
   - What will characterise the demand for educated labour in the future? What kind of flexibility in the education system will be needed to meet future labour market demand?

3. Policy makers dealing with these and other questions are calling for all kinds of information in the form of statistics, indicators and analysis providing insights to underlying processes and conditions. Indicators may also invite policy makers and analysts to raise questions and draw on evidence to make better choices.

4. Education is meant to have effects not only at the individual level but also at society level. In theory the overall investment in education will have positive effects on productivity growth and hence on economic growth. However empirical data give very little evidence for such an impact of education on economic growth. At the micro level, on the other hand, this impact can clearly be seen.

5. There has been a growing interest in measuring the amount of human resources in OECD countries in recent years. At the same time the interest in performing international comparisons has also been growing. Therefore the presentation of international indicators on education in Education at a Glance has been of great interest to policy-maker, researchers and other persons engaged in the ”education industry”.

6. This chapter contains a description of the developmental work undertaken by Network B (NwB) regarding indicators in the field of labour market outcomes. Among these indicators several have been used in the INES publication Education at a Glance (EAG). They are listed in the appendix. Some of the results of the indicators are also presented in this chapter. At the end of the chapter there is a presentation of ideas of possible additional indicators with connection to labour market outcome of education.

7. The area of labour market outcomes relates to other types of issues in connection to the results of education. Therefore some of the indicators presented here have clear connections to other chapters in the
compendium. Thus the indicators on earnings touch upon the area of rates of return to investment in education. The indicator on intergenerational change has connections to the issue of equity, whereas the indicators on labour force activities of leavers and labour market entry clearly are connected to the issue of transition from school to work.

**The role of education in economic theory**

8. Traditionally the factors of production are grouped in three categories: Land, Labour and Capital. In this chapter we focus on factors related to labour. In this connection labour refers to "all human attributes, physical and mental, that are used in production". Like land and capital labour is not a homogeneous factor as some jobs require very little training, whereas others require many years of training. Long time ago when all production was very labour intensive the main thing was the number of persons available for work. Today the skill level of the labour force is the crucial factor. That will certainly be even more the case in tomorrow’s knowledge society.

9. Education invested in trained labour is often referred to as human capital. The knowledge and skills accumulated in human capital also include the effects of informal learning and workplace experience. The theory of human capital dates back to Adam Smith in the 18th century and his theory of compensating wage differentials (A p. 24-25, B p. 121).

10. The theory of human capital offers an explanation of the incentives needed to promote investment in education. Generally the individuals are interested in investing in education if they can expect a subsequent rise in wages. This difference in compensation - the so-called education premium - varies greatly across the countries and the debate always goes on whether this premium is large enough to encourage individuals to invest in education. Other reasons for individuals to invest in education might be to reduced their risk of unemployment and to increase their job satisfaction.

**Existing indicators on educational attainment and labour market outcomes**

11. There is a set of indicators on educational attainment and labour market outcomes which can be regarded as "core indicators". Some of them have been included in all editions of EAG, others in almost all of these publications. Most of them are relatively stable over time since they are not effected by changes in education policy from year to year. These indicators are often referred to by policy-makers when they evaluate they evaluate their national education policies and when they present ideas for future education policies.

**Educational attainment**

12. One of the main goals for the education and training policies of the OECD countries is to enhance the level of educational attainment and the skill level of their population. The reason for that is to improve or retain the respective countries’ competitiveness. Therefore indicators on educational attainment are of great interest to policy-makers.

13. A group of indicators showing the educational structure of the adult population has been developed by Network B. The motivation for these indicators has been that "educational attainment - and, by extension, labour force qualifications - are important factors in determining economic outcomes and the quality of life for individuals and society as a whole".

14. The indicators belonging to this group have usually been labelled:

- Educational attainment of the adult population, 25-64 years
- Percentage of the population that has attained a specific level of education
15. The indicator called "Educational attainment of the adult population" is providing a fundamental picture of the education level in the OECD countries. It has been retained in all editions of Education at a Glance (EAG) from the very first one in 1992. This indicator is showing the population aged 25-64 distributed by completed levels of education. The indicator has been described as "largely based on formal educational qualifications and only partially captures the skills and competencies that are acquired through adult, or continuing education and training, or other non-formal ways of learning at home and in the workplace".

16. In most editions of EAG aggregations of the above described indicator have been presented. One of these aggregated indicators is showing the percentage of the population having completed at least upper secondary education. In the 1997 edition of EAG this indicator gave not only a picture of the situation in each country in 1995 but also a projection for the years 2005 and 2015. This was the case also for the indicator showing the percentage of the population having completed tertiary education in the same publication. In EAG 98 the latter indicator was not showing the proportion with tertiary education, but with education at university level. The indicator (A1.2a in EAG 98) was showing that the proportion having graduated from university still varies widely across the countries. Thus only 6 - 8 per cent of the population aged 25-64 in Austria, Italy, Portugal and Turkey had completed an education at that level, whereas this proportion was well above 20 per cent in the Netherlands and the United States. The picture changes considerably if countries are ranked in respect of the proportion having attained at least education at upper secondary level. Then countries with a large proportion having attained upper secondary level, such as Austria, Czech Republic, Germany, Poland and Switzerland, would be ranking much higher.

17. The data in EAG 98 were also revealing large differences in the educational attainment between generations. Thus the proportion having attained university level was much higher in all countries among 25-34 year-olds than among 55-64 year-olds. This means that this proportion will be steadily growing in the labour force for several years to come. In 11 out of 26 countries the proportion of university graduates was at least twice as high among those aged 25-34 than among those aged 55-64.

18. Data for age-groups by gender are showing another type of change between generations. In 1996 (presented in EAG 98) the proportion of people having graduated from university was on average twice as high among men as among women in the age 55 to 64. The average proportion was 10 per cent for men and 5 for women. Among 25-34 year-olds there was no difference at all between men and women regarding the average proportion of graduates. For the individual countries 11 had a higher proportion of graduates among women than among men at the age of 25 to 34 and in 14 countries the situation was the opposite.

19. The indicators in the EAG publications showing gender differences in educational attainment have been calculated on the same data as those showing the attainment level of the whole population. In EAG 92 the number of women in each education category was divided by the number of men in the same category. The quotient was then multiplied by 100. In EAG 93 -96 the proportion of women was shown by just dividing the number of women by the total population in the same educational group.

20. Differences between men and women in educational attainment have been presented in EAG 93 and 95 in terms of an index. In EAG 93 this index was showing the proportion of men and women that would have to change education level to bring about similarity in the educational structure of the two genders. In EAG 95 the index was calculated by weighting together the differences between men and women in proportions for each educational level with the largest weights for the two extremes (primary education and university education). In the latter publication calculations were made for two age-groups, 25-34 and 55-64 respectively. The gender differences were in almost all countries much larger among the
older persons than among the younger ones. Despite the differences in calculation methods between EAG 93 and EAG 95 virtually the same countries proved to be those with the largest gender differences. Australia, Austria, Germany, Netherlands and Switzerland were among the countries with the largest gender differences in both publications.

21. A new measure of educational attainment was presented in EAG 96, namely the indicator Cumulative years of schooling in the population aged 25 to 64 and 25 to 34. It was calculated by converting the attainment levels for men and women into years of schooling on the basis of the theoretical cumulative duration of the respective levels of education. This indicator was used to show differences in educational levels between men and women.

22. Labour force and household surveys have been used as data sources for the above mentioned indicators in most countries. In EAG 92 the number of countries having reported data on educational attainment was 20 out of a total of 24 OECD countries at that time. In EAG 98 attainment data was published for 26 out of 29 OECD countries. The missing countries were Iceland, Japan and Mexico. In EAG 98 attainment data were published not only for OECD countries but also for eight Asian and South American countries not being members of OECD but involved in the World Education Indicators programme (WEI).

**Labour force status by educational attainment**

23. In general there are significant differences in labour market performance between educational levels. It is of interest to see if these differences are the same across OECD countries and how they relate to initial education. Therefore NwB has brought together material that can shed light on how different educational groups fare in the labour market. A very simple way of getting an indication of differences in participation in the labour force is to compare the educational distribution of the population with the same distribution of the labour force. Such an exercise is verifying that those having an education at tertiary level are making up a considerably larger proportion of the labour force than of the population in all OECD countries. Tables showing labour force participation rates for different educational groups are leading to the same type of conclusions namely that those having attained higher levels of education are more active on the labour market. This doesn’t necessarily mean that they also are employed in larger proportions than those with lower levels of education. This is however indirectly proven in tables showing unemployment rate by education. This rate is usually considerably higher for those with a lower level of educational attainment than for those with a higher level.

24. All these labour market aspects were summarised in one indicator in EAG 97 and 98. This indicator was showing expected years in employment, out of the labour force and in unemployment for the different educational levels. By using other labour force data and converting rates into time it was possible to synthesise the labour force statistics into one indicator.

25. The data source for the above mentioned indicators is the national labour force survey for practically all countries. In addition to the above indicators NwB has developed indicators measuring the labour force status of persons leaving education. The data sources of these indicators are typically follow-up surveys describing what persons were doing one year or five years after leaving education. These surveys are not as harmonised as the labour force surveys and thus do not produce as comparable data as the labour force surveys.

26. The following indicators are showing labour force status by educational groups:

- Labour force participation rates
- Unemployment rates
- Employed and unemployed as percentage of the population
- Labour force activity of persons leaving education
Expected years in employment, out of labour force and in unemployment

27. A general finding regarding the labour force participation rates is that these rates are higher for men than for women at each age and level of educational attainment. This difference is largest in the oldest age-group (55-64). This is probably not a reflection of differences in health status between men and women but rather a reflection of woman’s role as it was seen earlier. Also young women participate less than men. For 25-34 year-olds this difference is certainly due to the fact that many young women leave the labour market when they have small children.

28. Another finding is that the high-educated participate in larger proportions in the labour force than those with lower levels of education. This is true for both men and women but much more pronounced among women. As an example one can look at data for Australia. In 1996 93 per cent of Australian men aged 25-64 with a university degree were in the labour force. The corresponding proportion for Australian women was 82 per cent. The difference in participation rate between university graduates and those with an education below upper secondary level was 12 percentage points among men and 27 percentage points among women. In some other countries the difference among women were even larger.

Earnings and education

29. Earnings is a type of labour market outcome of education that is providing information about the balance between supply and demand for different groups. Network B has developed an indicator showing the average earnings for each educational level related to the average earnings of those having attained upper-secondary education. This allows comparisons of earnings within each country but not across countries. What can be compared across countries is the structure of relative earnings of the educational groups.

30. It is not always easy to interpret the results of the calculations of relative earnings. The differences in earnings observed between educational groups include effects due to differences in hourly (weekly or monthly) earnings as well as differences in hours worked per year for each of the groups. During the developmental work undertaken on the earnings indicator the debate has been divided. On one hand it has been argued that the network should try to measure the impact of each of the components. On the other hand some network members have expressed the opinion that the indicator should include even more components e.g. unemployment and non-participation in the labour force without any split-up. This could be done by dividing total earnings of those at work and having attained a specific education level by all persons having attained that particular level. The motivation for doing so was that this indicator should combine a lot of different effects of education into one single measure. The indicator which has been presented in EAG over the years is combining different effects into one measure. It has however been restricted to measuring the earnings of those who had an income during the reference year.

31. In the other earnings indicator developed by NwB mean annual earnings of women are compared with mean annual earnings of men at each educational level. Also in this case different components come into play. Since women typically work less hours per year than men the number of working hours plays an important role to explain differences in earnings. In addition the hourly earnings of women usually are lower than those of men. Therefore the resulting index figure has in practically all cases been well under 100 for women in all countries and at all educational levels in all EAG publications where the index figure for men in the corresponding groups have been 100. There have been very few exceptions from the rule that men earn more than women. Looking at different groups it seems as if the differences between men and women are smaller in some countries among young persons than among older persons and among those with higher degrees than among those lower educated.
Areas for further development

32. This section of the chapter is including some indicators which have been developed some years ago and in some cases also have represented by data in an edition of EAG. Others have been discussed at network meetings but never been used for data collection. A third category of indicators contain ideas of indicators that have not been discussed at any network meeting.

33. The proposed indicators presented here can be grouped together in three groups, namely those shedding light on equity issues, those showing the impact of learning on labour market outcomes and those giving background information.

34. The first type of indicators is represented by two sets of indicators, namely:
   - Intergenerational and social equity in attainment
   - Migration and education (This set of indicators can be regarded as equity indicators if they focus on differences in for example educational attainment or job opportunities between the migrants and the main population).

35. The second type of indicators is represented by:
   - Low-educated/low-skilled
   - High-educated (the match between field of education and occupation)
   - Labour force status for leavers (by field of education)
   - Labour market entry (or Time to stable job)
   - “Inactive” population by educational attainment
   - Quality of working environment by educational attainment

36. The third group may consist of the following areas for further studies:
   - Knowledge economy (knowledge-based industries)
   - Migrants and education (if the focus is put on the brain gain or brain drain that the migrant represent to the individual country)

**Intergenerational and social equity in attainment**

37. In EAG 98 an indicator was presented on intergenerational change in completion of tertiary education. The indicator was showing the percentage of the population (16 to 65 years of age) who had attained tertiary education, by level of attainment of their parents (indicator A2.2)

38. As a policy context of the indicator was mentioned a high correlation between the educational attainment of individuals and their parents. Such limits to intergenerational mobility can create problems both in terms of equity and in terms of raising the stock of human capital.

39. Further development of this indicator would be desirable. A problem to address is where to find data for this indicator. The indicator in EAG 98 was based on data from the International Adult Literacy Survey (IALS). In order to be able to maintain the indicator for a number of years ahead some other data source has to be found.
Migration and education

40. What do the flows of manpower between countries look like in terms of competencies? Which countries are subject to brain drain and which countries can benefit from brain gain?

41. There are different reasons for moving from one country to another. One reason may be differences in job opportunities and wages. Other reasons may have to do with political instability or armed conflicts between countries or within countries. For example a lot of refugees have left their homes in former Yugoslavia and moved to other countries in Europe. What possibilities do these people have to get a job in their new country?

42. NwB should consider developing an indicator shedding light on the impact of migration on the labour markets of the OECD countries.

Low-educated/low-skilled

43. Education policy has to be focused not only on those enrolled in education or those already highly educated but also on those who have a low level of educational attainment. In economies which increasingly become knowledge-based those low-educated/low-skilled are at risk of being marginalised. Therefore it is of great importance to develop an indicator or a set of indicators which can shed light on the labour market situation of these groups. It is important to describe the needs of education and training that have not been met. This indicator should be connected not only to the issue of labour market outcome but also to the issue of barriers to participation in continuing education and training (CET).

High-educated

44. The provision of competencies for enterprises, local and central government agencies and organisations is an issue for education policy. Therefore it is of great importance to evaluate the present development in the creation of competencies. That can be done by studying the trends over time in the number of higher educated.

45. For the individual persons enrolled in education it is important to see how those already educated fare in the labour market. If demand for skilled labour doesn’t grow in the same pace as the supply of skilled labour unemployment will rise and wages will fall in real terms.

46. By developing an indicator with time-series on employment and growth of wages for skilled labour in each country NwB could provide information about important trends in demand for different educational groups.

Labour force status for leavers (by field of education)

47. This indicator has been mentioned already in paragraph 2.2 among existing indicators. Since the indicator needs some further refinement it is also mentioned here as an area for further development. The indicator is showing the labour force status of those having completed their education a certain number of years ago. The data on the leavers should be presented not only by education level but also by field of education. To achieve good comparability across countries the measuring of status of the leavers should be done after the same number of years in all countries involved. The existing follow-up surveys are done in different ways and after a varying number of months or years. Also the definition of each status certainly would have to be gone through to make sure that combinations of activities are classified equally.
**Labour market entry (or Time to a stable job)**

48. This indicator may as well be regarded as transition indicator. It also connects to the indicator on labour force status for leavers. It would measure for how long each person had to seek a job before he/she got his/her first stable job after having completed education. There are of course a number of factors to be defined before this indicator can be clearly specified. For example the concept of stable job has to be defined both in terms of duration measured in months or years and in terms of number of working hours per week or per month. The type of employment contract (limited or unlimited duration) which the employee has got could be used to judge if a work is stable or not. Also the level of the income could be one of the conditions to establish when a job has been stable. To provide policy-relevant information the data on the time to get a stable job should be presented by educational attainment, field of education and gender.

**“Inactive” population/marginalised population**

49. People with a low level of educational attainment and not being enrolled in education or being employed can be characterised as being part of the “inactive” or marginalised population. It is of great importance to policy-makers to have a measure of the magnitude and structure of this group and to see if people remain in the group for a long time or if they only temporarily are to be seen as inactive/marginalised. NwB should develop an indicator based on Labour Force Surveys to present information on the inactive population.

**Quality of working environment**

50. The working conditions of those employed differ a lot not only with regard to type of work but also with regard to working environment and working conditions. To get a better picture of the total outcome of education NwB should try to develop a measure of the quality of working environment of those educated at different educational levels. Such information could complement the information discussed above about differences in employment, unemployment and earnings. Most OECD countries already collect data regularly on working environment and working conditions. When these data are collected by adding questions to the regular labour force surveys one could easily combine results on e.g. stress, physical strain and injuries related to work with information about labour force status and/or educational attainment of the individual person. Information about existing surveys on working environment can be found in various publications from e.g. Eurostat (C and D).

**Knowledge-based economy**

51. Recently OECD published a report on knowledge-economies (E). Data on investment in knowledge and physical capital in the OECD countries were presented. To arrive at a measure of total investment in knowledge different types of spending were added up such as spending on research and development, investment in software and public spending on education. NwB could refine this type of indicator and have it calculated each year. That would give a picture of how the knowledge-based industries are developing in each of the OECD countries.

52. The work that NwB could do in this area should of course be co-ordinated with similar work going on in other groups related to INES or to other OECD projects.
REFERENCES


Björklund A., P-A Edin, B. Holmlund and E. Wadensjö: Arbetsmarknaden [SNS Förlag 1996]

Social portrait of Europe [Eurostat September 1998]


INDICATORS ON CONTINUING EDUCATION AND TRAINING

Anna Borkowsky, Swiss Federal Statistical Office, Switzerland

Why indicators on continuing education and training?

1. As industrialised countries move ever more swiftly into the information age, their economies are increasingly reliant on the skills and knowledge of their people rather than on physical resources. This reliance on "people skills" places a premium on human capital development. One part of this development, initial skill acquisition, has long been recognised as a crucial determinant of individual, community, and societal performance and success. But as the workplace changes in ways that make continuous skill and knowledge development more important, continuing education and training are gaining in importance as well and many countries recognise the importance of these "later" skill development efforts. However, while much is known about the efforts expended by governments and individuals across OECD countries to promote learning within formal education institutions, far less is known about the extent of learning that takes place in the workplace or other settings outside of formal education, and after the completion of initial education. While some of this learning takes place through formal training courses, other learning occurs informally, "on the job" or through self-regulated learning; even less is known about these informal methods of learning. Better measures are needed, both within and across countries, of the extent to which adults engage in different types of learning outside of the formal education system, which adults are in need of opportunities for continued learning, and the existing barriers and incentives to continued learning--particularly for adults who are at risk of exclusion from the labour market or society.

2. While the goal of lifelong learning for all is established on the political agenda, there remain serious questions about the ability of our societies to meet the challenge. Providing for lifelong learning for all will put new demands on the organisational capacity of educational systems. As yet unresolved is the issue of how to organise best this provision of learning opportunities for all, and the roles to be played by the “standard” education system, by market-oriented providers, by employers and by other agencies. The differences between OECD countries in the policies pursued in this respect seem fairly substantial. Empirical findings and indicators on the provision of CET and the effectiveness of such provision schemes would be most welcome. Such information can help policy-makers evaluate, for example, whether their country’s individuals, businesses, or governments are under-investing in adult learning, and the appropriate policy responses to stimulate learning behaviour.

Past developments

3. At the international level, the last years of the 1990s have seen developments and advances in several different areas – the development of conceptual frameworks, advances in definitions and methods, data gathering both at the national and the international level, and the publication of indicators in EAG (Education at a Glance). The results give a rather diverse and as yet uneven picture.

Development of a framework for indicators of lifelong lifewide learning

4. Conceptual work within the OECD has resulted in a framework for indicators based on the concept of lifelong, lifewide learning. “The idea of lifelong, lifewide learning means an expansion of education and learning investment mainly outside the public sector and the traditional education
institutions, by providing and recognising education and learning at work place and in other situations all over the lifecycle.” (Härde 1999)

- National Surveys on Continuing Education and Training, in the Briefing Book for the September 1999 Network B meeting in Athens, Greece

- Chapter 3 “Training of Adult Workers in OECD Countries: Measurement and Analysis“ in Employment Outlook 1999

**Advances in definitions and methods**

5. The definition of the scope of educational statistics in ISCED 97 does not exclude "adult education" anymore, as was the case with the previous standard classification. How regular education data collections will take into account these new definitions remains an open issue.

6. Important work on detailed definitions and descriptions of CET are published in: OECD (1997a), Manual for better training statistics. However, while the relevant questions are raised and solutions for many conceptual problems proposed, the manual is not specific enough to serve as formal guidelines for the collection of CET statistics.

7. A first draft of such guidelines has been developed by OECD and submitted to international scrutiny in spring 2000. (OECD 2000a).

**Data gathered in national CET household surveys**

8. Several among the OECD countries have regular national surveys on CET, some going back to the 1980s. National household surveys on CET (Canada, Finland, Germany, Sweden, Switzerland, UK, USA.) have several common characteristics (Hörner 1999, see also Phelps & Stowe 1998):

- No proxy interviews
- A reference period of several (usually 12) months
- Questions about all types of CET, not only job-related
- A varying amount of further detail on all or selected formal learning activities such as duration, provider, content, motive and other characteristics
- Some information on informal learning (more common in the late 1990s than earlier)

9. The data collected are not strictly comparable, but there is much convergence in the key issues addressed by national surveys of Continuing Education and Training. Some of the recurring questions, themes and dimensions relating to CET are (i) participation (incidence), (ii) time spent in training (volume), (iii) the nature of the training (formal, informal, job-related, other), (iv) the source of financial support for the training (public, employer, self), (v) the extent of demand (especially unmet) for training or learning opportunity, and (vi) barriers to participation experienced by non-participants.

**Data gathered by international surveys**


11. The harmonised European Labour Force Survey (ELFS) has long included a question on participation in job-related training during the 4 weeks preceding the interview. The module on participation in education has been changed to include information on all training. Data are available from
1998 onward. Part of the results from the ELFS has been included in the NWB data collections since 1994.

12. The data source that has been used most intensely in EAG is the International Adult Literacy Survey (IALS).

**Indicators published in Education at a Glance**

13. The first indicator on CET participation appeared in the third edition of Education at a Glance (OECD, 1995a). It was called P08 “Participation in job-related continuing education and training,” and included breakdowns by educational attainment and gender, and could be calculated for the employed population only. One of its limitations was the use of different reference periods for participation in CET. For this reason, all the data presented had to be split – one part for those countries using a 12-month reference period and another for those using a 4-week reference period. The indicator was based on a special data collection by network B of the INES project. Data came from national Adult Education surveys and from the European Labour Force Survey.

14. Indicator P08 was more appropriate for a human capital indicators (HCI) approach, but the network aimed even then for a broader definition of CET, certainly including participation in all formal CET, but also in more informal forms of adult learning. (Borkowsky, Van der Heiden & Tujinman, in OECD 1995b)

15. The first indicators relied on the diverse national and international data sources available, but the volatility of estimates from source to source undermined confidence in the reliability and comparability of these data. (OECD 1999b, O’Connell 1999) The main issue is: can data from different surveys be used in the same indicator?

16. From 1997 onward, the CET indicators published in Education at a Glance were based on data from the International Adult Literacy Survey (IALS). While this survey provided the most comparable data of the available sources, it does not occur regularly and is handicapped by small sample sizes. The indicators developed from these data have included overall and job-related participation rates, the volume of participation, and various information on financing CET, the providers, the media of instruction and other aspects of CET. Information on the reasons for non-participation was also included. (OECD 1997b, OECD 1999a. See also Berlanger & Valdivielso (1997) and Belanger & Tujinman (1997) for a more extensive analysis of the IALS CET data.) The indicators developed culminated in showing the expected hours of training over the life cycle in EAG 2000. (OECD 2000b)

17. A further development of this idea results in the measure of the total expected years in education, including continuing education and training (chart 1). For the conversion of hours of training into years, the benchmarks proposed in EAG 2000 (30 hours a week and 40 weeks a year for “equivalent” full-time participation) have been retained. The chart shows that at present CET adds a “learning time” equivalent to between one and two years of education to the already fairly long duration of initial education. Formal CET amounts in most countries to less than 5% of the total educational time.
Towards an integrated system of statistics on continuing education and training or a CET account

18. An integrated system of statistics and indicators on CET will still depend on different data sources – mainly the participants in CET and the providers of CET, but maybe others such as financing or accrediting agencies –, because there will still be a need for different kinds of information. Official statistics often work with national and regional aggregates that may in part be estimated. The best known examples of this kind of synthetic statistic are the national accounts or the labour accounts (see Buhmann et al. 2000 for a description of labour accounts). For the calculation and construction of indicators on CET at the national or regional level, we might also envision building a synthetic CET account involving different, but harmonised and linked data sources. The ideal system differs from the present one in that all the necessary data sources exist for all countries, that these data sources are harmonised in their use of concepts and definitions, that the information that is necessary to link one data source to the other is readily available, and that the coverage in each country is fairly complete. The rest of this paper will focus on the issues involved in building such a synthetic account. The least attention will be devoted to the

---

43 I thank Pierre Doray for suggesting the development of an “ideal” system for CET statistics. For the first germ of the ideas developed here see Borkowsky 1997.

44 For the purposes of this paper, “harmonised” will refer to the use of common definitions. Data sets are said to be “linked”, when it is possible to adjust them to cover a common universe of CET activities. For instance, data coming from a household survey covering all private households can be linked to an enterprise survey covering only large enterprises by omitting the data coming from individuals who are employed by small enterprises.
harmonisation of concepts, definitions and guidelines for data collection, because this issue is already widely accepted as important.

**Different data sources will still exist...**

19. Even in the ideal statistics system for CET, there is not one data source covering all needs. The different sources all have strengths and weaknesses, which will have to be taken into account during data collection and interpretation. For the creation of synthetic CET statistics, careful consideration would be given to who would be in a position to know what, or who would be the best informant about different CET aspects.

20. The following main sources of data will still exist: data coming from the providers of formal CET, and data coming from the participants i.e. the learners in CET. These two sources can also be seen as the two basic players in a market, covering the supply and the demand in CET. In addition there will be data coming from enterprises. Enterprises are difficult to place in the conceptual framework of supply and demand, because they typically play a double role: they are providers or suppliers of CET for their own employees, but they also are buyers of CET for their employees which is then supplied by other providers.

21. In the 90s, attempts to publish internationally comparable data on CET have used only two sources: surveys of individuals such as IALS or other national surveys, and enterprise surveys such as CVTS. The providers have not been used as a source of international data on CET, although they are the main data source for regular education statistics.

**Source one: Persons – participants**

22. Every learning activity involves a human being who learns or attempts to learn. This individual is probably the best source of information about his or her motivation for learning, the way the learning activity inserts itself in his or her life course and personal situation, and for subjective evaluations of the learning activity. The individuals who do not engage in formal learning are the only source of information about the obstacles they confront.

23. Individuals are also a fairly good source of information about their own socio-demographic characteristics – gender, age, education achieved, labour force status, marital status, family situation and so on. The employed are usually also asked about their employer and its characteristics: size of the firm, industry of the firm, and less frequently about characteristics of their individual workplace.

24. Insofar as the formal learning events we are interested in involve a definite time and place, individual participants should also be able to indicate the number of activities they are involved in and the volume of each activity, although getting valid, reliable volume data for activities that occurred over an extended period of time is quite difficult. The individual can also give a very good subjective indication as to the purpose of the learning activity, be it pleasure and enjoyment or some job-related purpose.

---

45 See also Altena et al (1999) for the report on the SETA project, conducted as a part of the LEU Leonardo da Vici programme. It links educational participation data from different sources, namely provider data, data from the LFS and data from CVTS for EU countries. The least comparable data source proved to be CVTS, which was then excluded from the attempt at harmonisation.

46 However, one of the major changes in ISCED 97 is its extension to all educational activities, including adult education. Investigations of the coverage of CET in the “regular” education statistics are underway, but the discussion on data coming from providers other than enterprises has still a long way to go. This reflects the priority given to job-related CET and the role of employers as providers in the discussion up to now. It may also be linked to practical difficulties- e.g., there is no universe list of providers in many countries, so who would you survey?
25. In Adult Education Surveys, individuals are usually also asked to give a description of the learning activities, consisting of information about the content of the course, financial support received, the provider or organiser, and the kind of degree or certificate aimed at by the learning activity. All or any of these themes may be missing in a given survey.

26. The content of the course does not usually pose a problem, although the coding of contents may involve a lot of work and not be undertaken for practical considerations.

27. Most Adult Education Surveys have refrained from asking for amounts of direct expenditures and financial support received, and restrict themselves to asking about the sources of financial support. However, it may be possible to exploit specialised surveys on earnings and consumption to arrive at a global estimate of expenditures for CET by private households.

28. The information on providers is judged to be difficult. The difficulty of the question and the reliability of the information gathered vary across countries, depending not only on the quality of the surveys (which we assume to be beyond reproach), but also on the heterogeneity and the dynamics of the national CET system. A fairly simple system in which the vast majority of all job-related CET is provided by the employers and most non-job-related CET offered by a few schools is easier to describe and to capture in a questionnaire than a heterogeneous system in which employers tend to out-source CET, individuals buy their own job-related CET from market-oriented providers, and market-oriented providers and non-profit organisations cater to the leisure-oriented segment of CET. A first hypothesis might be: the larger the market-oriented segment of CET, the faster the change in the roster of providers.

29. Judging from what has been done in surveys of participants, individuals are thought to be less than perfect sources of information about the degree to which learning activities are structured (i.e. involve an explicit curriculum and didactic planning).

Source two: Providers - enrolments

30. The providers include schools, colleges, universities, professional organisations and trade unions, publishers of learning programmes, organisers of hobby courses and enterprises. They are the ideal source of information about the intended content of programs or activities offered as well as the place of CET in the national education system. The providers would typically be able to report whether a given activity is classified as vocational or general in its intent.

31. The providers are the best source of information about the teachers and other personnel involved in CET activities. They would also be very well placed to give information about expenditures for and costs of CET activities. They would probably be a good source of information on financial flows, but this information might have to be supplemented by data coming from government agencies for the case of publicly funded CET.

32. The providers can give very good information about the number of activities offered and the amount of “class-room” work involved by them.

33. They can also give information on the number of enrolments, and on some of the socio-demographic characteristics of the enrolled such as gender, age and labour market status. In all probability, the range of socio-demographic information gathered by the providers is fairly limited. We would not expect them to have a large amount of information on the individual characteristics of those enrolled.

34. We purposely do not speak of “participants”, because we wish to distinguish the participant (the person who participates in at least one activity) from the entry in the student roster. Since one person can

---

47 Providers are the sources of most information in the „regular“ education statistics. The discussion on the information providers can or could provide is based on these statistics.
engage in more than one learning activity, the relationship between participant and enrolments is one to many. A harmonisation of the student rosters of the different providers is not possible in most countries, and so persons who frequent several providers are in all probability counted several times, if we just add the provider statistics.

35. Dropouts from CET activities present a further problem of as yet unknown magnitude: do the providers count them, and do the individuals also report learning activities which they have not finished?

*Enterprises – enrolments by employees*

36. In its role as provider, the enterprise can be a good source for the same kind of information as the other providers: content, volume, personnel and finances involved in CET activities provided for by the enterprise, if it has a fairly good accounting system.

37. In its role as a buyer of CET provided by others, we can also expect some information as to the content and the volume involved, both in terms of the number who enrol and the volume of the activities.

38. The enterprise would be the ideal source of information about itself: the size, the industry, whether it is growing or declining, etc. The enterprise would also be a good source of information about the employees enrolled in employer-sponsored CET: their place in the work-hierarchy, their occupation, the degree of qualification demanded by the work being done, the cost of a working hour of the employee. They could be questioned on their CET policies: Who is eligible? What is required versus voluntary? Do they prefer to organise in-house training or to buy CET on a market? Are informal training schemes such as job-rotation or quality circles provided? And other issues.

39. The enterprise would be a good source of information about some socio-demographic characteristics of the employees such as age and sex, but probably a less good one about other individual characteristics such as educational attainment. The enterprise would be expected to be a doubtful informant on more complex or subjective characteristics of the employees such as the range of learning activities engaged in, individual motivations to participate in learning activities or barriers experienced by the non-participants.
The following synoptic table gives an overview of the data sources and the types of information to be gathered from each source. The number of XXs is an attempt to indicate the degree of appropriateness of each source.

<table>
<thead>
<tr>
<th>Information</th>
<th>Data source and reported statistical unit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual: Participants and Enrolments</td>
</tr>
<tr>
<td>Characteristics of participants/enrolments</td>
<td></td>
</tr>
<tr>
<td>Basic socio-demographic data (gender, age)</td>
<td>XXXX</td>
</tr>
<tr>
<td>Labor market status</td>
<td>XXXX</td>
</tr>
<tr>
<td>Complex socio-demographic data</td>
<td>XXXX</td>
</tr>
<tr>
<td>Subjective data</td>
<td>XXXX</td>
</tr>
<tr>
<td>Number of events</td>
<td>XXX</td>
</tr>
<tr>
<td>Characteristics of EACH activity</td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>XX</td>
</tr>
<tr>
<td>Place in national education system</td>
<td>X</td>
</tr>
<tr>
<td>Volume</td>
<td>XXXX</td>
</tr>
<tr>
<td>Motivation</td>
<td>XXXX</td>
</tr>
<tr>
<td>Characteristics of provision</td>
<td></td>
</tr>
<tr>
<td>Type of provider</td>
<td>XXX</td>
</tr>
<tr>
<td>Personnel</td>
<td>XXXX</td>
</tr>
<tr>
<td>Expenditures</td>
<td>XXXX</td>
</tr>
<tr>
<td>Opportunity costs</td>
<td>XXXX</td>
</tr>
<tr>
<td>Characteristics of workplace of participants or persons enrolled</td>
<td></td>
</tr>
<tr>
<td>Enterprise (size, industry)</td>
<td>XXX</td>
</tr>
<tr>
<td>Skill demands</td>
<td>X</td>
</tr>
<tr>
<td>CET policy</td>
<td>XXXX</td>
</tr>
<tr>
<td>Results or outcomes</td>
<td></td>
</tr>
<tr>
<td>Subjective evaluation</td>
<td>XXXX</td>
</tr>
<tr>
<td>Usefulness, etc.</td>
<td>XXXX</td>
</tr>
</tbody>
</table>
... but the data will be harmonised and linked

41. It will be necessary to harmonise the data by using the same definitions and categories for their collection.

42. Because there is not one ideal data source covering all aspects of CET, there will always be a need for linking the data coming from different sources in order to arrive at a complete picture of CET activities in a country.

43. The following example lists the data necessary for an estimate of the total financial investment in human capital represented by job-related CET in a given year. We would have to include the direct expenditures of providers and participants, and the opportunity costs incurred by both enterprises and individuals. The data would have to come from different sources:

- The providers (including the enterprises in their role as providers) are the best source for data on the direct expenditures for personnel (teachers, trainers and others), materials and facilities for job-related CET activities.

- To avoid double-counting, the expenditures of the participants for tuition and fees should be excluded, since they go towards covering the direct expenditures of the providers, but participants may have additional direct expenditures for other goods such as travelling and accommodation or additional child-care. The participants are the only source for data on these expenditures, even if they may not be a very reliable source.

- The enterprises are the best source of information on the opportunity costs of the foregone results of productive work of the employees in training.

- Among the opportunity costs of individuals are the foregone earnings for time spent in training instead of gainfully employed. For many of the employed, these costs are zero, since their training takes place during working hours. For others, they may be negligible, since the volume of individually sponsored job-related CET activities is low. Still others may have high opportunity costs, because they reduce their working time considerably. Opportunity costs related to the fact that the participants do not have time to do their own housework and have to buy the equivalent services are of an unknown magnitude. And of course, the value of the reduced leisure time is zero. For all these costs, the individuals are the best source, but getting reliable estimates could be very difficult.

44. It might also be necessary to complete the coverage of one or the other data source by making estimates. Enterprise data for instance are frequently missing for small enterprises. It would be feasible to estimate the volume of employer-supported CET provided by small enterprises using data from household surveys. It might also be possible to estimate the financial resources invested by making assumptions based on data coming from the larger enterprises.

The basic unit – the hour of participation

45. Synthetic statistics such as the CET account need a common unit of measurement as a basis for linking the data coming from different sources. For national accounts and their satellites, this basis would be monetary units; for labour accounts, the basis would typically be employed persons or employees converted into full-time equivalents. For the system of education indicators developed within INES, one very important measurement unit is the enrolled student, sometimes also measured in full-time equivalents.

46. For the CET account, the basic unit of measurement would probably be a measure of the time spent by participants in a given type of CET activity. The easiest measure of time would be hours of
participation, but these hours could be converted into larger measures such as days, weeks or years\textsuperscript{48} spent in continuing education activities.

Why is it necessary to use a measure of time?

1. Participants or the number of persons enrolled in continuing education activities must be ruled out, because of the danger of double counting when adding data from different providers.\textsuperscript{49} This is obviously less of a problem in countries with registers that also cover continuing education. But these countries are few.

2. Participation events i.e. the instances of persons or enrolled participants engaged in a continuing education activity can only be used as the basic unit for harmonisation if the different data sources agree on the definition of “an event”. Whether this is the case is a question for further research, but there are grounds for doubting the congruence of definitions. A language course taking place once a week for the whole year might be counted as one course by the participant, since she considers it as a continuing activity, but as four activities by the provider, since the course is organised on a quarterly basis (enrolment is for one quarter at a time, the necessary resources – teachers, class-rooms – are reassigned each quarter, fees are billed on a quarterly basis ...).

3. Hours of participation: Hours have the advantage of a commonly agreed-on social definition, on which all data sources could converge. Of course, the different data sources would have to be aligned on the issue whether the basic unit is the “common hour of 60 minutes” or the scholastic hour or lesson, with a duration of typically less than 60 minutes.

\textit{The basic source – the individual}

47. In the ideal system, all data sources would cover extensively the total range of CET activities and issues within one country; lacking this ideal, the coverage would at least be similar for all countries one wishes to compare. Unfortunately this is not the case.

48. Household surveys typically strive for “complete” coverage of private households, but very often the questions of participation in CET activities are asked only of persons within a limited age range. The missing coverage would typically involve persons living in institutions, younger and older persons.

49. The extent of the coverage of CET activities in provider based data collections is very difficult to assess. We would assume that the bigger providers and those for which the organisation of continuing education is the main activity would have a better chance of being included. The missing parts of the system would involve mainly smaller providers and those with other core businesses. The size of the missing piece may differ very much from one country to another.

50. Enterprise surveys typically do not cover the all enterprises. CVTS and CVTS2, the largest international CET enterprise surveys, are restricted to enterprises with more than 10 employees in the secondary and tertiary sectors, excluding public administration and the health and education sectors. Obviously, the surveys also leave out small businesses and the primary sector. Again, the size of the missing piece of all employer provided CET is different from country to country.

51. Household surveys tend to be the most comprehensive of the surveys. Individuals are reasonably good informants about themselves. They can also give indications about the sector and the size of the enterprise employing them, thus allowing an adjustment to the data coming from enterprise surveys.

\textsuperscript{48} See chart 1 for an example of the conversion of hours into years in education.

\textsuperscript{49} “Regular” education statistics have in principle the same problem, but the possibility of double enrolments is generally simply ignored.
Individuals can also give indications about the provider of the CET activities they engage in, thus allowing an adjustment of the individual data to data coming from data collections based on provider data. A fairly comprehensive survey of individuals would therefore be a basic prerequisite for an integrated CET account.

52. The linking and harmonisation of data coming from different surveys also depends on common categories and classifications used by the different surveys. The linking of individual and enterprise data is of course greatly facilitated by the existing classification of industries and size of enterprises and the solid evidence about the degree of precision one might expect from individuals and enterprises gained within labour market statistics. In the field of CET, this kind of experience with harmonised categories and classifications is totally lacking at an international level, and at a preliminary stage at the national level.

One step towards the solution - the NWB module

53. A survey of individuals is the basis for a national CET account, and a set of harmonised surveys of individuals is needed as the basis for internationally comparable CET accounts. Actually, an international set of harmonised individual CET data is also of value in and of itself. There is much to gain from the use of common and agreed-upon definitions and standards to measure CET. Data collected according to such common standards will provide a secure basis for comparative analysis and a better understanding of the role of CET in countries’ education and work-force development efforts. They will also provide a platform for more ambitious research into the outcomes of different learning activities and the factors that motivate adults to engage in learning.

54. In this context, network B of INES proposes to invest in a project known as the NWB CET module. The goal of this project is to develop operational guidelines for gathering statistics on adult participation in continuing education and training in the context of household surveys. The module would include examples of “best practice” and examples of core questions. The module is intended to be used by OECD countries as a supplement to existing national surveys, such as national household panels or labour force surveys. Designing this CET module to be implemented within a national “carrier” survey avoids the burden of an extra national data collection, and provides a vehicle for linking CET participation data to other important characteristics of the adult population (such as age, sex, education level, and employment status) that are typically assessed in existing national surveys.

55. The first steps towards the development of such a module include stock-taking of specialised national surveys on CET, the specification of formal international guidelines on training statistics and a diagnostic analysis of estimates from different national/international sources, with a view to identifying the nature of the measurement problems.

56. The concepts and definitions underlying the questions and answer categories within the module will be designed to provide internationally comparable data for key indicators on CET, including the level of CET participation, the intensity (duration) of participation, providers of CET, incentives and barriers to participation, and important types of participation (such as formal versus informal, required versus voluntary). The module will focus primarily on organised learning, which is relatively easy to measure and relatively amenable to public policy intervention. However, since capturing less formal or less organised types of learning is also of increasing policy interest, these forms of learning will also be considered for inclusion in the module.

57. With data collections based on the NWB module it will for instance be possible to complete Table 1. The chart shows that participation in formal CET courses has tended to be fairly stable in the

---

50 The different surveys might also be linked directly, as when enterprises and their employees are surveyed conjointly. These surveys tend to be limited in the same way as stand-alone enterprise surveys, thus still leaving a gap in the coverage.
late 1990s. Results for Switzerland, on the other hand, suggest an increase in the incidence of informal CET. It would be of interest to know whether a similar trend holds for other countries.

**Table 1.**  
Percentage of the employed population 25 to 64 years of age that has participated in job-or career-related continuing education and training.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation rate 4 weeks</td>
<td>Australia</td>
<td>-</td>
<td>-</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Austria</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Belgium</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Denmark</td>
<td>-</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Finland</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>15</td>
<td>-</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greece</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hungary</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ireland</td>
<td>-</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Luxembourg</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Norway</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>-</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
<td>-</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td>-</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Participation rate 12 months</td>
<td>Australia</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Canada</td>
<td>28</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Finland</td>
<td>41</td>
<td>-</td>
<td>45</td>
<td>44</td>
<td>-</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>-</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Germany</td>
<td>-</td>
<td>33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Switzerland</td>
<td>38</td>
<td>-</td>
<td>35</td>
<td>-</td>
<td>38</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Participation rate 6 months</td>
<td>Sweden</td>
<td>-</td>
<td>-</td>
<td>44</td>
<td>42</td>
<td>39</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Switzerland: informal types of CET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>47  59</td>
</tr>
</tbody>
</table>

**REFERENCES**


Hörner, Walter (1999),

O’Connell, Philipp (1999)


EDUCATIONAL DETERMINANTS AND OTHER ASPECTS OF THE TRANSITION PROCESS

Patrice de Broucker, Chief, Analysis and Dissemination Section, Centre for Education Statistics, Statistics Canada, Ottawa, Canada

Michel-Henri Gensbittel, University-Enterprise Unit, Université Paris I, Panthéon Sorbonne, France

Christine Mainguet, scientific attachée, Department of Studies and Statistics, Ministry for the Walloon Region, Namur, Belgium

Introduction

1. Transition is determined by individual educational pathways but also by the social and economic context prevailing at the time. This chapter reports on work as part of the INES project to improve indicators on youth access to the labour market, and to take stock of how education policies affect the transition between school and the workplace.

Are there effective transition or education policies?

2. There are several aspects to transition-related policies, including initial education, in particular vocational training in all its forms; continuing education, special arrangements for young people experiencing problems (underskilled, long-term unemployed, etc.); and incentives to hire young labour market entrants or take on trainees. The purpose of transition indicators is to help to identify those policies (in particular education policies) that are more effective than others.

3. There is probably no point in aiming for effectiveness by trying to fit the initial education system to the production system. Production is constantly evolving, and enterprises find it very hard to define their medium-term labour requirements in terms of quality or quantity. In any case, school-leavers will meet only part of those needs. Furthermore, career guidance and information procedures in today’s world are persuasive rather than directive. It also takes time to introduce a new programme of initial education, and several years will pass before the first students qualify and enter the labour market. Their initial education would obviously be out of sync with developments occurring over the forty years of their working lives. So the effectiveness of initial education should be judged not on its capacity to give immediate access to employment but on its capacity to equip individuals to learn and adapt on a lifelong basis.

4. An analysis of the indicators also shows how difficult it is, even when making allowance for the social and economic context in each country, to identify stable causal links that could suggest universally successful policies. What is clear is that, in any country, high educational attainment is an asset in transition. Even if it were possible to raise youth attainment sharply, however, transition problems would still not be solved, in the short term at least. Besides, many apparently similar policies clearly produce different outcomes.

51. Representing Canada in the INES project.
52. Formerly represented France in the INES project.
53. Formerly represented Belgium’s French-speaking Community in the INES project.
5. Apprenticeship is often cited as an effective policy and countries where it is common have fewer problems with access to the labour market. But a correlation is not necessarily a causal link. We must look at the reasons why apprenticeship is widespread and effective: apprentices are given training, to some extent by those in the trade; they hold jobs and receive wages. The capacity of the production system to invest in young people probably explains why apprenticeship is successful, and why transition status does not deteriorate as badly when the economic situation declines. But countries without any form of apprenticeship, even those with no school-based vocational pathways at all, also perform well when it comes to transition because enterprises invest differently, by hiring young people and training them on the job.

**Transition from initial education to working life**

6. The term “transition” is generally used to describe the complex process that leads young people out of initial education and into working life. The very word suggests a passage from an initial to a final state. Yet it is hard to suggest a precise definition that would be identical in every country (and operational in statistical surveys) for this transition from school to the world of work. Increasingly, but to differing degrees across countries, young people’s lives are being marked by alternating periods of work and education and by periods of combined work and study, and their status often changes. The borders between work and education are blurred.

**Transition is a process**

7. With regard to transition, we have opted to be very pragmatic and extend the period of observation from the end of compulsory schooling to the end of the longest training courses (generally in higher education). In other words, observing and analysing transition amounts to observing and analysing what happens to young people between the ages of 15 and 29, be it education and training, employment or unemployment. The word “process” indicates the emphasis in this analysis on consecutive or combined spells of education, work, inactivity and unemployment. This is also a time when the young experience other transitions (moving out, setting up home alone or with a partner), which may have some impact on initial job options and mean switches from the pathways they have in many cases chosen before leaving compulsory education.

**Components of the transition process**

8. Continuing education beyond compulsory schooling is a significant feature of the transition process. Educational attainment, measured in terms of the highest qualification achieved or years of schooling, is a highly discriminatory variable. Continuing education takes various forms (not all of them available in every country), including general education, school-based vocational training, apprenticeship, on-the-job training, and courses for young jobseekers.

9. The second feature of transition is work experience. This can be gained in a number of ways, i.e. time spent in work environment while still at school (usually organised by the school, but without a real contract of employment), holiday work (or part-time jobs while studying), work as an apprentice, and stable or insecure jobs held without training.

10. Many young people experience spells of inactivity or unemployment during the transition process. This is a feature of youth mobility, whereas adult mobility can mean moving from job to job, without leaving the labour market or becoming unemployed. But while these spells are more common in the youth population, their number and duration can be extremely variable between the ages of 15 and 29.

---

54. See for instance the various indicators on unemployment, labour force participation and employment rates based on these criteria in successive editions of *Education at a Glance - OECD Indicators*, and the chapter on labour market outcomes in this compendium.
A distinction should be made between intentional mobility (job shopping) and unintentional mobility (jobs lost when firms close down; fixed-term contracts); the latter is far more detrimental, and more frequent among younger workers. The distinction is not unrelated to educational attainment, in that young people whose qualifications are recognised on the labour market have more control over their career development than those with fewer educational resources.

**Can transition be described as successful?**

11. When discussing entry into the labour market, transition can be assumed to be a success when young people find jobs to suit their abilities and tastes, and will continue their education on a lifelong basis and become upwardly mobile. These criteria, however, are very hard to identify from survey data.

12. Other aspects also enter the equation when defining successful transition, and relate not only to the individual experiencing the transition but, more generally, to society at large. Depending on the circumstances, transition can be protracted, arduous and in the long run quite expensive, even if the individual views it as successful. It incurs a variety of costs, financial of course but social and human as well. These are costs to the individual but also to the community, such as the government funds that may be required to support preventive or remedial policies on transition, and initiatives relating to how education and training are organised and funded.

13. Attention also needs to be paid to the processes which lead some young people into marginalisation; low attainment or failure at school, repeated spells of unemployment and inactivity or a lack of vocational training may all gradually push people out of the labour market. While unfortunate circumstances upon entering the labour market can lead to marginalisation, it is reasonable to think that the education system, by failing to detect such individuals early or carefully enough, and failing to offer the right solution to their problems, is sending into the labour market young people who are particularly ill-equipped to find work. Those who leave school early and with no qualifications merit very close attention.

**Setting transition in a broader context**

14. The impact of the economic situation on transition is also highly significant. Everything is easier in a prosperous economy: substantial funds are available for different forms of education, jobs are readily available to the young, spells of unemployment are short, it is easier to gain work experience while studying, and so on. Transition is often more arduous when the economic climate is poor and when young people have to compete with older members of the labour force. Intergenerational balance may play a role here, but what counts is the attention that society pays to its youth, and its ability not to let young people bear the brunt of a worsening economic climate. Effective transition systems cannot go as far as to afford young people full protection from economic hardship, but they can certainly prevent them from being the first (or only) ones to shoulder the burden of a slump and pay the price of the ensuing adjustments.

15. Transition should also be envisaged from the standpoint of the welfare system, hiring practices and more generally corporate attitudes to skills (scope for continuing education, internal mobility), which vary across industries and countries alike. What is needed is a multidimensional approach to transition, one that does not recommend policies in situations where they are not likely to succeed. Typological work, based on indicators that describe both the institutional/economic background and individual pathways, is required here. Analysing indicators to identify effective policies should divide countries into at least three broad categories, i.e. those with widespread apprenticeship, those with well developed school-based vocational training, and those in which general education predominates.

---

55. The Final Report of the Thematic Review of the Transition from Initial Education to Working Life distinguishes between countries where apprenticeship is a major youth training pathway, i.e. Germany and Switzerland, and countries with a quasi-apprenticeship system, i.e. Austria, Denmark, the Netherlands and Norway.
Transition indicators

16. Addressing every aspect of transition and assessing its quality requires numerous indicators. The classic method is to look at indicators on education, labour force participation, unemployment and employment for each of the age groups relevant to transition (e.g. 15-24, 25-29). Although readily available, these indicators do have significant limitations. They are too static, and fail to describe what is essential to any analysis of transition, namely the dynamics, changes in status, the gradual accumulation of useful or unfortunate experiences. Measuring unemployment at any given time, for instance, does not indicate whether it is always concentrated on the same groups (experiencing long and/or repeated spells of unemployment), or whether the pattern is broader. It has therefore been necessary to develop new indicators.

Early indicators

17. Prior to 1996, most work consisted in calculating, for the age groups concerned by transition, the classic indicators used for adults on the labour market, in particular labour force participation and unemployment rates. This was to some extent imposed by the lack of documented and comparable databases. The information came mainly from Labour Force Surveys, which provide regular data with which to monitor cyclical developments and obtain stock data for each age group. But these did not, in every country, make it possible to determine how much time people had spent on the labour market (since completing their initial education, for instance).

18. A low labour force participation rate in the adult population indicates that a high proportion of people are not in work and do not wish to or cannot be, for instance because they are not adequately trained. In the youth population, a low participation rate predominantly reflects high educational enrolment. It is therefore an ambiguous indicator, in that substantial investment in education is likely to be beneficial in the medium term, though many young people also remain in education because jobs are hard to find in the short term.

19. Unemployment rates can sometimes be very high among the young, well above adult rates. In fact, when many young people remain in education, there are proportionally fewer young members of the labour force, many of whom have left school early, and they usually encounter numerous problems on the labour market where highly skilled young people abound. Table 1, based on recent data, gives a clear illustration of this in the 20-24 age group, and it is even more striking among 15-19 year-olds.

56. The definitions used here are in accordance with the International Labour Office guidelines: the employed comprise all persons who, for at least one hour during the reference period, performed some work for wage or salary, in cash or in kind (by convention this includes conscripted members of the armed forces). The unemployed comprise all persons who are without work, actively seeking employment and currently available to start work. The labour force participation rate defines all those who are employed or unemployed as a percentage of the total population; the unemployment rate defines the number of unemployed persons as a percentage of the labour force.

57. Also available were administrative data on qualifications, student ages (UOE Questionnaire) and a very small number of national surveys on those leaving initial education, but these were not harmonised and covered only a few levels of attainment. The surveys served as a basis for the indicators on youth unemployment one year and five years after completion of studies, published in successive editions of Education at a Glance.
Table 1: Labour force participation and unemployment among 15-19 year-olds and 20-24 year-olds in 1998

<table>
<thead>
<tr>
<th>Country</th>
<th>Participation rate 15-19 yrs</th>
<th>Unemployment rate 15-19 yrs</th>
<th>Participation rate 20-24 yrs</th>
<th>Unemployment rate 20-24 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>56.1</td>
<td>12.1</td>
<td>79.4</td>
<td>5.8</td>
</tr>
<tr>
<td>Switzerland</td>
<td>22.0</td>
<td>m</td>
<td>70.0</td>
<td>7.6</td>
</tr>
<tr>
<td>United States*</td>
<td>41.2</td>
<td>15.4</td>
<td>77.2</td>
<td>8.0</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>21.7</td>
<td>19.4</td>
<td>70.9</td>
<td>8.4</td>
</tr>
<tr>
<td>Portugal</td>
<td>25.9</td>
<td>12.9</td>
<td>68.4</td>
<td>9.3</td>
</tr>
<tr>
<td>Germany</td>
<td>31.3</td>
<td>7.7</td>
<td>73.3</td>
<td>10.0</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>61.7</td>
<td>15.5</td>
<td>76.0</td>
<td>10.1</td>
</tr>
<tr>
<td>Denmark*</td>
<td>57.9</td>
<td>2.8</td>
<td>79.4</td>
<td>10.5</td>
</tr>
<tr>
<td>Australia</td>
<td>56.6</td>
<td>19.4</td>
<td>82.7</td>
<td>12.0</td>
</tr>
<tr>
<td>Canada</td>
<td>41.8</td>
<td>21.3</td>
<td>71.3</td>
<td>14.6</td>
</tr>
<tr>
<td>Turkey</td>
<td>39.0</td>
<td>12.5</td>
<td>53.9</td>
<td>15.7</td>
</tr>
<tr>
<td>Belgium</td>
<td>7.4</td>
<td>28.2</td>
<td>56.4</td>
<td>17.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>21.9</td>
<td>22.6</td>
<td>62.3</td>
<td>18.6</td>
</tr>
<tr>
<td>France</td>
<td>7.7</td>
<td>23.2</td>
<td>49.7</td>
<td>24.8</td>
</tr>
<tr>
<td>Finland</td>
<td>31.2</td>
<td>46.4</td>
<td>65.1</td>
<td>29.0</td>
</tr>
<tr>
<td>Greece*</td>
<td>14.6</td>
<td>39.6</td>
<td>60.3</td>
<td>30.3</td>
</tr>
<tr>
<td>Italy</td>
<td>17.9</td>
<td>39.1</td>
<td>54.5</td>
<td>30.6</td>
</tr>
<tr>
<td>Spain</td>
<td>23.7</td>
<td>47.2</td>
<td>59.6</td>
<td>34.4</td>
</tr>
</tbody>
</table>

Source: OECD, Ines-Network B*, 1999 data
* 1997

20. Consequently unemployment rates in a particular age group are not easy to interpret in isolation, and do not give a broad picture of the status of young people at that age. As labour force participation rates indirectly reflect educational enrolment, the two sets of information have to be put together. Table 2 shows how education and work status can overlap. This explains why the sum of educational enrolment and labour force participation often exceeds 100%, particularly in the younger age group.

Table 2: Labour force participation and educational enrolment among 15-19 year-olds

<table>
<thead>
<tr>
<th>Country</th>
<th>Participation rate</th>
<th>Educational enrolment rate</th>
<th>Combined rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>56.1</td>
<td>87.1</td>
<td>143.2</td>
</tr>
<tr>
<td>Switzerland</td>
<td>55.2</td>
<td>85.5</td>
<td>140.8</td>
</tr>
<tr>
<td>Denmark*</td>
<td>57.9</td>
<td>79.9</td>
<td>137.8</td>
</tr>
<tr>
<td>Australia</td>
<td>56.6</td>
<td>76.9</td>
<td>133.5</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>61.7</td>
<td>67.8</td>
<td>129.6</td>
</tr>
<tr>
<td>Canada</td>
<td>41.8</td>
<td>83.0</td>
<td>124.9</td>
</tr>
<tr>
<td>United States*</td>
<td>41.2</td>
<td>82.6</td>
<td>123.8</td>
</tr>
<tr>
<td>Germany</td>
<td>31.3</td>
<td>91.6</td>
<td>122.9</td>
</tr>
<tr>
<td>Finland</td>
<td>31.2</td>
<td>87.0</td>
<td>118.2</td>
</tr>
<tr>
<td>Sweden</td>
<td>21.9</td>
<td>86.4</td>
<td>108.3</td>
</tr>
<tr>
<td>France</td>
<td>7.7</td>
<td>93.1</td>
<td>100.8</td>
</tr>
<tr>
<td>Spain</td>
<td>23.7</td>
<td>76.4</td>
<td>100.1</td>
</tr>
<tr>
<td>Portugal</td>
<td>25.9</td>
<td>73.6</td>
<td>99.5</td>
</tr>
<tr>
<td>Greece*</td>
<td>14.6</td>
<td>82.3</td>
<td>96.9</td>
</tr>
<tr>
<td>Belgium</td>
<td>7.4</td>
<td>85.3</td>
<td>92.8</td>
</tr>
<tr>
<td>Italy</td>
<td>17.9</td>
<td>74.8</td>
<td>92.7</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>21.7</td>
<td>67.3</td>
<td>89.0</td>
</tr>
<tr>
<td>Turkey</td>
<td>39.0</td>
<td>38.3</td>
<td>77.3</td>
</tr>
</tbody>
</table>

Source: OECD, Ines-network B, 1999 data
* 1997

58. Network B of the OECD INES Project is more specifically concerned with indicators on education and labour market destinations.
21. The sum of the two rates may be less than 100% when the young are neither employed, seeking work or enrolled in education (i.e. when they are marginalised, or when society does not expect young women to enter the job market). Figures in excess of 100% indicate that young people are both studying and working (or seeking employment). This combined status (employment and education) may be due to apprenticeship or alternating schoolroom/workplace training, or to a labour market that offers scope for young people to study while holding what is usually a part-time job. Apprenticeship and alternating training are well developed in Germany (the dual system), Switzerland, Denmark and the Netherlands. In the last two countries, there are also numerous jobs available to young students. All of these cases provide initial work experience, which should be taken into consideration when analysing the transition process.

22. Youth unemployment rates are also hard to compare directly. In countries with an apprenticeship system, apprentices are counted as employed and included in the unemployment rate denominator; in those with predominantly school-based vocational training, school pupils are not counted as participating in the labour force and so unemployment rates will be higher, other things being equal. Consequently international comparisons of youth unemployment are also complicated by these disparities (Table 1).

<table>
<thead>
<tr>
<th>Country</th>
<th>ISCED level</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Rep.</td>
<td>26.5</td>
<td>3.3</td>
</tr>
<tr>
<td>Hungary</td>
<td>18.0</td>
<td>2.5</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>19.8</td>
<td>2.9</td>
</tr>
<tr>
<td>United States</td>
<td>12.1</td>
<td>1.9</td>
</tr>
<tr>
<td>Sweden</td>
<td>21.2</td>
<td>3.4</td>
</tr>
<tr>
<td>Ireland</td>
<td>15.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Australia</td>
<td>12.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Germany</td>
<td>20.4</td>
<td>4.9</td>
</tr>
<tr>
<td>New Zealand</td>
<td>17.4</td>
<td>4.1</td>
</tr>
<tr>
<td>Netherlands</td>
<td>5.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Canada</td>
<td>18.5</td>
<td>4.8</td>
</tr>
<tr>
<td>Belgium</td>
<td>20.9</td>
<td>5.7</td>
</tr>
<tr>
<td>Finland*</td>
<td>23.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Austria*</td>
<td>10.9</td>
<td>4.2</td>
</tr>
<tr>
<td>Iceland</td>
<td>7.4</td>
<td>2.9</td>
</tr>
<tr>
<td>France</td>
<td>26.7</td>
<td>11.1</td>
</tr>
<tr>
<td>Poland</td>
<td>12.6</td>
<td>6.5</td>
</tr>
<tr>
<td>Korea</td>
<td>14.5</td>
<td>7.9</td>
</tr>
<tr>
<td>Norway*</td>
<td>9.1</td>
<td>5.3</td>
</tr>
<tr>
<td>Denmark</td>
<td>10.4</td>
<td>9.5</td>
</tr>
<tr>
<td>Spain</td>
<td>24.6</td>
<td>28.6</td>
</tr>
<tr>
<td>Italy</td>
<td>18.9</td>
<td>27.0</td>
</tr>
<tr>
<td>Mexico (1)</td>
<td>3.4</td>
<td>4.9</td>
</tr>
<tr>
<td>Greece (1)</td>
<td>13.0</td>
<td>22.8</td>
</tr>
<tr>
<td>Turkey</td>
<td>6.4</td>
<td>11.3</td>
</tr>
<tr>
<td>Portugal</td>
<td>5.1</td>
<td>8.1</td>
</tr>
</tbody>
</table>

Source: OECD database

* 1997

60. The full table with all the ISCED levels, for three age groups and by gender, is published in the 2000 edition of Education at a Glance - OECD Indicators.
23. Nor do high rates of employment necessarily imply successful transition. For a variety of reasons depending on the business cycle, many young people may be tempted to accept jobs that are easy to find but insecure or underskilled, thereby jeopardising their education and possibly even their future career. Hence the importance of looking beyond the classic indicators to grasp the realities of youth status from a more qualitative standpoint.

24. Finally, educational attainment obviously affects the transition from education to employment. Table 3 clearly shows that unemployment in the 25-29 age group (the least debatable indicator, since far fewer are enrolled in education) is almost always lower among the more highly skilled. The opposite is true, however, in a few countries (Spain, Italy, Mexico, Portugal, Turkey and Greece), where it may be that the young leave school early because work is available (there is still a large farming sector, for instance), and the more highly skilled are handicapped by their lack of time and experience on the labour market. In these countries the labour market cannot offer all higher-education graduates jobs to match their qualifications. In many cases, these young people are unwilling to accept less skilled employment.

25. Another limitation when interpreting stock data, even by age group, is that at a given age the more highly skilled young people who have spent longer in education have necessarily spent less time on the labour market than early school-leavers. The attainment effect is therefore blurred by the effect of time spent on the labour market.

**New avenues**

26. To overcome these limitations, members of the INES Project’s “Education and labour market destinations” network (national experts and suppliers/users of statistics on employment, training and education) decided to develop better, comparable summary indicators on the educational determinants of the transition process.

**New indicators**

27. One drawback of the classic indicators was that they did not always clearly or accurately cover the combined status of work and education. In autumn 1996 a subgroup of the “Education and labour market destinations” network therefore launched an original data collection, cross-classifying enrolment in an educational or training programme with labour market status. The labour force survey data on enrolment and employment are gathered in such a way as to allow cross-classification, making it possible to envisage better unemployment indicators and describe how the young gain initial work experience: via programmes alternating classroom teaching with practical experience, or via student employment (excluding summer jobs) not directly related to the content of initial education. Table 4 below illustrates this type of data for the 15-29 age group as a whole.

28. The same data were collected again in 1999. This time, distinctions were made based on the highest qualification obtained, for the first time using the new ISCED-97 classification. These are the data shown in the tables below.

29. This was a major innovation, providing an accurate picture of education and work status with no double counting, and allowing new indicators to be built and tested. The data were published in the 1998 edition of *Education at a Glance*. The first breakthrough was the calculation of unemployment rates for young people not enrolled in education (where cross-country comparison is significant), which were then compared with educational enrolment (including apprentices). The second, was that, in the absence of cohort data for every country, this information could be used to show that unemployment rates among 20-
24-year-old non-students are often lower in countries where a high percentage of the 15-19 age group both study and work, although the situation varies substantially across countries (Figure 1).

Table 4: Labour market and education status of 15-29 year-olds (1998)

<table>
<thead>
<tr>
<th>15-29 yrs</th>
<th>Students on work-study programmes</th>
<th>In education</th>
<th>Not in education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employed</td>
<td>Unemployed</td>
<td>Inactive</td>
</tr>
<tr>
<td>Germany</td>
<td>10.7</td>
<td>15.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Australia</td>
<td>20.3</td>
<td>2.4</td>
<td>16.9</td>
</tr>
<tr>
<td>Belgium</td>
<td>1.0</td>
<td>3.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Canada</td>
<td>15.5</td>
<td>2.7</td>
<td>25.9</td>
</tr>
<tr>
<td>Denmark*</td>
<td>26.0</td>
<td>4.1</td>
<td>14.5</td>
</tr>
<tr>
<td>Spain</td>
<td>0.1</td>
<td>3.6</td>
<td>3.9</td>
</tr>
<tr>
<td>Finland</td>
<td>11.5</td>
<td>8.9</td>
<td>33.2</td>
</tr>
<tr>
<td>France</td>
<td>3.5</td>
<td>0.3</td>
<td>44.3</td>
</tr>
<tr>
<td>Greece*</td>
<td>1.0</td>
<td>0.8</td>
<td>39.1</td>
</tr>
<tr>
<td>Italy</td>
<td>0.2</td>
<td>0.6</td>
<td>36.2</td>
</tr>
<tr>
<td>Netherlands</td>
<td>3.6</td>
<td>18.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Portugal</td>
<td>5.3</td>
<td>0.9</td>
<td>33.1</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>0.8</td>
<td>0.1</td>
<td>26.9</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>3.3</td>
<td>16.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>8.9</td>
<td>3.5</td>
<td>34.0</td>
</tr>
<tr>
<td>Switzerland</td>
<td>13.8</td>
<td>22.7</td>
<td>m</td>
</tr>
<tr>
<td>Turkey</td>
<td>1.6</td>
<td>0.2</td>
<td>18.9</td>
</tr>
<tr>
<td>United States*</td>
<td>17.8</td>
<td>1.9</td>
<td>23.9</td>
</tr>
</tbody>
</table>

Source: OECD, Ines-Network B, 1999 data

NB: m = too small to be reported, considering the sample size.

* 1997

Note: In spite of the progress made in data collection, as reflected in this table, disparities persist as to how apprenticeship is identified and allocated to one of the three categories of activity. For instance, the Danish data do not include apprentices in work-study programmes.

63. For a breakdown by age group and gender, see 2000 edition of Education at a Glance - OECD Indicators. This edition also contains a brief presentation of the same data showing the number of years spent in each status.
Figure 1 - Estimated link between the unemployment to population ratio of 20-24 year-olds not in education and the percentage of 15-19 year-olds in education who are employed

Source: OECD, Ines-Network B, 1999 data

30. The diagram above, published in Education at a Glance (1998) and updated here, shows primarily that in some countries enterprises offer young people jobs while they are studying (including apprenticeships) and when they have completed their studies. Yet it does not enable us to tell how effective education policies are: can we advise countries like Belgium, France, Italy or Greece (which have the fewest jobs for 15-19 year-olds) to introduce widespread apprenticeship as in Germany, Switzerland or the Netherlands, or to give up school-based vocational training, as in Japan or Korea which do so well without it? Since education policies can have little short-term influence on the economic environment or society’s attitude to youth, it seems reasonable to compare education policies within groups of countries that are similar in terms of these criteria, since across-the-board comparisons cannot be drawn. This argues in favour of a typological approach, less approximate than a comprehensive approach, that would be based on the data gathered since 1996. Below is an example, which needs further development.
Table 5: Typological approach

<table>
<thead>
<tr>
<th>Study</th>
<th>Education and work</th>
<th>Education but no work</th>
<th>Work</th>
<th>Unemployment</th>
<th>Not in the labour force</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>1.4</td>
<td>83.9</td>
<td>3.9</td>
<td>1.9</td>
<td>8.8</td>
</tr>
<tr>
<td>France</td>
<td>3.9</td>
<td>89.2</td>
<td>2.0</td>
<td>1.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Greece</td>
<td>0.7</td>
<td>81.6</td>
<td>8.1</td>
<td>5.1</td>
<td>4.6</td>
</tr>
<tr>
<td>Group A average</td>
<td>2.0</td>
<td>84.9</td>
<td>4.7</td>
<td>2.9</td>
<td>5.5</td>
</tr>
<tr>
<td>Finland</td>
<td>10.6</td>
<td>76.4</td>
<td>6.1</td>
<td>3.2</td>
<td>3.7</td>
</tr>
<tr>
<td>Sweden</td>
<td>11.4</td>
<td>75.0</td>
<td>5.5</td>
<td>1.7</td>
<td>6.4</td>
</tr>
<tr>
<td>Group B average</td>
<td>11.0</td>
<td>75.7</td>
<td>5.8</td>
<td>2.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Italy</td>
<td>0.2</td>
<td>74.6</td>
<td>10.8</td>
<td>6.5</td>
<td>7.9</td>
</tr>
<tr>
<td>Spain</td>
<td>1.7</td>
<td>74.6</td>
<td>10.8</td>
<td>8.7</td>
<td>4.1</td>
</tr>
<tr>
<td>Portugal</td>
<td>3.1</td>
<td>70.5</td>
<td>19.5</td>
<td>2.7</td>
<td>4.3</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>1.7</td>
<td>65.6</td>
<td>15.8</td>
<td>4.0</td>
<td>13.0</td>
</tr>
<tr>
<td>Turkey</td>
<td>1.7</td>
<td>36.6</td>
<td>32.5</td>
<td>4.7</td>
<td>24.5</td>
</tr>
<tr>
<td>Group C average</td>
<td>1.7</td>
<td>64.4</td>
<td>17.9</td>
<td>5.3</td>
<td>10.8</td>
</tr>
<tr>
<td>Canada</td>
<td>23.5</td>
<td>59.5</td>
<td>9.4</td>
<td>3.2</td>
<td>4.3</td>
</tr>
<tr>
<td>USA</td>
<td>24.5</td>
<td>58.1</td>
<td>10.3</td>
<td>2.2</td>
<td>4.9</td>
</tr>
<tr>
<td>Australia</td>
<td>31.3</td>
<td>45.5</td>
<td>14.3</td>
<td>6.1</td>
<td>2.7</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>30.2</td>
<td>37.7</td>
<td>22.0</td>
<td>5.7</td>
<td>4.4</td>
</tr>
<tr>
<td>Group D average</td>
<td>27.4</td>
<td>50.2</td>
<td>14.0</td>
<td>4.3</td>
<td>4.1</td>
</tr>
<tr>
<td>Switzerland</td>
<td>43.4</td>
<td>42.1</td>
<td>9.6</td>
<td>m</td>
<td>3.2</td>
</tr>
<tr>
<td>Denmark</td>
<td>44.7</td>
<td>35.2</td>
<td>11.8</td>
<td>0.8</td>
<td>7.8</td>
</tr>
<tr>
<td>Germany</td>
<td>23.8</td>
<td>67.8</td>
<td>5.0</td>
<td>1.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Netherlands</td>
<td>38.5</td>
<td>48.6</td>
<td>10.8</td>
<td>0.8</td>
<td>1.3</td>
</tr>
<tr>
<td>Group E average</td>
<td>29.2</td>
<td>56.6</td>
<td>9.4</td>
<td>1.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Source: OECD, Ines-Network B, 1999 data

31. In this typology, countries have been grouped according to the importance of employment among 15-19 year-olds depending on whether or not they are in education. An institutional factor (widespread apprenticeship) also entered the equation for Group E.

- Group A: High educational enrolment, very few young people in work or alternating training.
- Group B: High educational enrolment, but a significant proportion of young people work while studying.
- Group C: Lower educational enrolment, with some young people already employed (or inactive) in this age group.
- Group D: Numerous young people employed (over 40% on average), combined with study or not.
- Group E: Countries focusing on apprenticeship, but with jobs available to young people.

32. The new data also make it possible to build new unemployment indicators. To overcome the drawbacks of using the classic unemployment rate, two avenues have been explored. One is to calculate the unemployment rate for those not in education, and the other the unemployment to population ratio for those not in education. The second of these two indicators seems by far preferable, as can be seen from Table 6.
Table 6: Comparison of unemployment indicators for 15-19 year-olds\(^64\) (1998)

<table>
<thead>
<tr>
<th></th>
<th>Unemployment rate (unemployed/ labour force)</th>
<th>Youth participation rate</th>
<th>Unemployment to population ratio</th>
<th>Ratio of unemployed non-students to unemployed youth</th>
<th>Ratio of unemployed non-students to total youth</th>
<th>Ratio of non-students not-in-the labour force to total population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>12.1</td>
<td>56.1</td>
<td>6.8</td>
<td>11</td>
<td>0.8</td>
<td>1.3</td>
</tr>
<tr>
<td>Denmark*</td>
<td>2.8</td>
<td>57.9</td>
<td>1.6</td>
<td>47</td>
<td>0.8</td>
<td>7.8</td>
</tr>
<tr>
<td>Switzerland</td>
<td>8.1</td>
<td>22.0</td>
<td>1.8</td>
<td>76</td>
<td>1.4</td>
<td>3.6</td>
</tr>
<tr>
<td>Germany</td>
<td>7.7</td>
<td>31.3</td>
<td>2.4</td>
<td>65</td>
<td>1.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>22.6</td>
<td>21.9</td>
<td>4.9</td>
<td>34</td>
<td>1.7</td>
<td>6.4</td>
</tr>
<tr>
<td>France</td>
<td>23.2</td>
<td>7.7</td>
<td>1.8</td>
<td>100</td>
<td>1.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Belgium</td>
<td>28.1</td>
<td>7.4</td>
<td>2.1</td>
<td>92</td>
<td>1.9</td>
<td>8.8</td>
</tr>
<tr>
<td>United States*</td>
<td>15.4</td>
<td>41.2</td>
<td>6.4</td>
<td>35</td>
<td>2.2</td>
<td>4.9</td>
</tr>
<tr>
<td>Portugal</td>
<td>12.9</td>
<td>25.9</td>
<td>3.3</td>
<td>80</td>
<td>2.7</td>
<td>4.3</td>
</tr>
<tr>
<td>Canada</td>
<td>21.3</td>
<td>41.8</td>
<td>8.9</td>
<td>36</td>
<td>3.2</td>
<td>4.3</td>
</tr>
<tr>
<td>Finland</td>
<td>46.4</td>
<td>31.2</td>
<td>14.5</td>
<td>22</td>
<td>3.2</td>
<td>3.7</td>
</tr>
<tr>
<td>Czech Rep.</td>
<td>19.4</td>
<td>21.7</td>
<td>4.2</td>
<td>94</td>
<td>4.0</td>
<td>13.0</td>
</tr>
<tr>
<td>Turkey</td>
<td>12.5</td>
<td>39.0</td>
<td>4.9</td>
<td>97</td>
<td>4.7</td>
<td>24.5</td>
</tr>
<tr>
<td>Greece*</td>
<td>39.6</td>
<td>14.6</td>
<td>5.8</td>
<td>88</td>
<td>5.1</td>
<td>4.6</td>
</tr>
<tr>
<td>UK</td>
<td>15.5</td>
<td>61.7</td>
<td>9.6</td>
<td>60</td>
<td>5.7</td>
<td>4.4</td>
</tr>
<tr>
<td>Australia</td>
<td>19.4</td>
<td>56.6</td>
<td>11.0</td>
<td>56</td>
<td>6.1</td>
<td>2.7</td>
</tr>
<tr>
<td>Italy</td>
<td>39.1</td>
<td>17.9</td>
<td>7.0</td>
<td>93</td>
<td>6.5</td>
<td>7.9</td>
</tr>
<tr>
<td>Spain</td>
<td>47.2</td>
<td>23.7</td>
<td>11.2</td>
<td>77</td>
<td>8.7</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Source: OECD, Ines-Network B, 1999 data
NB: m = not available

33. The classic unemployment rate (Table 6, column a) does not, as we have seen, take educational circumstances into consideration. Consequently an unemployed young person counted in the numerator may, in some countries, also be enrolled in education. The denominator may include young people in vocational training, provided they are apprenticed, but not similar people on school-based vocational courses. Finally, when educational enrolment -- the main factor affecting youth employment -- is very high (see Table 2), the unemployment rate may also appear very high, and yet it concerns very few young people. This is the case in Spain, Finland, Greece, Italy, Belgium, France and Sweden, although the situation in all of these countries is not quite identical (column b). So to give a more realistic picture of the situation, it is better to indicate the ratio of unemployment to population, rather than to the labour force as has been the case to date (column c).

34. Taking one step further the search for a better indicator of the youth unemployment “problem”, young non-students can be viewed as the unemployed most at risk on the labour market since they cannot even hope to see their lot improve in the medium term. The ratio of unemployed non-students varies substantially across countries (column d). The reason is probably as much institutional as measurement-related. This kind of combined status (unemployment and education) may not be measured in quite the same way in every country, depending for instance on whether or not a full-time or part-time student is regarded as available for work (one of the three criteria defining unemployment).

35. We can then calculate the ratio of non-students unemployed to total population (column e). These ratios lend themselves more easily to international comparison than classic unemployment rates, since they no longer include young people who are both working and studying, or the effects of apprenticeship on the unemployment rate. The indicator shows, for a given age group, how some young people could become unemployed if they were not enrolled in education.

\(^64\) For a breakdown by age group and gender, see 2000 edition of Education at a Glance - OECD Indicators.
people eventually come to be marginalised. In the United Kingdom, for instance, relatively low unemployment rates mask the fact that the proportion of young people out of work and receiving no education is as high as in Italy, for instance, and much higher than in Finland, two countries where classic unemployment rates are twice or three times as high. When defining transition policy, it is probably one of the most important indicators to consult; in our view this measurement of unemployment is the best way of grasping the real scale of youth unemployment as a policy issue.

36. This indicator can also be supplemented with information on young people who are neither in the labour market nor in education, often considered to be an indicator of marginalisation among specific groups. Care is needed here, however, for this status may be temporary or intentional and not necessarily a real sign of marginalisation. It is fair to say, nevertheless, that this is not a normal or favourable situation to be in at such a young age.

37. The 1999 data have made it possible for the first time to gather more qualitative information on the jobs held by young people: part-time or full-time, temporary or permanent. New indicators based on these data are included in the 2000 edition of Education at a Glance. They measure the importance of part-time and temporary work among 20-29 year-olds compared with averages for the population as a whole. Temporary jobs appear to be more common among young people only in countries where the labour market offers them little access. Part-time jobs are very often less common among young people than in the population as a whole, with the exception of a few countries where there is very little part-time work and its development appears to concern new young labour market entrants.

**Outlook**

38. In light of the lessons learnt in recent years, we believe that several avenues should be given priority.

**Considering the economic and social context**

39. It is crucial to take into account economic and social disparities, together with distinctive institutional features, when assessing the impact and relevance of initial education policies. For instance:

- An education policy such as apprenticeship is often deemed effective but may be hard to implement in a country where there are few jobs for the young and average skill requirements are low, for not many firms will be able to take on and train new labour market entrants. In countries like this, are some initial education policies more appropriate than others?

- Will certain forms of vocational training that are well suited to current skill requirements allow further economic development if the young people thus trained do not have sufficient general knowledge to evolve and train on a lifelong basis?

- Early contact with the world of work is often thought to be beneficial to young people. But rapid access to employment, when the labour market allows it, can be detrimental to training. In this case, is training policy for young wage-earners a good solution, allowing them to acquire the training they will one day need?

- Some countries have training systems based predominantly on general education. Yet they do not always have poor youth unemployment indicators. Would it still be in their interest to introduce more vocational training?

40. These are only a few of the questions raised. Clearly they can be answered only by taking into consideration the economic and social context and trends, and identifying scope for relevant comparisons. It is also advisable to avoid hasty generalisations about policies that have proved successful in some countries but cannot be transposed to others without economic and social changes that go well beyond the scope of education policymaking, and would need to be clearly identified.
41. The typological approach outlined briefly here should therefore be pursued. A vital criterion is the capacity of society to provide opportunities for its youth, and to invest -- even on a long term basis -- in young people by offering them education and employment.

42. This breakthrough is feasible in the relatively short term by building on available indicators and on the findings of the thematic review of transition. Furthermore, current and future work as part of the INES project on upper secondary (vocational) institutions is expected to provide useful insight into the national aspects of education pathways and relations between schools and the world of work. Also needed, however, is further information, using a new data collection system, on employers’ labour requirements and in particular sector-specific hiring practices, hiring opportunities for young labour market entrants, and trade-offs between the internal and external labour markets.

**More insight into the transition process**

43. We have placed considerable emphasis on the fact that transition is a process, in which work and education are complementary. In our view, consideration has to be given to the consecutive and combined experiences of transition if we are to gain insight into the joint impact of initial education and early work experience. It is clear, for instance, that in order to measure the extent to which early contact with the labour market (through alternating training or student jobs) affects the likelihood of finding work, depending on the level and type of attainment, data must be available on comparable cohorts in every country.

44. It is accordingly not only outcomes that need to be recorded (such as unemployment and employment one year and five years after completion) but also information on pathways, including initial education (level, type and field), first contact with the world of work, number of jobs held, average duration of contracts, number of months of unemployment and time spent out of the labour market in a given period, labour conditions in successive jobs, occupations, and periods of training. Surveys of this kind also generate more qualitative information, such as satisfaction/dissatisfaction with current status, career plans, and retrospective views on individual educational pathways.

45. The fact should be faced that building such tools to gather information on youth pathways through the education system and in the labour market will be an arduous task. Apart from the problem of developing tools that are comparable across countries, it should be borne in mind that even devising them requires substantial investment. As we have already indicated, it is hard to pinpoint exactly when initial education ends, it is now increasingly common to return to education, and people’s entire education cannot be summarised by the point at which they leave the initial education system. One solution might be to monitor panels of young people over many years, but results will be long in coming and entry into employment is staggered over time, during which the economic climate may change. Another approach is to draw up samples of young people who, at the same time, completed educational courses giving access to the labour market, and to ask them for a retrospective view of their educational pathways and experiences of employment, unemployment or return to education.

46. Besides this long-term work, the shorter-term prospects are interesting now that questions have been included in European labour force surveys on the date at which people obtained their highest qualification (and the date at which they left initial education, in the ad hoc module on transition). This will shortly enable new indicators to be built showing, in a sufficient number of countries, the amount of time spent on the labour market and the date of exit from the initial education system.

**Broadening the scope of the indicators**

47. Within the perspective we have just outlined, an effort needs to be made to supplement analytical work with additional information.
This is particularly true of wages, which are viewed as a priority indicator but have not yet been given the importance they deserve. Some experimental data collections launched in 1999 still cover only a small number of countries. They aim, for instance, to obtain information on youth wages one year after graduating, for those leaving upper secondary and post-secondary non-tertiary education ISCED (3/4), tertiary type B (5B) and tertiary type A and advanced research programmes (5A/6), and to identify the labour market status of recent graduates from higher education (5A/6), according to their fields. These sets of data, based on monitored graduate cohorts, are to be extended in the future.

Although it is important to take into account initial education policies, consideration must also be given to the potential and capacity of young people to undertake lifelong learning, and one way of analysing this is with indicators on their enrolment in continuing vocational training or second-chance education, depending on their initial education and the type of transition they are engaged in. This will mean looking at the capacity of skilled youth to master new technologies, which are substantially transforming corporate skill requirements, and monitoring their progress.

But extending the scope of the indicators in this way may also mean including new, cross-cutting themes. For instance, greater emphasis will be placed on the notion of equity when developing new indicators by including disparities in social background as well as gender differences. We suggest the development of relative indicators comparing, for each country, the status of younger and older people, women and men, etc. In the same vein, the network will be focusing in years to come on 20-24 year-olds who are no longer in education and have not attained upper secondary level. Experts in the INES network are working on the collection of a new set of data on this target population, ensuring that it includes the factors behind early exit from the school system, in terms of both the context and remedial arrangements prevailing in each country.

REFERENCES


OECD (1998), The transition from school to work, in Education at a Glance – OECD Indicators.

NEW INDICATORS FOR ECONOMIC RETURNS TO EDUCATION

Zhongren Jin, American Institutes for research, United States

A Brief history of the study

Background

1. The economic returns to education have long been a concern not only to educators and economists, but also to the general public and policy makers. While few doubt the positive impact that education has on the economy, there are heated debates in many OECD countries about the level of social resources that should be devoted to education, the allocation of finite resources to different areas of education, and, more broadly, the sharing of the costs and benefits of education among individuals, the government, and society. To address these policy concerns, in 1996, the OECD Ministerial Council requested that OECD use existing data to conduct a study on the economic returns to investment in education, with a particular emphasis on economic returns to government expenditures on education. Network B of OECD’s Indicators of Education Systems (INES) project assumed responsibility for conducting the study. Ten countries, including Australia, Belgium, Canada, Denmark, France, Germany, Norway, Portugal, Sweden, and the United States, participated in the study. The National Center for Education Statistics (NCES) of the US Department of Education provided technical support.

2. In the past three years, this study has gone through two developmental phases. The first phase began in 1996 and ended in 1998. The current phase began in 1999.

The first phase of the study

3. The first phase of the study focused on estimating the internal rates of return (IRR) to education. The internal rate of return to education may be interpreted as an annual rate of monetary returns to the expenditure on education. Building on previous work done by Alsalam and Conley (1995), Network B designed a conceptual framework for estimating the IRR. Using this framework as a guide, a data collection instrument was developed and distributed to OECD countries for gathering the data. Each participating country submitted information concerning earnings, property income, social transfers, and income taxes for population cohorts in different age, gender, and education attainment categories. Countries also submitted information about individual and government expenditures on education. A cost-benefit analysis then was performed to estimate the internal rates of return to education from three different perspectives -- individual, governmental, and societal. Table A shows selected findings from this phase of the study, which has been reported in Education at A Glance (OECD, 1998).
Table A. Estimates of private, fiscal and social rates of return to education at university tertiary level, by gender: 1995

<table>
<thead>
<tr>
<th>Country</th>
<th>Private</th>
<th>Fiscal</th>
<th>Social</th>
<th>Private</th>
<th>Fiscal</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>14</td>
<td>10</td>
<td>11</td>
<td>21</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Belgium</td>
<td>14</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Canada</td>
<td>14</td>
<td>7</td>
<td>9</td>
<td>21</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Denmark</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>France</td>
<td>20</td>
<td>11</td>
<td>13</td>
<td>28</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Sweden</td>
<td>—</td>
<td>6</td>
<td>9</td>
<td>—</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>United States</td>
<td>11</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>9</td>
<td>11</td>
</tr>
</tbody>
</table>

—Not available.


4. The major advantage of estimating the internal rate of return to education is that this indicator provides summary information about economic benefits that education brings to individuals, society, and governments. The internal rate of return is a widely used and closely examined indicator in the economic and business world, especially when choices among important investment alternatives are available. However, using IRR as an education indicator poses some limitations. First, as an economic concept originated from human capital theory, the IRR is not fully accepted by society in general and by educators in particular. For some individuals, it does not seem appropriate to consider the financial returns to education as a measure of the success or value of education. Second, estimation of the internal rates of return to education involves various assumptions, some of which are very strong or unrealistic. For example, to estimate the IRR to education, it is necessary to assume the number of years required for a particular level of education, the typical age of students, the average costs of attaining that particular level of education, the typical length of a person’s working life, etc.\(^{65}\). The IRR estimates thus obtained may be contingent on the assumptions made, therefore the magnitude of the IRR estimates may not have much intuitive meaning. These complications obscure the policy significance of the IRR estimates as an education indicator.

The second phase of the study

5. In view of the limitations of the IRR analysis, Network B decided in early 1999 modify its approach to developing indicators to measure the economic returns to education. Based on suggestions made by the OECD secretariat, Network B “unpacked” the various components of the IRR analysis. Instead of estimating the IRR to education, the emphasis now is on examining the specific economic benefits from education that are combined to derive the IRR estimates. By focusing on specific factors such as average annual earnings, property income, social-transfers income, as well as income taxes, the new approach minimizes reliance on complex economic concepts or unrealistic assumptions. The findings from this approach are easier to understand and the policy implications of the findings are more intuitive, particularly when making comparisons across countries with different educational systems and social structures.

6. As in the IRR analysis, this new approach also involves analyses from the individual, societal, and government perspectives. The choice of perspective has implications for which specific costs and benefits will be involved in the analysis. If the perspective of an individual is chosen, education costs will

\(^{65}\) One observed trend in education among OECD countries is that education is increasingly becoming a life-long experience. More older students are going back to school for further education, and a large proportion of them enroll as part-time students while continuing to work. The existing IRR model cannot handle situations like these.
include private direct expenditure on tuition and fees, after-tax foregone earnings, reduced annual income from social transfers, and increased income taxes. Education benefits will include higher salaries, higher probability of being employed, and increased annual income from property. If a governmental perspective is chosen, education costs will include government direct and indirect expenditures on education, and lost income tax revenues on students’ foregone earnings. The benefits will include increased revenues from income taxes, and decreased expenditures on social transfer programs. In this analysis, we included all three perspectives. The reason for expanding the analysis to include government and societal perspectives is that much is already known about the individual perspective from the many studies of economic returns focusing on this perspective; little is known about the government and social perspectives, as few studies have examined the issue from these perspectives. This lack of research is largely due to the difficulty in obtaining appropriate data. Given that the OECD has a rich source of information about earnings, property income, social transfers, and income taxes from participating countries, the OECD is uniquely positioned to examine the economic returns to education from the governmental perspective.

7. In September 1999, preliminary findings from the second phase of the study were reported to Network B at its annual plenary meeting in Athens, Greece. The remainder of this chapter describes the current study and its findings, and then lays out an action plan for Network B to continue the study in the years to come.

**The impact of education attainment on various sources of income**

8. This study examines the economic returns to education from individual, governmental, and societal perspectives. To achieve this objective, various sources of people’s income were included in the examination. In past studies, two indicators were typically used to measure the economic benefits of education: the average annual earnings of people with different education attainment levels, and the labor force participation and employment rate of people with different education attainment levels. Although both indicators reflect important aspects of the economic benefits of education, they mainly focus on the economic benefits from individuals’ perspectives and do not reflect the education benefits from societal and governmental perspectives. For example, without including social-transfers income and income taxes, the existing indicators overstate individuals’ economic benefits from education and neglect economic contributions that education brings to the society as a whole.

9. To overcome these limitations, this study included a broader set of income measures to reflect various aspects of the economic benefits of education. In addition to average annual earnings, property income was included in the analysis. Observations suggest that individuals with more education not only receive higher average annual earnings, but also have more resources and greater ability to invest. In the long run, they may obtain more income from property assets. This study provided an opportunity to test this hypothesis empirically. This also expands traditional analyses from the individual perspective. Another source of income included in the study is income from social transfers. Findings using traditional earned-income measures suggest that persons with less education are more likely to rely on social-transfers income. This study provides empirical data to test this hypothesis. Finally, this study also includes data on individuals’ contribution to government income tax revenues. Given that all OECD countries have a progressive income tax system, it was hypothesized that persons with more education pay more income taxes in both absolute and relative terms. This study used empirical data to test this relationship.

10. In summary, this study examined the impact of education on individuals’ average annual earnings, property income, income from social transfers, and income taxes. These items collectively reflect the most important aspects of the economic returns to education, and they provide a more complete account of the economic returns to education. By including social transfers income and income taxes into the analysis, this study explicitly examined economic returns to government expenditure on education. Although this analysis did not include an indicator reflecting the relationship between education and labor force participation and employment, the findings from this study implicitly reflect the impact of education.
on labor force participation and employment. Later in this chapter when we discuss the findings, we will further elaborate on this point.

11. A caveat. Education brings many benefits to individuals as well as to society. However, not all benefits of education can be measured in monetary terms. For example, statistics show that better educated people live longer. Unless this longer life span extends the length of an individual’s working life, the benefit would not necessarily be reflected in economic returns. A society with a higher level of education also may enjoy a higher degree of social cohesion. But unless the higher level of social cohesion is reflected in higher levels of economic activity, this benefit could not be easily measured in economic terms. To the extent that this study focuses on economic benefits and overlooks the non-monetary benefits of education, this analysis underestimates the comprehensive benefits of education.

Data and design of the study

12. The data used in this study were collected from ten participating OECD countries. However, after careful check of the accuracy and reliability of the data, data from eight countries were used in the analysis reported here. The relevant government agencies that are responsible for gathering education statistics in each country provided information about population size, number of students, aggregated annual earnings, aggregated property income, income from social transfers, and income taxes. For each of the data items, the information was provided by age-range, gender, and education attainment (or enrollment) level.

13. This study focused on three education attainment levels: education below the upper secondary level, upper secondary level, and university level and above. Upper secondary education was used as the reference level for all comparisons. In other words, people with either less-than-upper-secondary education or university education were always compared to people with upper secondary education when examining the economic benefits of education. To avoid the complicated problem of comparing across different national currency units, indices of income were created for persons with various levels of education attainment. The income index for people with upper secondary education attainment was always set to equal 100. Income indices for other levels of education attainment were expressed as percentages of the income for people with upper secondary education. For example, if the average annual earnings for people with below upper secondary education are 78 percent of that for people with upper secondary education, the average annual earning index for people with less-than-upper-secondary education is 78. Similarly, if the average annual earnings for university graduates are 125 percent of that for upper secondary graduates, the relevant index for people with university education is 125.

14. The analysis focused on the population cohort aged 35-to-44, because the majority of this cohort would have completed their formal education credentials, and would have been in the prime of their working career. At this stage of their working life, the majority of the cohort should have a stable job, accumulated working experience, and could expect a stable income stream. Put another way, this population cohort is at its prime for reaping the benefits of education. The economic benefits obtained by this cohort provide a good indicator for the economic benefits of education enjoyed by the population as a whole.

15. The economic benefits of education were examined for both men and women. As the first step of the study, the examination of education benefits by gender was largely concentrated on within-gender differences rather than cross-gender differences. However, since education and economic equity are always important policy concerns, a more thorough cross-gender comparison of education benefits should be conducted in the future.

Preliminary findings

16. In this section, preliminary findings are reported. Interpretations of the findings are also provided, and policy implications are discussed.
Earnings are positively correlated with education attainment level

17. Figure 1-1 shows the relationship between individuals’ average annual earnings and their education attainment level. For the seven OECD countries with data available, for both men and women, the average annual earnings for persons with less than upper secondary education are significantly lower than the average annual earnings for their upper secondary counterparts. On the other hand, the average annual earnings for persons with university education are much higher than the average annual earnings for their upper secondary counterparts in all of the eight countries with data. There is a strong positive relationship between average annual earnings and education attainment level.

**Figure 1-1.** Mean annual earnings of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean annual earnings of individuals with a secondary education, by gender and country (Earnings for people with secondary education=100)

**Men**

![Graph showing mean annual earnings for men across different educational levels in various countries]

**Women**

![Graph showing mean annual earnings for women across different educational levels in various countries]

*Note: Data for people with less than upper secondary education were not available for Australia.*
18. Table 1-1 shows that, compared to the earnings of their counterparts with an upper secondary education, the average annual earnings for men with less than an upper secondary education range from 67 percent in the United States to 88 percent in Belgium. For women, the same comparisons range from 62 percent in Canada to 79 percent in Norway. On the other hand, people with a university education earn more than that earned by those without a university education. Table 1-1 shows that, compared to the earnings of their counterparts with an upper secondary education, the average annual earnings for men with a university education range from 142 percent in Belgium to 189 percent in France. The same relationship also holds for women.

**Table 1-1. Mean annual earnings of individuals 35 to 44 years old with various levels of educational attainment**

As a percentage of mean annual earnings of individuals with an upper secondary education, by gender and country (Earnings for people with an upper secondary education=100)

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th></th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less than upper secondary</td>
<td>Upper secondary</td>
<td>University</td>
</tr>
<tr>
<td>Australia</td>
<td>---</td>
<td>100</td>
<td>162</td>
</tr>
<tr>
<td>Belgium</td>
<td>88</td>
<td>100</td>
<td>142</td>
</tr>
<tr>
<td>Canada</td>
<td>82</td>
<td>100</td>
<td>147</td>
</tr>
<tr>
<td>Denmark</td>
<td>74</td>
<td>100</td>
<td>147</td>
</tr>
<tr>
<td>France</td>
<td>81</td>
<td>100</td>
<td>189</td>
</tr>
<tr>
<td>Norway</td>
<td>82</td>
<td>100</td>
<td>149</td>
</tr>
<tr>
<td>Sweden</td>
<td>86</td>
<td>100</td>
<td>173</td>
</tr>
<tr>
<td>United States</td>
<td>67</td>
<td>100</td>
<td>183</td>
</tr>
</tbody>
</table>

—Not available.

19. If we define a positive earnings differential as an earnings premium, in five of the eight countries, including Australia, Canada, France, Norway, and the United States, women with a university education enjoy a higher earnings premium than that enjoyed by their male counterparts. For example, in the United States, women with a university education had a 103 percent earnings premium over their upper secondary counterparts, while men with a university education only had an 83 percent earnings premium. Similarly, in France, the earnings premiums for men and women with a university education are 89 and 122 percent, respectively, over their counterparts with an upper secondary education.

20. One point worth mentioning is that the earnings premiums are a product of both a higher wage rate and more stable employment. Much research shows that higher educational attainment is related to both increased earnings and lower unemployment rates. While women who want to work may seek more education, they may also become more productive at their work when they gain more education. Women with university education are more likely to have a full-time, decent paying professional job than their less educated counterparts (OECD, 1998(2), Pole, 1995). Nevertheless, women in general, are more likely to work part time or work on temporary basis than their male counterparts. Many other factors, such as differences between men and women in college major, or career and occupational choices may also explain the earnings difference. For example, women without a college education are more likely in lower paying jobs than men without a college education (mainly the low-paid service jobs, rather than higher paying trades jobs), while college educated men and women have more equitable incomes—this would lead to a greater earnings premium for women. Furthermore, child-bearing and rearing obviously affect earnings capability. Since women without a college education tend to have more kids than women with a college education, the earnings premium for women with university education will be even larger than one would expect based on education level alone Further studies are required to determine the various causes that lead to a larger earnings premium for better educated women.
Property income is positively correlated with educational attainment level

21. Among the OECD countries with data available, property income as a percentage of a person’s gross income (i.e., the total income from all sources) ranges from roughly 3 percent in the United States to roughly 12 percent in Belgium. Figure 2-1 shows that in all the OECD countries with data available, the higher the individual’s education attainment level, the higher the individual’s property income.

Figure 2-1. Mean annual property income of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean property income of individuals with a secondary education, by gender and country (Property income for individuals with an upper secondary education=100)

Men

Women

NOTE: Data for people with less than upper secondary education for Australia were not available. Data for people with less than upper secondary education for Canada were not reliable because of small sample size.

22. Table 2-1 provides information about the magnitude of the property income premiums of education. Specifically, the premiums in property income obtained by men with a university education over their counterparts with an upper secondary education range from 50 percent in the United States to 289 percent in Australia. The same premiums for women range from 55 percent in Denmark to 151 percent in Australia. In six of the eight countries involved in the study, including Australia, Belgium, Denmark, France, Norway, and Sweden, men with a university education enjoy a higher premium in property income than their female counterparts. Examination of the data presented in Tables 1-1 and 2-1
seems to indicate that the education premiums for property income are higher on average than those for annual earnings.

Table 2-1. Mean annual property income of individuals 35 to 44 years old with various levels of educational Attainment as a percentage of mean property income of individuals with a secondary education, by gender and country (Property income for individuals with an upper secondary education=100)

<table>
<thead>
<tr>
<th>Non-students ages 35-44</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less than upper secondary</td>
<td>Upper secondary</td>
</tr>
<tr>
<td>Australia</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>Belgium</td>
<td>84</td>
<td>100</td>
</tr>
<tr>
<td>Canada</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>Denmark</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>France</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>Norway</td>
<td>64</td>
<td>100</td>
</tr>
<tr>
<td>Sweden</td>
<td>85</td>
<td>100</td>
</tr>
<tr>
<td>United States</td>
<td>82</td>
<td>100</td>
</tr>
</tbody>
</table>

—Not available.

23. There are two possible sources for the premiums in property income received by the better educated. The first of the possible sources is social inheritance. Empirical data suggest that people with higher levels of education tend to come from families of higher socio-economic status. Better educated people may inherit more property assets from their parents than those with less education, thus giving them an edge in their ability to receive income from property assets. The second possible source for the premiums in property income is that the better educated may have more resources to invest in property assets, which results in more property income in the long run. In terms of resources for investment, the better educated not only have more financial resources than the less educated due to their higher average annual earnings, they may also have a greater willingness or ability to invest in property as a result of their education. Strictly speaking, only the property-income premiums from this second source can be considered as an economic benefit of education, while the premium from the first source is merely social inheritance. However, practically it is not easy to attribute the premiums in property income to their original sources. Data on inheritance income would be needed to help disentangle these sources.

The less educated are more likely to rely on social-transfers income

24. Social transfers and income taxes are two major means for society to redistribute income and to improve economic equity among various social groups. As such, one would expect to find that the less educated (who have lower earnings) would receive more social transfers and pay less income taxes than those with more education. Figure 3-1 shows that there is a strong negative relationship between individuals’ social-transfers income and their education attainment level. The higher an individual’s education-attainment level, the less an individual’s income from social transfers.
Figure 3-1. Mean annual social transfer income of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean social transfer income of individuals with a secondary education, by gender and country (Transfer income for individuals with an upper secondary education=100)

Men

[Bar chart showing social transfer income by gender, country, and level of education for men]

Women

[Bar chart showing social transfer income by gender, country, and level of education for women]

Note: Data for people with less than upper secondary education for Australia were not available.

25. Table 3-1 shows that, compared to men with an upper secondary education, the average social-transfers income for men with less than upper secondary education ranges from 111 percent in Sweden to 260 percent in the United States. The similar comparisons for women range from 116 percent in France to 296 percent in the United States. On the other hand, persons with a university education receive much less social-transfers income than their counterparts with upper secondary education. For example, the average social-transfers income received by men with university education range from 29 percent in Australia to 90 percent in Belgium. The same comparisons for women range from 40 percent in the United States to 93 percent in Belgium.
Table 3-1. Mean annual social transfer income of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean social transfer income of individuals with a secondary education, by gender and country (Transfer income for individuals with an upper secondary education=100)

<table>
<thead>
<tr>
<th>Non-students ages 35-44</th>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>29</td>
<td></td>
<td>100</td>
<td>49</td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>127</td>
<td>100</td>
<td>90</td>
<td>126</td>
<td>100</td>
<td>93</td>
</tr>
<tr>
<td>Canada</td>
<td>140</td>
<td>100</td>
<td>46</td>
<td>140</td>
<td>100</td>
<td>70</td>
</tr>
<tr>
<td>Denmark</td>
<td>200</td>
<td>100</td>
<td>62</td>
<td>142</td>
<td>100</td>
<td>77</td>
</tr>
<tr>
<td>France</td>
<td>126</td>
<td>100</td>
<td>74</td>
<td>116</td>
<td>100</td>
<td>77</td>
</tr>
<tr>
<td>Norway</td>
<td>143</td>
<td>100</td>
<td>40</td>
<td>120</td>
<td>100</td>
<td>82</td>
</tr>
<tr>
<td>Sweden</td>
<td>111</td>
<td>100</td>
<td>56</td>
<td>126</td>
<td>100</td>
<td>87</td>
</tr>
<tr>
<td>United States</td>
<td>260</td>
<td>100</td>
<td>38</td>
<td>296</td>
<td>100</td>
<td>40</td>
</tr>
</tbody>
</table>

—Not available.

26. Not only do the less educated receive, on average, a larger absolute amount of social-transfers income, but also the social-transfers income composes a large percent of their gross income. According to Table 3-2, for all the countries with data available, for both men and women, the lower the individual’s education attainment level, the higher the percentage of gross income that comes from social-transfers income. The percentages of social-transfers income for persons with less than upper secondary education are more than double the percentages for their counterparts with university education in every country, for both men and women. For example, in the United States, social transfers constitute less than one percent of gross income for women with a university education, while the relevant percentages for women with upper secondary education or less than upper secondary education are 3.4 and 14 percent, respectively. Similarly, in France, the relevant percentages for women with university education, upper secondary education, or less than upper secondary education are 4.5, 12.1, and 19.9 percent, respectively. The same pattern holds for men.

27. Several observations can be drawn from these data. First, the less educated seem more likely to receive a larger amount of social transfers income than their better educated counterparts. Second, since income from social transfers composes a much larger percentage of their gross income, the less educated seem to rely more on social-transfers income than their better-educated counterparts. Third, women are more likely to rely on income from social transfers than men at all education attainment levels. For example, Table 3-2 shows that social transfers income as a percentage of gross income for women is always larger than that for men, and is often more than double the relevant percentages for men.

Table 3-2. Mean annual income from social transfers as a percentage of gross income

<table>
<thead>
<tr>
<th>Non-students ages 35-44</th>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5.5</td>
<td>1.0</td>
<td></td>
<td>11.7</td>
<td>3.0</td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>11.9</td>
<td>8.5</td>
<td>5.5</td>
<td>22.8</td>
<td>14.3</td>
<td>10.7</td>
</tr>
<tr>
<td>Canada</td>
<td>9.4</td>
<td>5.8</td>
<td>1.7</td>
<td>21.4</td>
<td>11.0</td>
<td>4.5</td>
</tr>
<tr>
<td>Denmark</td>
<td>14.3</td>
<td>5.8</td>
<td>2.5</td>
<td>28.4</td>
<td>17.3</td>
<td>10.1</td>
</tr>
<tr>
<td>France</td>
<td>10.4</td>
<td>6.8</td>
<td>2.7</td>
<td>19.9</td>
<td>12.1</td>
<td>4.5</td>
</tr>
<tr>
<td>Norway</td>
<td>9.6</td>
<td>5.7</td>
<td>1.6</td>
<td>26.5</td>
<td>19.1</td>
<td>10.6</td>
</tr>
<tr>
<td>Sweden</td>
<td>17.6</td>
<td>14.2</td>
<td>5.0</td>
<td>40.6</td>
<td>29.3</td>
<td>19.0</td>
</tr>
<tr>
<td>United States</td>
<td>6.2</td>
<td>1.7</td>
<td>0.4</td>
<td>14.0</td>
<td>3.4</td>
<td>0.7</td>
</tr>
</tbody>
</table>

—Not available.
28. Several policy implications may be drawn from these observations. First, since there is a strong negative correlation between welfare dependency and education attainment level, an increase in expenditures on education (including drop-out and teen-age pregnancy prevention programs) may be an effective way to reduce welfare dependence. For example, the data show that women with less than upper secondary education have the highest dependency on social-transfers income. Governments could introduce programs targeted to this group to encourage girls and women to remain in school and reduce their dependency on social-transfers income. Although in the short run this may require extra government investment in education, in the long run, it may dramatically reduce expenditures on welfare assistance. Second, the higher degree of dependence among women on social-transfers income may reflect the social division of labor in which women are more likely than men to leave or stay out of the labor force to assume responsibility for family and child care. For example, in Denmark, Norway, and Sweden, women usually take a long paid or unpaid maternity leave, and government subsidies to children are also recorded under the mothers’ name. All of these may increase social transfers income to women. In this case, governments may need only assure that their social policies do not discriminate against men, or otherwise discourage men from also assuming child-care responsibilities. Finally, the higher degree of dependence among women on social-transfers income may also reflect the economic inequality that women face today in the labor market. Women may not be paid equally for their employment (OECD, 1998). In this case, in order to reduce women’s dependence on social-transfers income, improvements are needed in the legal and economic systems to guarantee that women have an equal opportunity in employment and pay.

The better educated tend to contribute more to society through income taxes

29. Income taxes paid by individuals are revenues to the government. The income tax systems in all the OECD countries are progressive systems in the sense that people with more income pay income taxes at a higher rate. These systems thus function not just to provide funds for common social needs, but also as an important policy tool to redistribute wealth among social groups. As a consequence, income taxes could reduce individuals’ incentives to invest in education, because the better educated earn more and therefore pay more income taxes. However, from the government’s point of view, the larger the income-tax revenue, the more resources available to the government for spending on the common goods of society. In this sense, income-tax revenue provides an important incentive for the government and society to invest more in education.

30. Figure 4-1 shows a widespread positive relationship between individuals’ education attainment level and their income tax contribution. In general, those with more education pay more income taxes, not only because they have higher average annual earnings, but also because they pay income taxes at higher tax rates. Figure 4-1 also suggests that education has a much larger impact on women than on men in terms of increasing their contribution to government income-tax revenues, although this is largely due to the greater increase in earnings for better educated women relative to men.
Figure 4-1. Income taxes of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean income taxes of individuals with a secondary education, by gender and country (Income taxes for individuals with an upper secondary education=100)

Men

![Income Taxes Bar Chart for Men]

Women

![Income Taxes Bar Chart for Women]

Note: Data for people with less than upper secondary education for Australia were not available.

31. Table 4-1 shows that, compared to their counterparts with upper secondary education, the average income taxes paid by men with less than upper secondary education range from 62 percent in the United States to 85 percent in both Belgium and Sweden. The average income taxes paid by men with university education range from 156 percent in Belgium to 216 percent in France. For women, this gap is even larger. For example, if the income taxes paid by women with upper secondary education are set to equal 100 percent, the average income taxes paid by women with less than upper secondary education range from 43 percent in Belgium to 83 percent in Denmark. The average income taxes paid by women with university education range from 147 percent in Denmark to 289 percent in France.
Table 4-1. Income taxes of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean income taxes of individuals with a secondary education, by gender and country

(Income taxes for individuals with an upper secondary education=100)

<table>
<thead>
<tr>
<th>Non-students ages 35-44</th>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>—</td>
<td>100</td>
<td>174</td>
</tr>
<tr>
<td>Belgium</td>
<td>85</td>
<td>100</td>
<td>156</td>
</tr>
<tr>
<td>Canada</td>
<td>80</td>
<td>100</td>
<td>192</td>
</tr>
<tr>
<td>Denmark</td>
<td>81</td>
<td>100</td>
<td>164</td>
</tr>
<tr>
<td>France</td>
<td>75</td>
<td>100</td>
<td>216</td>
</tr>
<tr>
<td>Norway</td>
<td>81</td>
<td>100</td>
<td>178</td>
</tr>
<tr>
<td>Sweden</td>
<td>85</td>
<td>100</td>
<td>189</td>
</tr>
<tr>
<td>United States</td>
<td>62</td>
<td>100</td>
<td>189</td>
</tr>
</tbody>
</table>

Women

<table>
<thead>
<tr>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>100</td>
<td>199</td>
</tr>
<tr>
<td>58</td>
<td>100</td>
<td>232</td>
</tr>
<tr>
<td>83</td>
<td>100</td>
<td>147</td>
</tr>
<tr>
<td>65</td>
<td>100</td>
<td>289</td>
</tr>
<tr>
<td>77</td>
<td>100</td>
<td>174</td>
</tr>
<tr>
<td>83</td>
<td>100</td>
<td>157</td>
</tr>
<tr>
<td>58</td>
<td>100</td>
<td>231</td>
</tr>
</tbody>
</table>

—Not available.

32. Investment in education in the short run may increase government expenditures. However, in the long run, these findings suggest that the government benefits by receiving more income-tax revenue from its better-educated citizenry.

Progressive income tax systems and social transfer programs reduce income gaps

33. As mentioned before, progressive income tax systems and social transfer programs are important mechanisms for societies to redistribute social wealth and to reduce social inequality. Figure 5-1 shows the impact of income redistribution on income equality. Specifically, Figure 5-1 compares the income gaps between various education groups before income redistribution with the income gaps after redistribution. Among all countries with data available, for both men and women, income redistribution reduces the income gaps between the better educated and the less educated. The data suggest that these income redistribution policies often have a larger impact on reducing the income gaps for women than for men.

Figures 5-1. Mean annual income of individuals 35 to 44 years old with various levels of educational attainment as a percentage of mean annual income streams of individuals with a secondary education, by gender and country

(Income for individuals with an upper secondary education=100)

Men (Less than an upper secondary education)
After tax and social transfer income
Before tax and social transfer income

Men (University education)
Figures 5-1. Mean annual income of persons 35 to 44 years old with various levels of educational attainment as a percentage of mean annual income streams of person with a secondary education, by gender and country (Income for people with an upper secondary education=100)—Continued

Women (Less than an upper secondary education)

Women (University education)

Note: Data for people with less than upper secondary education for Australia were not available.

34. Table 5-1 presents information about the magnitude of income gaps among various social groups before and after income redistribution. For example, before redistribution, in the United States, the income of men with less than upper secondary education was 67 percent of that of their counterparts with an upper secondary education. However, after redistribution, this percentage increased to 73 percent. For another example, in Sweden, before income redistribution, the income of women with university education was 153 percent of that made by their counterparts with an upper secondary education. After redistribution, this percentage declined to 126 percent. In both cases, income redistribution reduces the income gaps between the better educated and the less educated, by increasing the after-tax and after-transfer income of the less educated, and reducing the after-tax and after-transfer income of the better educated. In different countries, the degree of the income redistribution differ, resulting in differences in the extent of the redistribution.
Table 5-1. Mean annual income streams for individuals 35 to 44 years old with various levels of educational attainment as a percentage of annual income streams of individuals with a secondary education according to before or after income tax and social transfer income, by gender and country (Income for individuals with an upper secondary education=100)

<table>
<thead>
<tr>
<th>Non-students ages 35-44</th>
<th>Less than upper secondary</th>
<th>Upper secondary</th>
<th>University</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>—</td>
<td>100</td>
<td>166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>88</td>
<td>100</td>
<td>144</td>
<td>71</td>
<td>100</td>
</tr>
<tr>
<td>Canada</td>
<td>83</td>
<td>100</td>
<td>167</td>
<td>64</td>
<td>100</td>
</tr>
<tr>
<td>Denmark</td>
<td>74</td>
<td>100</td>
<td>148</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>France</td>
<td>79</td>
<td>100</td>
<td>197</td>
<td>64</td>
<td>100</td>
</tr>
<tr>
<td>Norway</td>
<td>81</td>
<td>100</td>
<td>153</td>
<td>79</td>
<td>100</td>
</tr>
<tr>
<td>Sweden</td>
<td>86</td>
<td>100</td>
<td>175</td>
<td>76</td>
<td>100</td>
</tr>
<tr>
<td>United States</td>
<td>67</td>
<td>100</td>
<td>182</td>
<td>64</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Before income tax and social transfer income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
</tr>
<tr>
<td>Belgium</td>
</tr>
<tr>
<td>Canada</td>
</tr>
<tr>
<td>Denmark</td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td>Norway</td>
</tr>
<tr>
<td>Sweden</td>
</tr>
<tr>
<td>United States</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>After income tax and social transfer income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
</tr>
<tr>
<td>Belgium</td>
</tr>
<tr>
<td>Canada</td>
</tr>
<tr>
<td>Denmark</td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td>Norway</td>
</tr>
<tr>
<td>Sweden</td>
</tr>
<tr>
<td>United States</td>
</tr>
</tbody>
</table>

—Not available.

35. To the extent that income redistribution reduces the income gaps between the better educated and the less educated, progressive income tax systems as well as social transfer programs may reduce individuals’ incentive to invest in education, because this reduces the income premiums enjoyed by better educated people. However, these data show that this income redistribution rarely eliminates income premiums, so the incentive for individuals to invest in education may be smaller than it could potentially be, but should still exist. In addition, income redistribution increases social and economic equity among various social groups. It may reduce poverty related social problems and create a more cohesive society. The disadvantaged social groups that benefit from the income redistribution may have a better chance to improve their education and earnings as a result. All of this may benefit the society as a whole in the long run.

Discussion and direction for further study

36. In this section, we first provide a brief discussion of the findings, then lay out some options for further exploration of this topic.
**Discussion**

*Education benefits individuals, the government, and the society*

37. The findings from this study suggest that there is a strong positive relationship between people’s education attainment level and their ability to earn more income. As a consequence, not only individuals, but also the government and society benefit from an investment in education. If higher annual earnings and higher property income reflect mainly the economic benefits for better educated individuals, the decline in dependency on social-transfers income and the increase in government income-tax revenue reflect primarily the economic benefits of education for government and society. The findings of this study provide strong support for both individuals and government to invest in education.

*Inability to isolate confounding factors*

38. There are some limitations to this study. One important limitation is that this study cannot remove the effects of confounding factors. Better educated people are more likely than those with less education to come from advantaged socioeconomic backgrounds, and are more likely to have higher congenital abilities as measured in IQ (Ashenfelter & Kreuger, 1994). Both factors are believed to have a positive impact on a person’s performance in school and in the labor market. The current study implicitly assumes that the higher earnings of the better educated are monetary benefits from education rather than from pre-existing genetic or social conditions. But past research has found that part of the higher earnings of the more educated results from the more advantaged social backgrounds and higher IQ of these individuals (Ashenfelter & Kreuger, 1994). Thus, this study overstates the economic benefits of education. Unfortunately, in the absence of data produced from an experiment in which individuals are randomly assigned to pursue different levels of education, there are no simple solutions to this problem. This problem plagues the interpretation of results of all studies that are based on non-experimental data and that cannot appropriately control for the confounding factors.

39. Being aware of the potential problem of overstating the economic benefits of education, one should exercise caution when interpreting the findings from this analysis. As outlined when discussing the possible sources of higher property income for the better educated, both socioeconomic background of the better educated as well as their higher education attainment level may contribute to the observed findings.

*Differences in economic returns reflect differences in socioeconomic systems*

40. Countries have different social and economic systems that are largely shaped by their distinct cultural and political traditions. As a consequence, it is quite natural that countries may have different working-compensation systems as well as income distribution and redistribution systems. For example, both Figure and Table 1-1 show that the earning gaps between persons with different educational levels are much smaller in Belgium and Denmark than in France and the United States. This indicates that the labor market in Belgium and Denmark rewards education less with earnings than does the labor market in France and the United States. On the other hand, Figure 5-1 and Tables 5-1 suggest that the income redistribution through a progressive income tax system and social transfer programs results in a greater reduction of income gaps in Sweden than in the United States. This means that the degree to which the government intervenes in income redistribution is stronger in Sweden than in the United States. The cross-country differences in economic returns to education reflect differences in the social and economic systems in these countries in addition to differences in their education systems.

41. In view of the fact that the economic benefits of education are determined by various social and economic factors, the cross-country differences in economic returns to education cannot be simply attributed to the efficacy or the effectiveness of the education systems in these countries. Thus, in Belgium and Denmark, where better educated people get smaller earning premiums than their counterparts
in France and the United States, one cannot infer that universities are less effective or secondary schools are more effective in Belgium and Denmark than in France and the United States. The smaller earning premiums, in this case, are probably a consequence of broad socioeconomic arrangements. This case provides a good example of the social determination of the economic benefits of education.

**Broadening the scope of the study**

42. In the current study, comparisons mainly focus on the economic benefits across various education attainment levels. In the future, this study could be expanded to include comparisons across various social groups. Examples include comparisons of the economic benefits between men and women, between advantaged and disadvantaged social groups, between full-time and part-time students, between students of different subject areas, or between traditional school-age students and adult-students. One can also explore the interaction among education attainment levels, subject areas of study, and technological progress, and its impact on the economic benefits of education. More specifically, one can examine whether graduates from some subject areas benefit more than graduates from other subject areas, and whether the better educated benefit more from technological progress than do their less educated counterparts. All of these studies can provide important insights regarding the economic consequences of education to both individuals as well as to society as a whole.

**Improving the accuracy and comparability of the data**

43. One critical step to improve this study is to improve the quality of the data provided by countries. The data used in this study were collected in 1997, and, in most cases, the data reflect information as of 1994. This database should be updated very shortly. However, before we embark on a new round of data collection, Network B needs to carefully examine the conceptual framework for the study. Based on this review, the data collection instrument needs to be refined to better serve the information requirements of the study. Depending on the specific research agenda adopted by the Network, appropriate information items need to be collected and cross-country comparability of the data items needs to be ensured. Finally, the new round of data collection should conform to the new international standard classification of education (ISCED 97).

**Developing new indicators based on this study**

44. The findings of this study provided fresh insights into the impact of education on people’s property income, income from social transfers, and income taxes. All of these findings have significant policy implications. Network B may want to consider forming a subgroup to continue the developmental work to define, develop and present new indicators based on these findings.

**REFERENCES**


SECTION III

CONTRIBUTIONS FROM NETWORK C
INTRODUCTION
THE LEARNING ENVIRONMENT AND ITS RELATIONSHIP TO OUTCOMES

Jaap Scheerens, University of Twente, The Netherlands

The learning environment

1. Virtually all aspects and elements of national educational systems can be seen as part of “the learning environment” and be expected to have an impact on educational outcomes. Even characteristics that are deeply rooted in the national culture as well as basic values with respect to learning and education have been used to explain differences in educational achievement between nations. The same can be said of important structural differences in the organization of whole sectors of national education systems, for example, whether lower secondary schools are comprehensive or categorical, and whether or not vocational education at upper secondary level is organised in a “dual system”, where part of the training occurs in the workplace.

2. Environmental factors can be distinguished according to the degree to which they are malleable or appear to be given, that is to say, determined by history or societal and demographic aspects that are beyond the control of educational or even government policy.

3. A second characteristic of environmental factors is the level of aggregation at which the factors are defined. Educational systems can be depicted as a set of nested boxes. National policy measures, and structural and cultural aspects of the national system are the outermost box, moving inward towards state, provincial or regional characteristics, local (i.e. municipal) conditions, the level of schools and educational organizations, the level of the classroom and, finally, the level of the individual student.

4. These two dimensions, the degree of malleability and the level of aggregation, have important implications for the way the relationship between environmental conditions and educational outcomes is seen. The underlying assumption is causality, that is, certain states are expected to have an impact on outcomes. When factors are defined at high levels of aggregation, the causal factors may tend to be uncertain or difficult to ascertain precisely. An example of a highly aggregated outcome measure is the percentage of students graduating without delay. But even when factors are defined at lower levels of aggregation and the sample design allows for the estimation of relationships at multiple levels, causality may still be hard to determine. In dealing with indicators on the educational environment, the issue of causality is nevertheless relevant for the selection of environmental, or “process” indicators as well as for their interpretation.

5. Despite the fundamental importance of the causality principle, in practice there are limits to its application, which means that it tends to be used in a heuristic sense in the selection of indicators and in their interpretation. Internationally work has focused on the identification of factors that, by general consensus, are considered important in understanding the nature of education. Indicator development work has concentrated on three main areas:

- indicators on the locus of decision-making in educational systems;

---

The indicator work described in this introduction and in the following chapters was carried out in the context of the OECD INES (Indicators of Educational Systems) Project. The network of national experts responsible for developing “process” indicators is commonly referred to as “Network C”.

139
– indicators on teachers and the conditions of teaching at national level;
– indicators on conditions of schooling and on instruction at school and classroom level.

6. In the following four chapters aspects of these three areas of work will be discussed. Since the issue of causality inevitably arises in the discussion, some further background and perspective to this issue will be given in this introduction.

**The research literature on school effectiveness and educational productivity**

7. The question of which aspects and elements of learning environments “work” with respect to optimizing outcomes has been dealt with in various strands of educational research, varying from studies on “the education production function” (economics), to research on effective teaching (educational psychologists) and on unusually effective schools (educational sociologists). Reviews of the results of these three strands of educational effectiveness research are given in Hanushek, 1997; Wang, Haertel & Walberg, 1993 and Sammons et al, 1995, respectively. It is beyond the scope of this introduction to present an overview of the integration and the “state of the art” in these fields (cf. Scheerens, 2000, in press). It suffices to say that the results of this educational research have been used in identifying relevant process indicators. This has been the case when material from available international (IEA) surveys was used in order to develop indicators on classroom processes and in the preparation of the 1995/1996 international survey of primary schools. For future applications the knowledge base on educational effectiveness and recent developments in this field are considered to be of prime relevance. At the same time a review of the research findings points to important limitations. Some major points that need to be kept in mind are the following:

1. Research reviews and meta-analyses indicate that factors that are close to the actual teaching and learning situation matter most with respect to educational outcomes, compared to more remote factors such as managerial conditions and conditions above the school level.

2. Obvious factors of classroom management and curriculum, such as time on task and opportunity to learn, do indeed appear to be central in optimizing outcomes.

3. With respect to instructional strategies the evidence in favour of structured “teacher-centered” approaches appears to be most convincing, despite the more modern emphasis on constructivist approaches and “independent learning”.

4. International comparative studies indicate that factors that enhance effectiveness are not easily generalisable across countries, in other words, what appears to “work” in one country does not necessarily do so in the next.

5. In research on education production functions, recent studies challenge the dominant view that “money does not matter” that held sway in this line of research until a few years ago. Experimental studies on class size and correlation studies on the effect of teacher qualifications have shown sizeable effects attributable to these factors (e.g. smaller classes at the elementary level and subject matter mastery and verbal skills of teachers).

6. The knowledge base on educational productivity and school effectiveness exists primarily for the elementary and lower secondary level; efforts to replicate the research findings at higher ISCED levels have been only partly successful (see the chapter by Brandsma in this volume).

---

67 Constructivism is a form of pedagogy that emphasizes independent learning, i.e. sees the learning task as a construction by the learner.
Implications for future international work

8. Given the relevance of the above mentioned strands of educational research, the relevant literature will remain a source of inspiration. The recent findings and the ongoing debate in the field of education production functions is a case in point. In this debate the distinction between class size and pupil-teacher ratio is an important one, as is the need to obtain school-level information on these variables in addition to aggregate information at the national level. The same conclusion applies to teacher qualifications, where the research results suggest that the development of meaningful indicators will be a challenge (see Hendriks’ chapter).

9. At the same time it should be realized that there exist evaluation perspectives other than effectiveness, such as the equitable distribution of resources, the adaptiveness and responsiveness of schooling, and more intrinsic aspects of good organizational functioning, such as clarity of regulations and procedures and teacher satisfaction. In interpreting the information on the locus of decision-making these additional perspectives could be important. At the upper secondary and tertiary levels, adaptability in facilitating transition to the labour market is an additional important criterion for assessing the functioning of educational organizations.

10. In planning future work on process indicators of school functioning, it should also be realized that indicator projects are not usually designed as research studies, so that verifying cause and effect relationships empirically is not possible in most cases. The PISA project may be an important exception in this respect. However, the descriptive comparison of process indicators across countries with a view to gaining some understanding of the conditions of schooling and learning environments, coupled with rather modest aspirations with regard to the attribution of causality, is likely to remain the usual practice. In the chapter on the use of surveys to measure indicators on school organizational and instructional characteristics, examples of such descriptive interpretations are given.

11. Even if causality cannot be addressed, opening the black box of schooling by means of indicators on the educational environment and on processes is a relevant and fascinating enterprise, as the ensuing chapters will demonstrate.
REFERENCES


MEASURING PROCESS INDICATORS ON SCHOOL FUNCTIONING BY MEANS OF SURVEYS

Jaap Scheerens, University of Twente, The Netherlands
Carmo Climaco, Inspeção Geral da Educação, Portugal
Bodhild Baasland, Ministry of Education, Research and Church Affairs, Norway

Introduction

1. Educational data systems tend, perhaps naturally, to be geared towards information on educational inputs. This reflects in part the requirements of establishing education budgets and of distributing available resources among schools. With the increasing focus on accountability has come greater attention to outputs and outcomes, in other words, to how well education systems are carrying out their assigned mandate. A question of interest in this regard is by what processes available education inputs translate into a given set of outcomes at the school level. In the absence of an explicit linking of outcomes to inputs, one possible approach is to focus on what are generally considered “good practices” with regard to the organisation of instruction and the functioning of schools.

2. Ideally school process indicators should be able to predict relatively high performance in terms of student achievement. This is also the basic rationale for selecting factors that have received some empirical support in studies of educational productivity or school effectiveness research. The fact that such factors are mostly malleable conditions ensures their policy relevance as a basis for innovation and school improvement. Apart from the fact the implicit causal assumptions are not quite met in the state of the art of the field of educational effectiveness, stand-alone surveys do not allow for an empirical test by relating process indicators to output indicators. Fortunately the PISA-project does offer these possibilities (cf. Bosker, 1999). But in other cases, like the earlier primary school survey and the currently planned upper secondary school survey process indicators have to be interpreted “on their own”. Given the existing uncertainty of the predictive validity of even a carefully selected set of process indicators they cannot be simply used as “substitute output indicators”.

3. An alternative way to interpret and value process indicators in their own right is to consider them as internationally recognized characteristics of educational good practice. Their development in international networks warrants this interpretation. Comparative use of the information provided in such indicators provides international benchmarks of educational good practice. Since absolute standards on these indicators are hard to come by, average “scores” across countries are the most likely choice of an international benchmark. More in-depth information of a qualitative descriptive nature is needed to provide context to this information.

4. This chapter assesses the possibilities of providing process indicators on school functioning through surveys. The first part of the chapter reviews variables collected for four recent and current OECD surveys of schools, as a basis for determining what are the appropriate variables to include in a survey for the development of indicators on the functioning of schools. The effectiveness of the school-survey approach is then considered. The strong and weak points of surveys that depend on school management information and responses from principals are reviewed. Alternatives, like the inclusion of data from teachers and students, are presented. The final section deals with the interpretation of data from schools
surveys and their place in international benchmarking, by means of material from the 1995/1996 Network C survey of primary schools in twelve European countries.

**Selection of variables in school surveys**

5. Table 1 provides a schematic overview of the coverage of variables that have a place in the research literature on school effectiveness and educational productivity (see also the introduction to this section of the Compendium) in four OECD surveys. The pupil and school questionnaires of the PISA project and the questionnaires from the Network C primary schools survey (1995) and the Network C upper secondary school survey (draft version 1999) reflect the choice of variables in these studies.

6. The number of questions indicated is a rough estimate, since the questionnaires differ in the complexity of the questions: in the case of the upper secondary survey, for example, the questions that were counted usually consist of several sub-questions. Since the PISA pupil- and school-questionnaire are part of the same study, the choice of variables in both instruments could be seen as complimenting each other. When the union of both sets is taken it appears that they cover the theoretically relevant variables to a fairly large extent.

7. The OECD-Network C primary school survey has been most consistent in including school level variables that have been designated as relevant in school effectiveness research. When considering the selection of variables in the PISA pupil questionnaire the possibilities for variables at classroom level that reflect instructional processes have only been exploited in a relatively sparse way. For example no questions on evaluation practices and opportunity to learn have been included.

8. The overall picture presented in the table illustrates the lack of consensus and establishment of what one could call “instrumental educational theory”. Variables that are most consistently represented are class size, differentiation, evaluation practices, orderly climate and parental involvement.

9. Striking variables that are left out in all of these studies are “opportunity to learn”, teacher qualifications and school level per pupil expenditure.
**Table 1: Coverage of “process conditions” in four OECD surveys in terms of number of questions**

<table>
<thead>
<tr>
<th></th>
<th>PISA student</th>
<th>PISA school</th>
<th>C/Primary</th>
<th>C/Upper secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equipment, resources, computers</td>
<td>1</td>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Class size, pupil teacher ratio</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Teacher qualifications</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>School budget, ppe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional leadership</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Staff consensus &amp; cooperation</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Achievement orientation, high expectations</td>
<td>2</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td></td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Evaluation frequency &amp; use, feedback</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Orderly climate/ stability/absenteeism</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Structured teaching</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opportunity to learn</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross/net teaching time</td>
<td>3</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Parental involvement</td>
<td></td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Private/public school</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>% of school budget acquired by the school itself</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average student ses</td>
<td></td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional development of teachers INSET</td>
<td>1</td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Transition to labour market and higher education</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

10. Apart from the uncertainty concerning the establishment of knowledge on “what works” in education, there are other reasons for the somewhat shifting priorities in the choice of variables in these surveys. First of all, other perspectives, apart from “internal effectiveness”, may play a role. For example the heavy emphasis on student home background characteristics in the PISA pupil questionnaire seems to be inspired by interest in equity rather than effectiveness questions. The attention to transition issues in the upper secondary school survey is inspired by a broader perspective on organizational effectiveness, including adaptation to environmental demands and longer-term outcomes of schooling.

11. What can be learned from these experiences and which recommendations should be given with respect to the planning of future surveys? There are perhaps four major recommendations:

a) use a broadened concept of organizational (i.e. school) effectiveness;

b) keep track of recent research findings in the domain of educational effectiveness and educational productivity

c) design questionnaires on the basis of operational definitions of core indicators

d) maintain Network procedures and expert review as the procedural approach to design questionnaires
e) strive to maintain common definitions (meaningful across countries) for the target population of the survey; use established survey research procedures and questionnaire design principles to ensure the quality of surveys across countries.

12. Concerning point a), a broader concept of organizational effectiveness can be used as the basic model to generate relevant process (and output) indicators (Scheerens, 1995). Table 2, for example, uses the Quinn & Rohrbaugh (1983) framework to generate a more extensive set of indicators.

**Table 2: Additional factors for process indicators generated from the Quinn and Rohrbaugh framework**

<table>
<thead>
<tr>
<th>Human relations model</th>
<th>Open system model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quality of work life indicators</strong></td>
<td>- entrepreneurship</td>
</tr>
<tr>
<td>- respect</td>
<td>- collegiality</td>
</tr>
<tr>
<td>- participation in decision-making</td>
<td>- capacity for self-evaluation and learning</td>
</tr>
<tr>
<td>- professional interaction</td>
<td>- overt school marketing activities</td>
</tr>
<tr>
<td>- performance feedback</td>
<td>- parental involvement</td>
</tr>
<tr>
<td>- opportunity to use skills</td>
<td>- boundary-spanning positions</td>
</tr>
<tr>
<td>- resources</td>
<td>- external change agents</td>
</tr>
<tr>
<td>- congruence personal/organizational goals</td>
<td>- student enrolment figures</td>
</tr>
<tr>
<td></td>
<td>- resources (buildings, equipment)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Internal process model</th>
<th>Rational goal model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>(school effectiveness research)</em></td>
</tr>
<tr>
<td>- planning documents</td>
<td>- educational leadership</td>
</tr>
<tr>
<td>- disciplinary rules</td>
<td>- success-oriented ethos</td>
</tr>
<tr>
<td>- management information systems</td>
<td>- monitoring of student's progress</td>
</tr>
<tr>
<td>- formalization of positions</td>
<td>- time on task</td>
</tr>
<tr>
<td>- continuity in staffing and leadership</td>
<td>- content-covered (opportunity to learn)</td>
</tr>
<tr>
<td>- integrated curricula</td>
<td><em>(broader set of educational goals)</em></td>
</tr>
<tr>
<td>- attendance rates</td>
<td>- non-gradedness</td>
</tr>
<tr>
<td>- lessons &quot;not given&quot;</td>
<td>- team teaching</td>
</tr>
<tr>
<td></td>
<td>- individualization, differentiation</td>
</tr>
<tr>
<td></td>
<td>- continuous learning route</td>
</tr>
<tr>
<td></td>
<td>- time spent on social, emotional, creative and moral development</td>
</tr>
<tr>
<td></td>
<td>- &quot;learning to learn&quot; activities</td>
</tr>
<tr>
<td></td>
<td>- diagnostic testing</td>
</tr>
</tbody>
</table>
13. With respect to point b), examples of recent research findings in the field of educational effectiveness research concern *class size* (Galton, 1998, Grissmer, 1999), *teacher qualifications*, and (<biblio>) *and classroom composition*. As stated before, teacher qualifications have not received much emphasis in earlier school surveys by the OECD. The same is true for classroom composition. Research results concerning class size suggest that the effects differ for different age groups and for classes that have different SES-composition.

14. On c), experience suggests that surveys that are meant to generate indicators are designed as descriptive surveys and, despite good intentions, do not start from operational definitions of *indicators*. In the task forces that design the surveys, discussions are about substantive areas and individual questionnaire items, and less about indicators. The result is certain inefficiency in the sense that there are more questions and variables than are used for the actual indicators. The fact that individual countries may be interested in all the descriptive data from the surveys may be seen as a compensation for this inefficiency.

15. Regarding d), given the uncertainties with respect to the knowledge-base on educational effectiveness that were referred to above, the selection of process indicators that are considered relevant with respect to educational outcomes depends on consensus in relevant groups of representatives and experts. However, the criterion of policy relevance of the process indicators concerned is likely to differ between countries. Group decision-making should thus try to focus on indicators that are relevant in most national contexts.

16. Finally, on e), in the case of the preparation of the Network C upper secondary survey, a specific preparatory study was deemed necessary in order to delineate a comparable core of programmes (see Brandsma, this volume).

17. As was already stated in the introduction, the higher one ascends on the ISCED scale, the less compelling are the ideas and empirical evidence concerning the impact of characteristics of educational organizations and learning environments on outcomes. This suggests that what can be learned from the earlier experiences discussed above for surveys at the tertiary level is procedural rather than substantive. International networks like the current INES-Networks and expert consultations are considered to be essential in the selection of indicators and the design of surveys.

**Methodological questions regarding school surveys**

18. School surveys, administered to school heads or directors, have several methodological drawbacks. The possibility that responses may be affected by social desirability cannot be ruled out. A general strategy should be to ask questions that have a verifiable factual basis rather than to solicit opinions. On the other hand, asking factual questions about, for example, staff in full-time equivalents, number of pupils enrolled, absenteeism and different indicators on teaching and *instruction time* presupposes that schools keep administrative records containing this information. Despite the availability of elaborate computerized school management information systems, it is often found that the quality of this information leaves much to be desired. This implies that checks on the reliability of this type of information ought to be carried out, in order to have an estimate of the seriousness of this problem. Reliability studies, such as re-interview techniques (cf. Forsman & Schreiner, 1991) should likewise be considered.

19. A third problematic area has to do with the substance of the questions. Roughly speaking, one should distinguish between characteristics of the school organization, overall school policy and characteristics of management procedures, on the one hand, and instructional strategies and practices, on the other. Principals are well equipped to answer questions of the first kind, but the second domain, namely instructional strategies and practices, is much more problematical. In large schools the heads may simply have insufficient information on instructional strategies and classroom practice. In addition, given
the autonomy of teachers, there is likely to be variation between teachers and practices at different grade-levels. Under these circumstances, it is difficult for heads to arrive at reliable estimates of “average practice” in the school. The simplest remedy for this problem would be to leave out questions about classroom practice as much as possible. On the other hand, the direct learning environment has the largest impact on student achievement, compared to more “distal” factors (Wang, Haertel & Walberg, 1993) and leaving out this category completely may be a considerable loss in educational relevance. A solution that has been proposed is to prescribe procedures through which heads are guided to systematically consult samples of teachers on certain questions. Such a procedure makes the data-collection procedure more complex and time-consuming, however, and is therefore likely to reduce the response-rate.

20. Such problems, with only partially feasible remedies, mean that school surveys administered to head teachers may have many pitfalls and trade-offs.

21. Practices in other international surveys and in (national) school self-evaluation projects may offer some interesting possibilities to overcome some of these problems. Procedures whereby teachers rate the behaviour of head teachers and students that of teachers do away with some of the threats of social desirability inherent in self-descriptions. Rather than focussing on teacher descriptions of school management aspects, a case can be made for using pupils’ perceptions of classroom processes, especially because of the greater relevance of instructional as compared to school organizational factors. In several projects positive experiences have been obtained with this procedure (Scheerens, Hendriks & Bosker, 2000; Kuyper & Swint, 1996; Van Os, 1998; Hill, Rowe and Jones 1955).

22. An example of this approach is provided in Table 3. Evidence on the reliability and validity of this short rating-scale, on which primary school pupils evaluated their teachers, is provided by Kuyper & Swint (1996).

Table 3: Students’ opinion about the school, subjects and teachers (lower secondary level)
Source: GION, Groningen

How do you like or dislike this school?

1) How do you like or dislike these subjects?
2) How easy or difficult do you find this school?
3) How easy or difficult do you find each of these subjects?
4) How do like or dislike the teacher of each of the following subjects?
5) How well do you think the teacher each of the following subjects keeps order?
6) How well do you think the teacher of each of the following subjects explains things?
7) Does the teacher of each of the following lessons make you feel at ease during the lessons?
8) How many times are students asked to answer questions in front of the class in each of these subjects?
9) How interesting do you think the teaching of the teachers for each of these subjects is?
10) Do you think that the teachers of each of these subjects make it clear why the subject is important?
11) Do you think that the teacher of each of these subjects is fair to the students?
12) Is it possible to talk about personal matters to the teacher of each of these subjects?

23. If the intention is to get some information on the quality of the learning environment on the basis of school surveys, then a limited data collection among a sample of students per school should be seriously considered. A similar case can be made with respect to the inclusion of teacher-level information.

24. This review suggests that, for future application, a broader notion of a school survey, one that actually includes the idea of teacher surveys and pupil surveys, should be considered. In this way a deeper understanding of classroom management practices, attitudes and perceptions of teachers and students, professional development and instructional strategies may become feasible.
Process indicators as international benchmarks of educational “good practice”

25. The use of process indicators as benchmarks of educational “good practice” is illustrated by means of a few sections from the unpublished report on the Network C primary school survey. Core variables from the survey were used to construct “country profiles”. Each national score on a particular variable/indicator was presented together with the country average. The illustrative information contains the country-profiles of the Netherlands, Norway and Portugal.

26. A country profile is made up of 7 indicators for each country: stability, educational leadership, cooperation among staff, evaluation, differentiation, achievement orientation and parental involvement. In fact a selection of the results on each of the indicators is provided, whereby the results on a key variable on which the indicator was based is shown.

27. The following variables were selected to provide country profiles:
   1. The percentage of pupils in schools where over 75% of the teachers was also employed five years previously (row four of table S1);
   2. The percentage of non teaching headteachers’ time spent on educational leadership (in terms of pupils attending primary schools lead by these headteachers) (row 12 of table L1);
   3. The percentage of pupils attending primary schools in which entire staff meetings take place every week or more (row 11 of table C1);
   4. The percentage of pupils attending primary schools where information from pupils’ records is used for evaluating the functioning of the school (row 5 of table E1);
   5. The percentage of pupils in primary schools which practice ability grouping within classes in some or all classes (row 6 of table D1);
   6. Percentage of pupils in schools where achievement standards are set at the pupil level (row 4 of table A1);
   7. Percentage of pupils in primary schools with structures to involve parents in decision making on the curriculum (row 1 of table P2).

28. It should be noted that the country profiles are based on a limited set of variables, with a somewhat “narrow” interpretation of each of the indicators. In the comments that are provided with the charts and tables representing the profiles an explicitly broader scope is chosen, in which countries provide contextual information that relate to the broader definition of indicators used in the survey (Scheerens & Ten Brummelhuis, 1995).
### Country profile Netherlands

#### 95% confidence interval  
average of participating countries

<table>
<thead>
<tr>
<th></th>
<th>value in %</th>
<th>Standard error</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stability in teaching staff</strong></td>
<td>63</td>
<td>2.513</td>
<td>58 68</td>
</tr>
<tr>
<td><strong>Time for educational leadership</strong></td>
<td>25</td>
<td>2.254</td>
<td>21 29</td>
</tr>
<tr>
<td><strong>Entire staff meetings every week</strong></td>
<td>81</td>
<td>2.042</td>
<td>77 85</td>
</tr>
<tr>
<td><strong>School self evaluation</strong></td>
<td>60</td>
<td>2.550</td>
<td>55 65</td>
</tr>
<tr>
<td><strong>Ability grouping within classes</strong></td>
<td>32</td>
<td>2.428</td>
<td>27 37</td>
</tr>
<tr>
<td><strong>Achievement standards for pupils</strong></td>
<td>74</td>
<td>2.283</td>
<td>70 78</td>
</tr>
<tr>
<td><strong>Parent involvement in curricular decision making</strong></td>
<td>63</td>
<td>2.513</td>
<td>58 68</td>
</tr>
</tbody>
</table>

29. In the Netherlands staff stability in primary schools is fairly high, educational leadership is about average and cooperation, measured in terms of meetings of the entire teaching staff occurs quite frequently. The Netherlands is also high on the use of records for school evaluation purposes, the setting of achievement standards (highest of all countries) and parent involvement. With respect to ability grouping within classes the Netherlands is among the lower scoring countries, although this relative appraisal does not mean that ability grouping is a rare phenomenon (average of 32% of the pupils in schools where this practice occurs).
Staff stability

1. Staff stability is relatively high in the Netherlands. Most schools use the principle ‘last in, first out’ when it comes to firing teachers. At the time the study was conducted (1995/96) there was a decline in enrolments, which was considered to make it risky for teachers to move from one school to the other. Even experienced teachers can lose their jobs after changing schools when student numbers at their new school drop. In government policy, mobility rather than stability is regarded as a desirable goal. By bringing more schools under the authority of the same school board, it will become easier for teachers to change schools. Recently enrollments have picked up again and there is a shortage of teachers so that the conditions for stability vs. turnover have changed considerably.

2. The average age of teachers was also considered as a factor related to staff stability. The average age of teachers is rising, and will continue to do so in the next decade. This condition will make the problem of teacher shortage even more compelling in the near future.

3. The stability of head teachers in the Netherlands is higher than anywhere else. Head teachers are appointed by the school board. Unlike the situation in some other countries, being a head teacher is a position that is held until retirement. Moreover, the vast majority of head teachers are male, and men tend to have longer working lives.

School leadership

30. Due to decentralization policies there is an increased emphasis on managerial tasks for head teachers. He or she is forced to develop from ‘best among their equals’ into a manager, who has to coach and to judge the performance of teachers. The percentage of their employment that is spent on teaching is decreasing. Head teachers of small schools spend more time on teaching than do head teachers of large schools.

Evaluation

31. Pupils leaving primary school receive an advice as to which of the four kinds of secondary schools they can attend. According to the law this advice must also be based on an objective test. Each year the majority of the schools use the same test, which is provided by the National Institute for Test Development, CITO. This institute sells other standardized tests to schools, especially for formative evaluation of pupil achievements. A (still growing) number of schools apply these tests. A further development of these tests for school self-evaluation purposes is foreseen for the next years.

Differentiation

32. Multi-graded classes are still widespread in the Netherlands, despite important changes with respect to increase in school size. Multi-graded classes often assign their pupils to age groups rather than ability groups.

Achievement orientation

33. Comparison of pupil achievement with national averages (or with the average of comparable schools) is quite common for a few core competencies, like reading and arithmetic. The instruments for such comparisons are not yet available for all subject areas.

34. A high percentage of respondents say their schools use achievement standards for individual pupils. Probably part of them is referring to the national attainment targets. These apply to all pupils in all primary schools.
Parent involvement

35. In the Netherlands two thirds of primary schools are governed by private school boards, consisting of (former) parents and (other) representatives of the local community. The other one third are governed by the municipal authorities. Private school boards and municipal authorities have a large degree of autonomy to make decisions on curriculum, planning, finance, personnel and organization. This way, parents do have considerable influence in school policy. 

36. Moreover, all Dutch schools have a School Participation Council, where both parents and school personnel are represented. Schools are obliged by law to involve this council in decision-making. Decisions about the curriculum cannot be taken without the agreement of the parents in this council.

37. Recent policy will prescribe schools to inform all parents on the goals, activities and results of the school.

38. Parental involvement is not restricted to decision making, discussion and advice. A number of parents assist in educational tasks like reading in small groups or working with computers in the classroom, and other tasks like the opportunity for children to stay at school during lunch-time, assisting at excursions and helping to maintain the school building.

39. The amount of this work is not reflected in the indicator ‘percentage of parents’ since only a small number is actually involved in doing the work.
Country profile Norway

95% confidence interval  average of participating countries

* As stated in chapter 2 Norway did not provide data on evaluation practices

Norway

<table>
<thead>
<tr>
<th></th>
<th>value in %</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>country</td>
<td>average of countries</td>
</tr>
<tr>
<td>Stability in teaching staff</td>
<td>57</td>
<td>54</td>
</tr>
<tr>
<td>Time for educational leadership</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Entire staff meetings every week</td>
<td>88</td>
<td>30</td>
</tr>
<tr>
<td>School self evaluation</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>Ability grouping within classes</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>Achievement standards for pupils</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Parent involvement in curricular decision making</td>
<td>64</td>
<td>31</td>
</tr>
</tbody>
</table>

40. Compared to other countries the stability of staffing in primary is slightly above average. Norway is about average in educational leadership and highest in cooperation in the sense of the frequency
of entire teaching staff meetings. The items on evaluation were not scored in Norway since such evaluation practices do not fit in the Norwegian educational system (see commentary). Ability grouping within classes and the setting of achievement standards are quite low, as compared to other countries. Parental involvement in curricular decision-making is common in most Norwegian primary schools.

41. The following further comments were made.

**Stability of school staff**

42. In Norway, high turnover of teachers, sometimes also combined with shortage of formally qualified teachers altogether, has been seen as a threat to an equitable provision in particular in the remote rural areas of coastal northern Norway. A number of measures, therefore, has been taken over the years to stimulate teacher recruitment in such areas, including the dispersal of teacher training institutions throughout the country and salary inducements and travel allowances for those willing to serve small, outlying schools. In spite of such costly schemes, the problem has not, as yet, been completely solved. Even though Norway in the present study demonstrates a reasonably high stability level both regarding teachers and head teachers in comparison to the other countries included, there are certainly great geographical differences within the country. A recent study covering all the 300 schools in one of the northernmost counties showed that one third of the small schools (defined as primary 1-6 grade schools having less than about 110 pupils) had a turnover of at least 90% over a three year period. Such findings demonstrate the kind of within country variation that obviously may be present in many countries. From an equity point of view, inequalities in the education provision is damaging even though the proportion of pupils suffering from them may be relatively low.

43. The Norwegian study referred to above, also ventured into the possible effect of high versus low stability rates on the schools' ability to adjust to new organizational demands and recommendations related to such aspects as staff cooperation, school - home and school - community relations, timetable flexibility and curriculum development work. The main findings may be summarized in this way: In all schools, whether small or large, well staffed schools having some annual influx of new teachers are, on average, the more fruitful ones in terms of organizational development. Very high turnover is demonstrably detrimental in this respect across school sizes. The effect of very stable teaching forces (including head teachers), however, is markedly different on small and large schools. Large schools enjoying a very stable staff situation were only moderately less able to implement organizational improvements than large schools with a somewhat lesser degree of stability. The very small schools (2-4 teacher schools) with a very stable staff were the schools, which showed the poorest ability to adjust to new organizational and curricular demands and recommendations. These findings point to the need for support structures and network building particularly concerning the very small schools apparently enjoying a good staff situation.

**School leadership**

44. Since the early 1980s in Norway there has been an active strive for stimulating head teachers to spend more of their non-teaching time on educational leadership. Three nation-wide comprehensive programmes for leadership in-service training have been carried out. Against this background it may be surprising that the estimated time addressing educational leadership is just about the average for those countries included in the study. The findings do, however, correspond closely to the above-mentioned Norwegian investigation carried out during 1989-91 in which the average non-teaching time for head teachers was categorized as educational leadership, administrative work, and staff management. The amount of administrative work, averaged over schools, was reported to take about 55% of head teachers' non-teaching time whereas educational leadership made up for 25%. In this study a visiting research worker administered the questionnaire. In addition to the head teacher, a teacher-union representative among the staff was also involved as a respondent in estimating the percentages. The relative high correspondence as regards the Norwegian responses between these two studies may indicate a reasonably high reliability as regards the head teachers’ estimates in the present (OECD) comparative study.
45. The fact that the relative amount of head teachers' time spent on educational leadership is not higher, and does not seem to be increasing over the last years, may reflect another development in Norwegian schools especially noticeable over the last 10 years. This development concerns a tendency to delegate more power and responsibility from local education authority level to school level, generally without strengthening the administrative resources at school level.

46. According to an agreement, settled in 1977, any head teacher in a Norwegian primary and/or lower secondary school should have a certain amount of teaching depending on the size of the school, but never less than 25% of a full-time teaching post. This compulsory teaching commitment on behalf of a head teacher was ruled out in a new agreement from March 1996 to be implemented by 1st of August 1996. It is somewhat surprising that only 77% of the Norwegian head teachers in the present study claim to be teaching at all. Apparently the new regulations seem to have been anticipated, again an indication on the increasing demands on the head teachers as to leadership and management commitments that have been felt since the mid-80s.

Staff cooperation

47. Cooperation was a key word in the new 'Curriculum Guidelines for Compulsory Schooling in Norway' implemented in 1987. The present findings indicate that the effort to develop schools from being largely an institution comprising a number of individual teachers dealing more or less independently with their respective classes, to an organization of cooperating professionals accepting collective responsibilities, may have paid off. The relatively high level of formal and informal cooperative activities in Norwegian schools also reflects the fact that, since 1994, an agreement on the working conditions for the teachers states that a certain amount of the teachers’ working time (190 clock hours annually) should be available for various kinds of whole staff or sub-group activities, decided upon by the head teacher. Thus there are resources available for this kind of activities. However, as the Norwegian study referred to above demonstrated, the quality of the cooperative work going on in the schools varies a lot from being «...a millstone around your [the teachers'] neck »to being «the engine» for the whole running and development of the school, to quote some of the expressions coming up during interviews with teachers and head teachers. This study also clearly indicates that the schools demonstrating a high amount of qualitatively good joint staff activities, also were the schools in which a strong culture in line with the intentions of the (national) curriculum guidelines were most likely to develop.

Evaluation

48. Since formal evaluation, for instance in terms of marks, is not part of Norwegian primary education, this section was left out of the questionnaire employed in Norway. Nation-wide standardized tests particularly in Norwegian, mathematics and English, which were in common use during the 1960s and 70s, are not part of present day primary education in Norway. This, of course, does not imply that Norwegian primary school teachers are not evaluating the work and progress of their pupils. According to formal regulations there have to be conference hours, at least twice a year, between the school and the individual parent(s), where the pupil's general development as well as her/his progress in the various school subjects are discussed. During the last few years we may observe a greater willingness to reintroduce more formal assessment procedures. As part of a more comprehensive programme for school evaluation, The Ministry of Education has initiated the development of standardized tests for diagnostic purposes in Norwegian (grade levels 1-2 and 5) and in mathematics (grade 3-4). Work on diagnostic tests in English is also in progress.

Differentiation

49. Norway has a long tradition for delaying formal differentiation in classes or groups within classes. This means that the pupils should be restricted in their options for choices at the next educational level, for instance when being transferred form primary to lower secondary schools or from lower to
higher secondary education. In line with this tradition the composition of classes or groupings within classes according to ability levels rarely occurs in Norwegian schools.

50. Considering the high proportion (45%) of primary and/or lower secondary schools which has classes with two or more grade levels, it appears to be somewhat surprising that only 18% of the pupils do belong to schools having such classes. The finding is mainly explained by the fact that those schools having multi-grade classes are often very small ones (cf. introductory comments above), but also reflects a disbelief among Norwegian teachers and school leaders in multi-grade teaching. In Sweden, with a more centralized school structure and on average having considerably larger schools, more than 50% of the pupils are in schools that practice multi-grade teaching. This difference between Norway and Sweden has to be seen as a deliberate policy in Sweden to maintain multi-grade classes in schools having enough pupils to establish single-grade classes.

Achievement orientation

51. According to the criteria used in the present study, Norway, along with Sweden, appears to be only moderately achievement oriented in primary schools. The findings illustrates an ideology in Norwegian primary education of not cultivating the kind of open competition that might produce losers and young people lacking self confidence. On the other hand the principle of adapted education, strongly underlined in the Curriculum Guidelines, implies that every child shall be stimulated to develop according to her/his potentials. The relatively high percentage of Norwegian pupils experiencing schools where public recognition is related to the pupil's individual progress, therefore, may seem consistent with the above-mentioned ideology and priority.

Parent involvement

52. Related both to the comments on evaluation and achievement orientation it is interesting to observe that most Norwegian pupils belong to schools where the parent seem to be kept informed about the pupils' progress. The parents also seem to be informed on and to some extent involved in matters related to aims and objectives, curricular content and the general running of the school.
Country profile Portugal

95% confidence interval average of participating countries

Portugal

<table>
<thead>
<tr>
<th></th>
<th>value in %</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>country</td>
<td>average of countries</td>
</tr>
<tr>
<td>Stability in teaching staff</td>
<td>24</td>
<td>54</td>
</tr>
<tr>
<td>Time for educational leadership</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>Entire staff meetings every week</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>School self evaluation</td>
<td>36</td>
<td>42</td>
</tr>
<tr>
<td>Ability grouping within classes</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>Achievement standards for pupils</td>
<td>54</td>
<td>40</td>
</tr>
<tr>
<td>Parent involvement in curricular decision making</td>
<td>14</td>
<td>31</td>
</tr>
</tbody>
</table>
53. The stability in staffing of Portuguese primary schools is fairly low. The country is average in educational leadership. Entire teaching staff meetings occur very infrequently. School evaluation is average, whereas ability grouping within classes is below average. Achievement orientation, in terms of standard setting is above average, whereas parental involvement in curricular decision-making is relatively low.

**Contextual information**

54. In the specific Portuguese context pupils served at ISCED 1 (from 6 up to 11 years of age) are to attend first cycle schools (primary schools) for the first four grades (grades 1 to 4, for 6/9 year-olds), and second cycle schools and/or comprehensive schools for the following two grades (grades 5 and 6, for 10/11 year-olds).

55. 1st cycle schools are settled all over the territory and are characterised by the small size of each unit and the dispersion and isolation of the school network. The public schools consist of ca. 9300 establishments, 50% of which are very small schools merging grades 1 to 4 in the same class.

56. One single teacher is in charge of the class and is responsible for teaching the basic subjects. Generally classes take pupils from the same age group. However, for practical reasons, a class may include pupils from different age groups and grade levels.

57. Second cycle of compulsory basic education (grades 5 and 6) is delivered at 2nd cycle schools (grades 5 and 6) and comprehensive schools (from grade 5 up to 9).

58. In spite of the intended interdisciplinary approach to the curriculum in both types of schools, pupils are taught by a team of teachers, usually with one teacher per subject.

59. Both these categories of public schools consist of ca. 550 differently organised establishments admitting all pupils having completed the first cycle.

**Stability in school staff**

60. Due to the characteristics of the school network and the predominance of small and isolated localities, mainly in the rural areas, the teacher turnover is a major constraint to the development of education in Portugal.

61. On the whole, the degree of (un)attractiveness of particular socio-geographic areas plays a major part in the fairly high (73%) mobility of teachers: the most qualified, experienced and stable teachers concentrating in urban areas; schools located in semi-urban and rural areas having a substantially less qualified, less experienced and less stable teaching force.

62. Other such factors like demographic changes resulting in a certain degree (25%) of school structural change (fusion, suppression, creation) as well as in the staff retirement, and the teachers’ right to mobility being consecrated by the present Teachers’ Professional Status and Careers (Decree 139-A/90) also account for the low stability of the teaching staff.

63. Moreover, the structural effect, created by the high proportion of (very) small first cycle schools in the Portuguese school network (for example, when one teacher leaves a school of two staff members, this means a 50% turnover), has no doubt to be kept in mind when indicating and comparing its stability values to other countries’.
64. In terms of stability of the school head teacher, Portugal ranks lowest (33%) relative to the other countries. This figure, being to a large extent the direct result from a low stability in the staffing of Portuguese basic education schools, is also to be interpreted in the light of a formal policy. This policy allows for the election of the head teacher, every one year, in the 1st cycle schools (grades 5 and 6), and every two years, in the 2nd cycle (grades 5 and 6) as well as in the comprehensive schools (grades 5 up to 9).

65. However, since head teachers can be reelected after their first mandate, ca. 50% of the schools keep their head teachers for longer periods.

**School leadership**

66. Formal and explicit job specifications for head teachers in Portuguese schools cover two quite different situations particularly when time investment is the issue.

67. In 1st cycle schools (grades 1-4) the “school director”, although expected to invest some time in managerial activities, is formally appointed to a full-time (100%) teaching-position.

68. In 2nd cycle (grades 5 and 6) and comprehensive schools (grades 5-9) the “school board chairperson” is formally employed both for teaching and managerial activities. Depending on the size of the school (in number of pupils), 2nd cycle head teachers are to be appointed teaching duties from about 50% down to 0% of their total working time.

69. Portuguese pupils attend schools led by head teachers, whose teaching duties represent quite more than half (67%) of their total working time. As, at the same time, the available time for managerial activities represents just 33%, data from the study provide us with an overall average of time spent within the system, highlighting a rather strong focus on the pedagogic role of the head teacher.

70. Lately, in the new school management perspective (Decree 172/91), particular attention is being paid to the professional training of head teachers for managerial functions as a means of definitely supporting the improvement of the organizational, administrative and financial management of schools.

71. Portuguese schools are reckoned to rank average in the self-estimated time head teachers spend on educational leadership (as a matter of fact the report points out that countries are remarkably uniform in the relative emphasis on educational leadership).

72. Within the structural changes introduced by the reform of school administration, of which autonomy and the new responsibilities granted to schools are the most expressive ones, some expectations on pedagogic leadership are to be developed. Portuguese schools are currently expected to actually exercise their new responsibilities as regards pedagogic organization and pedagogical functioning. This comprises the management of curricula, syllabi and educational activities, the student assessment, their tutoring and guidance, the management of school spaces and learning time, and the teaching staff education and management (Law 43/89), as a means of operationalizing “pedagogical leadership”.

**Staff cooperation**

73. Entire staff meetings convened by the head teacher are reported to occur rather infrequently in Portuguese basic education schools. In fact, rather than to meet with the entire school staff, Portuguese head teachers are expected to convene sector-based staff meetings to deal about pedagogic, administrative or organizational matters.

74. Average figures point out to a very slight majority (51%) of pupils attending schools in which such formal meetings take place every month. Most other pupils (39%) attend schools in which meetings
of the entire staff are convened from one to four times a year. The observed distribution of values is to be directly associated to the different organization categories of basic schools involved in the present study.

75. As far as informal co-operation is concerned, Portugal performs much closer to the other countries, with 65% of pupils attending schools in which informal meetings take place every week or more often. Once again, this average value hides two rather different co-operative practices among teachers depending on the organization category of the schools involved. When analysed per stratum, data indicate that 57% of the pupils attending 1st cycle schools (grades 1-4), in which informal co-operation occurs every week or more.

76. The great majority of pupils (83%) in 2nd cycle and comprehensive schools (grades 5 and 6) regularly benefit from such frequent co-operation among teachers.

Evaluation

77. In the current Portuguese curricular context, the assessment of basic education pupils (grades 1-9) focuses on the realization of the general objectives of each cycle of studies and on the specific objectives of each subject area or subject (Norm 98-A/92).

78. In its capacity as a regulator of educational practices, assessment in its various modalities, formative, summative, standard and specialized, must be carried out in a systematic and continuous way. These are aimed at: the choice of the most appropriate methodologies and educational resources; the necessary curricular adaptations; the right response to the pupils’ special educational needs; adequate interaction between the teachers and the other parties involved in the educational process; the pupils’ judgements about their learning process and at the introduction of curricular modifications or procedures, deemed necessary for the improvement quality of the education system.

79. Accordingly, pupil evaluation practices in Portuguese basic schools show a quite high level of attention paid to individualized teaching (91% of pupils attend schools where information from pupils’ records is used for preparing individual study plans; 81% of pupils are in schools in which such information is used for selecting pupils for special programs). The use of pupils’ records for evaluating school is not such an usual practice among schools, Portugal ranking average when compared with other countries.

Differentiation

80. Under the terms of their late pedagogic autonomy, Portuguese schools are to develop mechanisms allowing for the detection in due time of the basic difficulties, different learning paces or other needs of pupils that call for specific adjustment and support educational measures and procedures. Thus “streaming” policies within classes are adopted as a general strategy for temporarily supporting pupils’ learning difficulties: individual tutoring; aptitude grouping for a period of time; specific scheduling arrangements; or alternative curricula are some of the different modalities of support recommended.

81. However, the reported values for ability grouping within classes point out to a considerable proportion (28%) of Portuguese pupils attending basic schools where this differentiation practice appears to be used on a permanent basis.

Achievement orientation

82. According to the present study, Portuguese basic schools are reported to be rather high achievement oriented. This is the case as far as comparing pupils’ achievement over time (54%) is concerned, as in terms of comparisons among schools (Portugal along with Greece rank high with a value of 35% of pupils in schools which compare pupils’ results among schools).
Within the current educational administration reform, autonomy being defined as the school’s capability to design and put into practice its own specific development programme (Decree 43/89), schools are granted extended responsibilities namely in the pedagogic field. Carrying out these responsibilities effectively calls upon a deeper knowledge of the school identity, according to such dimensions as its context, available resources, its functioning, and its pupils’ results. In this context, data on pupil assessment, along with much broader school information, are collected by individual schools on a yearly basis, thus enabling them to compare academic performance of their pupils within and among schools. Though this is not a generalized procedure, more and more schools are encouraged to involve in self-regulatory projects, based on the control of their information.

Another reported aspect of this high achievement orientation is the setting of achievement standards for individual pupils. According to the findings, 54% of the Portuguese pupils attend basic schools that follow this practice.

As far as standard setting is concerned, the reported values appear to be surprisingly high for the Portuguese context, therefore demanding particular caution in their interpretation. As referred in the main body of the report, Portuguese head teachers (as their colleagues in the other countries) seem to “have interpreted ‘setting standards’ relatively loosely, also judging attainment targets as such”.

Parent involvement

For reasons of community integration, the Comprehensive Law of Education (Law 46/86) determines that school administration and management be guided by the principles of democracy, representativeness and participation of everyone involved in the educational process, in the respect for school autonomy. In line with these general guidelines, the new model for basic and secondary schools’ management attempts “to reconcile the strict requirements of representation, democracy and community involvement” and “the imperatives of a stable, efficient and responsible administration and management” (Decree 172/91).

Within an enlarged “educational community”-concept of school, parents are expected to become active partners.

Actually, reported practices regarding parent involvement in school matters as estimated by Portuguese head teachers illustrate how weak their role as partners really is.

Data point out to an extensive and regular practice in most basic schools of informing parents on the progress of their children (98%), on the nature of the school’s objectives (83%) and pedagogical mission (79%). Regularly engaging parents in supporting their children’s learning at home is also reported to be a common practice among a majority (78%) of Portuguese basic schools.

As far as parents’ involvement in the school decision-making process is concerned, figures point out to a rather low and differentiated degree of participation: parents’ involvement greatly varying with the very object of their participation. Among such co-decision-making areas as curriculum, school planning, finance, personnel, and organization, parental involvement in curricular issues scores lowest (14%), whereas school personnel and financial issues reveal a much higher involvement from parents (62%; 55%). The weight of first cycle schools in the sample has to be mentioned to explain the differences in those scores. In one or two teacher schools, the many small ones, parents often try to solve some of the problems teachers face and cannot solve. Therefore, in many cases, parents get someone to cooperate in the supervision of their children’s lunch, to support teachers in school trips, to be in charge of janitor-services, etc. In these cases it implies extra personnel and financial resources.
Conclusion

91. School surveys within the context of an international education indicator project exist at the border of educational research and education statistics. When preparing and developing such surveys in task forces and networks there is a continuous struggle to concentrate on relatively simple core issues, or rather attempt to reduce quite complicated phenomena to their bare essence. The fact that these attempts are always only partly successful maybe points to a specific added value of these surveys, namely the provision of more in depth qualitative comments on what is really going on in the educational structures, cultures and policies of the member states.

92. In the current chapter the research literature on school effectiveness and educational productivity has once again been referred to as a remaining source of inspiration in selecting process indicators that have relevance with respect to educational outcomes.

93. The available knowledge base is stronger for the primary and lower secondary level than for higher ISCED-levels, although there is a beginning of a set of recommendations for vocational upper secondary education as well (see this section’s introduction). Broadened perspective on organizational effectiveness may be particularly relevant for the choice of process indicators at the higher levels. For the rest the procedural solution of work in international task forces and Networks is a quite potent solution to capture issues that are policy-relevant from an above-national point of view.

94. Particularly when capturing characteristics of the immediate learning environment, measuring student’s perceptions is to be seen as an improvement of the methodology of school surveys. As long as such efforts are strictly targeted the increased practical burden of data collection, using within school samples of students, can remain within acceptable limits. Including the questioning of teachers as part of school surveys should also be considered for future applications.

95. The actual use and interpretation of school process indicators, obtained from surveys, has been given specific attention in this chapter. Limitations with respect to a strong instrumental and “causal” interpretation of even those indicators that have been derived from the knowledge on educational productivity and effectiveness call for a more modest type of interpretation. Interpretation in terms of “educational good practice” has been illustrated on the basis material from the Network –C primary school survey (1996). As the comparison of the “profiles” from three countries shows us there is great variability on some core process indicators. The added country-comments provide flesh and blood to these figures, indicating that they should be seen against structural, cultural and demographic characteristics of national education systems.

96. Despite the limitations that were mentioned, and particularly when some of the new perspectives on methodology and interpretation are followed, school surveys are considered essential in their function of “opening the black box of schooling”, and the principal mechanism to provide new insights into education processes across the world.

References


King (19..)


Os, W. van (1999). *Bruikbaarheid en Effectiviteit van Studentenoordelen over het Onderwijs* [Usefulness of Student Ratings]. Amsterdam: Free University of Amsterdam


163
INDICATORS ON TEACHERS AND THE CONDITIONS OF TEACHING

Maria A. Hendriks, University of Twente, The Netherlands
Thomas M. Smith, OECD

Introduction

1. Although technology has changed the way that much of the developed world conducts its economic affairs, the technology of teaching remains primarily a face to face interaction between teachers and students. The considerable share of total educational expenditure that is dedicated to teacher compensation alone necessitates the close monitoring of how this critical human resource is utilised in the education enterprise. Teaching staff represents about three per cent of the total labour force, on average, in OECD Member countries, while the percentage of all staff employed in education amounts to more than five (OECD, 1998). Teacher salaries continue to be the largest single component of the cost of providing education (OECD, 1998). Policy makers are no longer content, however, to use educational expenditure and teacher salaries as the sole measures of the investment in teaching. The need to find ways of improving educational quality at a time of tight public budgets has raised the relevance of indicators of teacher quality and the conditions of teachers’ work to an unprecedented level.

Policy relevance

2. Ensuring that there will be enough skilled teachers to educate all children is an important policy concern in all OECD countries. The supply and demand of teachers is affected by demographic factors shaping the size of the labour force, as well as factors influencing the attractiveness of teaching relative to other professions. Many OECD countries are facing the problem of a greying teaching force, as a large number of teachers will be reaching retirement age in the next two decades (OECD, 1996b). At this moment, about 40 per cent of teachers in public primary and secondary education are between 40 and 49 years of age. Although this does not necessarily imply that there will be a teacher “shortage”, it points to the increasing need to monitor the factors influencing teacher supply. Efforts to recruit new teachers, as well as to retain current teachers, depend highly on the conditions of teachers work (OECD, 1998). Factors underlying the attractiveness of teaching in different countries, relative to other professions, can inform policymakers about alternative models for structuring incentives to teach.

3. Although a wide range of external factors influence teaching effectiveness, the training and background of individual teachers is linked to their ability to provide quality instruction. Empirical findings, coming out of school effectiveness research, suggest that differences between teachers account for more variation in pupil achievement than differences between schools, that is, teacher effects may be more important than school effects (Luyten & Snijders, 1996).

4. The quality of initial teacher training, as well as the attractiveness of teaching relative to other professions are two factors shaping the effectiveness of new teachers. Opportunities for professional development, both pedagogical and subject related, can help teachers already in the classroom maintain or improve their teaching skill. Educational commentators in a number of countries are currently advocating improvements and reforms in teacher training, licensing and professional development in order to improve the quality of instruction at all levels (Darling-Hammond, 1997; Sykes, 1999). These debates could be better informed through comparable, cross-country indicators on teacher training and the condition of teachers’ work.
State of the art on staffing indicators

5. Network C has prepared indicators for each edition *Education at a Glance* to date. Appendix A lists the indicators relating to staffing that have been published since the General Assembly in 1995. The current staffing database is heavily dependent upon the types of data available from administrative records in OECD countries. Data on teacher compensation and teaching time typically come from collective bargaining agreements, from laws or administrative regulations on teacher staffing and the conditions of work or from payroll systems. Each year these data, in addition to data on instructional time and the curriculum, are reported to OECD through the Network C Teacher and Curriculum Survey. Since the start of INES project, the Network has invested considerable time in developing definitions and reporting strategies in order to improve the comparability of these data. The Network’s ability to collect data on the conditions of teacher’s work is constrained, however, by variability in the types of information that national or regional administrations collect in different countries. For example, comparable data on teacher working time, i.e., the amount of time that they work each week to receive their regular salary, is unavailable in a consistent form from national administrations. Because of these constraints, the Network has sought alternatives to administrative data for its indicator development.

6. While the Network, in co-operation with the OECD Secretariat, has made use of existing sample survey data to develop indicators on staffing (see the indicators based on TIMSS and IEA SITES in the 2000 edition of *Education at a Glance*), these sources have been inadequate to address important policy issues. Consequently, the Network has developed its own school surveys, the first of which was administered in 1995 at the primary level. This survey examined school level staffing data on stability in school staff, school leadership and staff co-operation, the results of which were published in the 1996 edition of *Education at a Glance*. A Survey of School at the Upper Secondary Level is planned for administration in 2001. The main issues to be examined in this survey are how schools organise the transition from upper secondary school to further education or labour market, the quality of education, the human resources and the availability and use of information and communication technology.

7. Data collected through this Survey will be used to examine human resources in several ways:

   - differences in the allocation of teacher resources to different types of upper secondary education;
   - pre-service training level of teachers in the school;
   - barriers to recruiting or keeping quality teachers;
   - forms of in-service training opportunities available to teachers;
   - ways in which in-service training opportunities are linked to education reform goals

8. These issues were selected by the Network, and approved by the INES Steering Group, because they address important issues in education policy for OECD Member countries.

**Improvement and extension of the current indicator set**

9. Although much of the indicator development within the Network has focused on issues defined by Network participants as reflecting important national issues, such an ad-hoc approach is not appropriate for long term planning of teacher indicators. Three purposes/perspectives can be useful for evaluating the policy relevance of staffing indicators in relation to the efficacy and efficiency of education systems:

   1. *Investment in the quantity of human resources.* Descriptive statistics on the size of the teaching force and the non teaching force can provide benchmarks of the magnitude of the investments of countries in human resources for education, in other than strictly monetary terms. Measures of the number of instructional staff, school management personnel, and non-instructional staff per student can be viewed as a counterpart to financial indicators on the relative size of educational budgets and expenditure per pupil. Whereas expenditure data
measure the total cost of carrying out education activities, personnel data measure the quantity of human resources allocated to education activities. Observed differences in staffing patterns are starting points for examining alternative staffing models.

2. **Political economy of the teaching profession.** This perspective regards the motivation and satisfaction of the members of an organisation (i.e., school) as either a criterion for organisational effectiveness or as a means to the ultimate objective of organisational productivity. Much of the current debate on the status of the teaching profession and the concern over the conditions of teachers’ work can be interpreted as a desire to influence various factors affecting the supply of qualified teachers. In a number of OECD countries a decline in teacher satisfaction and a decrease in commitment have been linked to 1) a high workload, 2) a lack of career opportunities, 3) inadequate compensations and 4) the generally low standing of the profession.

Indicators that identify key differences in the conditions of teaching, relative to other occupations, and the potential sources of supply of new teachers are useful to policymakers and educational planners who are trying to ensure the adequate supply of quality teachers.

1. **Educational effectiveness.** Educational effectiveness perspective focuses specifically on the educational inputs or processes that have been shown to affect educational goal attainment. With respect to staffing, strong educational leadership and a success-oriented ethos in schools have been shown positively associated with educational achievement. With regard to teachers, this perspective focuses on those teacher characteristics and instructional behaviour that are directly linked to student achievement.

10. A high level of motivation and strong a commitment to teaching have been suggested as the qualities that make all other factors influencing teaching effectiveness possible (OECD, 1994; Sergiovanni, 1987). Other teacher characteristics linked to educational effectiveness appear to matter in the pursuit of quality in education include teacher preparedness, both in subject matter knowledge (Monk, 1994) and pedagogical knowledge (Scheerens, 1995); teachers’ intentions, objectives and expectations (Brophy & Good, 1986); as well as subject-specific didactic skills (OECD, 1994). Indicators that identify both organisational differences in the way that teachers receive their initial and continuing training and how classroom instruction and curriculum are organised can help us to understand what underlies differences in organisational effectiveness. Indicators of that reflect how schools are organised and governed can provide policymakers with alternative models for school level decision making.

11. These perspectives can be utilised to aid decisions about how to improve and extend the current indicator set. Over the past several years, the Network has undertaken several developmental activities in order to 1) develop a framework for designing indicators on teacher quality (including pre-service training, qualifications, induction and professional development) and 2) extend the current indicator set to tertiary staffing. Each of these activities will be discussed below, as well as opportunities to further develop work in these areas.

**Conceptualisation, development and preparation of indicators on teacher quality**

12. Wide-spread recognition of the importance of teacher quality to the educational outcomes of students has led to a variety of policies and proposals to enhance pre-service training and on-going professional development, as well as alternative methods for certifying teachers. A review of the literature on teacher quality, pre-service training (prior to being hired as a classroom teacher), and on-going professional development (often referred to as in-service training) has served as a foundation for conceptualising indicators of pre-service training and professional development.
13. One way to gain insight into different models of educational organisation is to examine the ways in which teachers are trained in different countries. While setting the standards for entry into the teaching profession too low can introduce individuals with inadequate subject matter or pedagogical skills into the classroom, some researchers have argued that setting the barriers to entry too high can discourage high ability people from considering teaching (Ballou and Podgursky 1997). Cross country comparisons of the level, length and structure of teacher training programs, paired with information on teacher compensation and working conditions, can suggest alternative models for how countries attempt to ensure an adequate supply of new teachers.

14. It is also important to distinguish between the requirements to qualify as a teacher today (from initial teacher education programs) and the actual education levels of the teaching workforce. It cannot be assumed that all those employed have completed a level of education equivalent to what is required today. For many countries, pre-service teaching requirements have changed substantially over time. While it is typical now to require a university-level (ISCED 5A) degree from prospective teachers, this has not always been the case. Furthermore, it is common in some countries for teachers to earn higher level degrees while teaching. For this reason, current pre-service requirements cannot serve as a proxy for the qualifications currently held by the teaching force.

15. The adequate matching of teachers’ skills to students’ educational needs is also a dimension of educational quality that is gaining in importance in a number of countries. While the proportion of teachers that are “certified” or “fully qualified” to teach may be legally equivalent to 100%, this does not guarantee that the match between teacher background and students’ educational needs is adequate. In the United States, it is common for teachers without specific training in a subject to teach that subject (Ingersoll 1999). Cross country comparisons of the match between a teacher’s training and the subjects they teach can help us form a more complete picture of how teacher supply and demand are balanced in different countries.

16. Research on teacher knowledge and preparation suggests that some kinds of preparation may make more of a difference than other kinds. For example, some studies show that standard knowledge of subject matter is important up to a threshold level: implying that teachers teaching courses in subjects that they have inadequate training are less effective than teachers who have been prepared to teach a given subject. Beyond that threshold, however, it has been argued that greater preparation in child development, learning theory, curriculum content and teaching methods has a stronger influence on teacher effectiveness than additional subject matter preparation (Darling-Hammond, 1997b, 1999). While this literature is suggestive, the optimal mix of subject matter and pedagogical training is not known. Quantitative and qualitative data on how teachers are prepared in these areas in different countries would help to clarify the debate.

17. The public education systems in several countries also allow alternative routes to teaching in order to facilitate entry into the teaching force for individuals who have not attended traditional teacher education programmes. Darling Hammond (1990) makes a distinction between (AR) alternate route programmes, which change the route to certification but not the certification standards themselves, and (AC) alternative certification programmes, which change the rules by which certification is granted. In the United States, for example, the vast majority of teachers hold regular certification, although some states have implemented new types of teaching credentials, including “emergency credentials” to broaden the pool of potential teachers (Sherman, 1998). In the Netherlands recruits who have experience outside the education system and want to become teachers can undertake in-service training during their first two years on the job in order to meet the initial qualifications for new teachers. Until that time, they hold temporary certification. As OECD countries become increasingly concerned about the source of supply of new teachers, information on successful alternative routes can be of great benefit to policymakers.

18. There is also growing interest in a number of countries in finding ways to increasing the effectiveness of teacher induction programmes. These programs, which typically involve mentorships with
more experienced teachers and other forms of support, help guide new teachers in curriculum planning, classroom management, academic knowledge and diverse teaching methods. According to Huling-Austin’s review (1990), the goals of induction programmes typically include improving teacher performance, increasing the retention of promising beginning teachers, promoting the personal well-being of beginning teachers, transmitting the culture of the system to beginning teachers and satisfying mandated requirements related to induction and certification. The support offered by these programmes varies widely, and there is some concern about how mentors are trained (Darling-Hammond, 1997, 1999). The inclusion of the quantifiable aspects of “induction” practices in indicators on pre-service training can provide insight into how different countries facilitate the transition of new teachers into the classroom.

The current state of indicators on pre-service training

19. In 1995, Network C prepared an indicator on the duration of initial teacher education, i.e. the number of years of education required to begin work as a fully qualified teacher according to the current formal policy in different countries. For the 2000 edition of EAG, Network C prepared an indicator on the current educational and pedagogical requirements for new teachers in OECD countries. For each ISCED level (0,1,2,3), the indicator presents the type of educational qualification required to become a teacher today (e.g., ISCED 3, 5B, 5A), the number of full-time equivalent years of teacher training, and whether or not pedagogical training is undertaken consecutively or concurrently with subject matter instruction. Information on mandatory induction periods and/or work experience required in some subject fields to become a fully qualified teacher is also presented. The information was collected through qualitative descriptions of the systems of pre-service training provided by Network C members.

20. The key results presented in the 2000 edition of Education at a Glance include:

- with regard to the duration of teacher training:
  
  At the primary level, the duration of pre-service training for teachers varies from 3 years to 5.5 years. At the lower secondary level, the duration of pre-service training tends to be similar to the primary level, although a number of countries require longer periods of study. In most countries, pre-service training is longer for teachers at the upper secondary level than the duration of pre-service training for teachers at the primary and lower secondary levels.

  Furthermore, the length of teacher training can be affected by the sequencing of pedagogical instruction, that is, whether or not pedagogical instruction is consecutive or concurrent with subject matter instruction. Completing a full academic qualification prior to pedagogical studies tends to take longer than if pedagogy and subject matter are combined. At the primary and lower secondary levels of education, the concurrent model of pre-service training is utilised in most countries. In this model, pedagogical training is taken at the same time as studies related to the specific subject area to be taught. At the upper secondary (general) level, the organisation of pre-service training varies more across countries.

- with regard to practical training periods

  In Germany, prospective teachers are required to work 18 - 24 months in a practical and professional training period (preparatory service) as part of their training. In most other countries the “teaching” component of teacher education tends to be shorter.

- with regard to induction periods and/or required years of work experience

  Mandatory induction periods (a period during which new teachers are intensively followed and guided) are required in France and Italy. In France, for example, at ISCED 2 and 3 prospective teachers who already hold the academic degree are required to follow a pedagogical teacher training for 1 or 2 years. The 2-year teacher-training programme is a special, intensive preparation where in the second the student teacher takes responsibility for a class in a lycée. This option is intended for prospective teachers wishing to take the competitive examination for
the higher level certification Agrégation. In Italy, teachers in upper secondary education are required after their university degree to follow an individual post-graduate study to prepare for a test of aptitude and a competitive entrance examination (the Esame di Abilitazione and the Concorso), which must be successfully completed before they can become fully qualified.

Related work experience is required to become a vocational teacher in Austria, Denmark, Finland and Norway. In Finland, for example, individuals wishing to become an upper secondary vocational teacher must hold either a higher vocational diploma or Master’s degree from university and have 2-3 years of related work experience.

Developmental activities on pre-service training

21. In addition to developing an indicator on the current requirements to become a teacher today, the feasibility of collecting data on the current qualifications of teachers (stock comparison) at the primary and secondary levels was explored, including:

- the ISCED-97 level of education to which different teacher qualifications have been mapped;
- the total (minimum) number of full-time equivalent (FTE) years of primary, secondary and tertiary education needed to start working with this qualification as a (fully) qualified teacher;
- the total (minimum) number of FTE years of pre-service teacher education needed to start working with this qualification as a fully qualified teacher;
- the number of classroom teachers (head counts) at each ISCED-97 level.

22. In order to investigate the availability of data on the current qualifications of teachers at the national level Network C undertook a 'quick survey'. The results showed that few countries collect data on the level of educational qualifications held by the current teaching force. This led the Network to explore alternative sources for data. In the 2000 edition of EAG, data on the educational qualifications of 8th grade students’ mathematics teachers were published. While the qualification level of lower secondary mathematics teachers is not necessarily representative of the qualification level of all teachers, this indicator does show the range of qualifications held by both new and experienced teachers in an important subject area. The following patterns could be identified in the TIMSS data:

- The vast majority of eighth grade students are taught mathematics by teachers who have at least a university-level degree or its equivalent. In 9 out of 18 countries, more than 90 % percent of students’ mathematics teachers have at least a bachelor’s or higher degree. Notable exceptions are Hungary and the Netherlands, where approximately 8 out of 10 students are taught by teachers with a secondary qualification and between 1 and 4 years of teacher training.
- Most mathematics teachers have at least some formal teacher training in addition to (or as part of) their educational qualifications. The exception is Greece where 88% hold a bachelor’s degree or equivalent but have had no teacher training.
- With the exception of Australia, the Czech Republic, France and the United States, it is relatively uncommon for 8th grade mathematics students to be taught by teachers who have a master’s degree or Ph.D.
- While in some countries older teachers have lower qualifications than younger teachers do, the opposite is true in other countries.
- Teachers with higher educational qualifications are more likely to be assigned to teach students with higher mathematics ability in most countries. This may reflect the practice of giving teachers with higher qualifications the more advanced classes.

23. Although concerns have been raised about the comparability of the coding of teacher qualifications in TIMSS, this is the most comprehensive source of data on the qualifications of the teachers.
force available today. While both PISA and the Network C upper secondary school survey have staffing questions as part of the questionnaire that will be administered to school principals, these data probably will not provide more accurate information. While many schools have teachers’ resume or academic records on file, a school-level database would be needed to determine how many teachers actually hold different levels of degrees, as well as the fields in which the degrees are held (Sherman, 1998). Questions about teacher qualifications that can be asked of school administrators need to be designed in a way in which they don’t require a database of qualifications at the school level.

Questions about teacher qualifications in the Network C Survey of Schools at the Upper Secondary Level

24. In the Network C Survey of Schools at the Upper Secondary Level, two questions regarding teacher qualifications have been proposed:

- the percentage of teachers in a certain programme in the school that are not “fully qualified to teach” and;
- the percentage of teachers qualified to teach but teaching courses in a certain programme for which they were not trained and/or qualified.

25. In the first question the meaning of “fully qualified” is linked to the formal regulations in each country for different educational programmes at the upper secondary level. It is recognised, however, that this linkage introduces comparability problems. For example, how stringent should a school administrator be in deciding whether or not a teacher has an educational qualification in a particular subject? Would it cover only those teachers teaching mathematics that had maths as their primary field of study in their pre-service training or should minor or secondary fields of study count as well? If a teacher has an educational qualification in pedagogy, but a specialisation in the teaching of mathematics and science, should they be counted under both of these categories? Furthermore, having a relatively high number of teachers certified in mathematics does not guarantee that all students are taught maths by a teacher knowledgeable in this subject area. If we want to know how many teachers are teaching subjects that they do not have adequate knowledge of, or conversely how many students are taught a particular subject by a teacher that has inadequately subject matter knowledge, a targeted question to teachers would be necessary.

26. For the second question, the administrator will be asked to indicate the percentage of class periods covered by a teacher who was not trained and/or qualified for covering them during last month. As the responses to this question will be rather subjective, they will need to be interpreted with caution. Because schools probably won’t have sufficient systematic data on this topic the responses will be illustrative of head teacher/principal’s perceptions of the adequacy of teacher training in their school.

Indicators that could be explored in the medium term

27. In order to provide greater contextual information on how different countries organise pre-service training for teachers, the following areas could be developed further:

- regarding pre-service education:
  - the balance of curriculum content presented in pre-service education (e.g. the time spend on subject matter versus the time spend on pedagogical training);
  - the amount of time spent on theory and practice;
  - the costs of pre-service training, i.e. the weight of teachers’ education in the education budget.
- regarding qualifications and certification:
  - the percentage of teachers who have a degree in the disciplinary area in which they are teaching
– the percentage of courses in different fields (e.g., mathematics) taught by a teacher without an educational qualification/specialisation in that field;

– the percentage of teachers that hold different types of certification (including temporary and emergency certification.

regarding teacher induction programmes;

– the percentage of new teachers that have participated in teacher induction programmes;

– the type of experiences that teachers have in teacher induction programmes

– the average length of participation in teacher induction programmes.

Issues regarding data-availability

28. It is likely that questions relating to teacher certification and the frequency and length of teacher induction programmes could be answered in a school survey by head teachers/principals. Questions relating to academic degrees, fields of study, pre-service training and the content and learning experiences in induction programmes would need to be gathered through a sample survey of teachers.

Professional development

29. Enhancing the professional development of the current teaching force is a prominent reform issue in a number of OECD countries. Professional development experiences in the past have been synonymous with “in-service” workshops, whereby knowledge is presented to teachers by an outside “expert” in the field and teachers assume the role of passive learners. While this type of direct instruction may be effective for conveying factual information, training in well-defined skills, and introducing teachers to new ideas, researchers argue that such instruction provides a questionable means of promoting critical thinking or deeper levels of understanding (Sykes, 1996; Ball, 1996). In the new paradigm, staff development is a shared, public process; promotes sustained interaction, emphasises substantive school related issues; relies on internal expertise, expects teachers to be active participants; emphasises the why as well as the how of teaching; articulates a theoretical research base and anticipates that lasting change will be a slow process (Hawley and Valli, 1999). The new paradigm would encompass all kinds of external and internal support, varying from in-service training in the traditional sense to advising, coaching, reciprocal class visits or the support of internal coaches.

30. Professional development has traditionally focussed on the characteristics of teachers and teaching. As the effectiveness of education systems is increasingly called into question, however, there has been a tendency to be concerned more directly about student learning. This has affected teacher development in different ways. In countries where there is particular concern about insufficiently high standards, attempts have been made to link in-service training and professional development to the raising of pupil achievement levels. In countries like Japan, where students do well on tests and which have an extremely strong national system for maintaining standards, in-service training and professional development are meant to create more diversity and creativity in the system. Here the reforming efforts are aimed at a more flexible thinking among students and teachers (OECD, 1998c, p. 23-24).

31. Professional development is a new area for Network C. So far no indicators have been published. As a first step in the developmental process, Network C members provided country descriptions of the national system of in-service training and professional development, including the following aspects:

– The policy purpose of professional development. In some countries teachers are obliged to participate in professional development, in other countries teachers have the right and participation is voluntary. Furthermore, in some countries professional development is regulated by law, while in other countries no specific legislation exists.
The purpose of in-service training and professional development. The development of teachers beyond their initial training can serve a number of different objectives, varying from updating individuals' knowledge of a subject in the light of a recent advance to enabling schools to develop and apply new strategies concerning the curriculum and other aspects of teaching practice. Professional development activities can also be specifically targeted to help weaker teachers enhance their effectiveness (for an overview see also OECD 1998c, p. 33-34);

The content of in-service training and professional development activities. This encompasses the introduction of new teaching techniques or curriculum content, the expansion and consolidation of the pedagogical, psychological and didactic knowledge, intervention programmes for disadvantaged pupils, guidance and counselling and information and communication technology;

The formats of in-service training and professional development. Training courses range from short one-day or evening seminars to a number of days in blocks or sequence. In some countries, professional development activities they may also take place during the summer. Many different other forms are possible, for example, distance study, self-tuition, team teaching, school-wide or grade-wide school development programmes.

The providers of in-service training and professional development. Providers could be, among others, university teacher training institutes and higher education institutes, education support centres, teacher associations, private contractors and individual schools.

Methods of funding in-service training and professional development In some countries, such as Sweden, there are no longer earmarked funds for professional development and it is up to the municipalities and/or the schools to decide how much will be allocated to these activities. In other countries, the Ministry of Education allocates funding and/or grants a specific earmarked budget to the responsible authorities.

32. From the country descriptions prepared by Network C participants, it is clear that only a very small number of countries have a formal national policy on professional development. In addition, the shift in focus from 'in-service workshops and courses' to professional development makes it less relevant to collect data on traditional measures, such as the percentage of teachers involved in in-service training activities and the mean number of hours that teachers are involved in-service training.

33. For the medium and long term, developmental work could be undertaken on one or more of the following aspects of teacher professional development:

- The types of professional development undertaken in schools in different countries and the extent to which teachers participate in these activities (e.g., number of hours and kinds of activities).
- At what level are decisions made about the design of professional development programmes? Does it occur at the school site or at a higher administrative level? What role do teachers play in determining the professional development content and strategies?
- The sources of funding for professional development. How are professional development resources allocated? What do they support (e.g., conference registration, tuition subsidies, stipends, materials, travel reimbursement, outside experts, substitutes, materials, a portion of administrators’ or specialists salary devoted to professional development)?
- The extent to which professional development is aligned with local, state or national standards or curriculum frameworks.
- The providers of professional development (e.g., in-house staff, external providers).

68 For an elaborated view the reader is referred to expert review on Human Resources, prepared by Sherman (1998) for the Survey of Schools at the Upper Secondary Level.
The policies that support or impede collegial activities such as mentoring, peer coaching, and time for teachers to share, discuss, and reflect on their practices.

The kinds of follow-up that are built into the professional development activities offered (e.g., classroom observations, consultations, mentoring, or follow-up workshops).

The incentives and supports provided for teachers to participate in professional development.

The measurement of the outcomes of professional development.

A number of the above mentioned issues can be addressed through a questionnaire administered to school principals/directors (i.e. the more policy related issues such as decision making or funding, as well as the kind of activities offered and the methods used to evaluate outcomes). Areas which may be more difficult for a school administrator to answer the types of professional development activities teachers have participated in, the extent of their participation, and the impact of these activities on their classroom practice. Obtaining such information would require a sample survey of teachers in different countries.

Professional development and the Network C Upper Secondary School Survey

In the Network C Survey of Schools at the Upper Secondary Level, four questions on professional development will be asked of the school principals/directors:

- In the past 12 months have teachers in the programme to which this questionnaire relates participated in the following activities related to education?

- The activities are: courses and workshops, conferences, observational visits to other schools, regularly scheduled collaboration among teachers on issues of instruction, mentoring and/or peer observation and coaching, individual or collaborative research on a topic related to education, participation in a network of teachers, visits to companies lasting longer than one day.

- Does the your school provide teachers with time to professional development during regular contact hours (answer categories: yes or no)?

- Consider the previous school year. Please indicate the percentage of teachers in the programme related to this questionnaire who spent at least two weeks at any of the activities as mentioned in the first question (answer categories: none, 1-10%, etc.)

- To what extent is professional development of teachers a priority in the policy of the school? (four answer categories varying from low or medium priority to exclusive top priority)

While these questions only examine a limited number of the core issues outlined above, the data resulting from the Upper Secondary School Survey should aid our understanding of how professional development is implemented at the school level in a number of OECD countries.

Extension of the current indicator set to tertiary staffing

When the INES project first started, internationally comparable information on tertiary education was quite limited, mostly to gross enrolment ratios and some rough estimates of education expenditure. Over the past few years the OECD has concentrated its efforts on improving the amount of policy-relevant information on tertiary education, as this level is seen as a critical component of the expansion of life long learning opportunities. The current edition of Education at a Glance examines tertiary issues relating to demographic changes in the age relevant population, access, participation and progress through tertiary programmes, expenditure on tertiary institutions and tertiary students, as well as graduate output and the economic returns to tertiary degree completion. Regarding staffing, however, indicators have been limited to those based on basic personnel and enrolment data, such as faculty-student ratios.
38. As issues relating to efficiency and quality in teaching have parallels between tertiary and lower levels of education, Network C has attempted a first cut at developing indicators of tertiary staffing by exploring the feasibility of expanding components of its primary and secondary data collection to the tertiary level. For reasons of practicality, the Network chose to start with ISCED-97 level 5A, which comprises mostly university-level and equivalent programmes.

39. Indicators that the Network felt were most adaptable to the tertiary level included the number of years of initial training for faculty, the number of (non-) teaching hours per year and faculty salaries and benefits. In consultation with tertiary experts with their respective countries, Network C members also proposed developmental work to develop the following indicators:

- the proportion of working time academic staff devote to teaching, research and administrative/other duties;
- the existence of or accessibility to teaching development and support units;
- the matching of institutional procedures for career advancement, promotion and recognition in teaching and research.

40. As staff at the tertiary level fulfils many different functions, the network decided to first develop a conceptual framework for collecting internationally comparable personnel data. The initial framework covered three basic categories of personnel at universities: academic staff, teaching/research assistants and different types of professional and support personnel. As the primary area of interest was in “tertiary teaching” it was agreed to start with academic staff and to develop comparable subcategories that would be meaningful for international comparisons.

41. In the international comparative study by the Carnegie Foundation for the Advancement of Teaching, five defining characteristics are mentioned when the diversity of the academic profession is addressed (Teichler, 1996). For the Network the most important characteristics to examine were:

- the constituting professional functions, i.e. research, teaching or any combination of both;
- the occupational ranks and tasks related to it, i.e. professors, intermediate-rank and junior staff.

The constituting professional functions

42. In the Carnegie study, research and teaching are viewed as the constituting functions of the academic profession. Administrative and service activities are viewed as additional and not characteristic for the identity of the academics. However, the role and scope of the academic profession has expanded significantly (Altbach, 1991). Many academics play much wider roles in society and tensions can be seen between the different tasks: teaching, research, administration, consultancy and community service (Altbach, 1991; Kogan, 1994).

43. In addition, changes in task have brought with them changes in academic structures and appointments. Part-time appointments, temporary full-time appointments, and full-time appointments restricted to either research or teaching have enabled the system to be more flexible (Kogan, 1994). In the Netherlands for instance, the holding of traditional academic ranks, with titles of assistant professor, associate professor and full professor accounted for fewer than half of all academic staff in 1991 (Geurts, Maassen & Van Vught, 1993). More than half of the academic staff held other titles such as teaching associates, research associates and research trainees, positions that accounted for only 16 percent of academic staff a decade ago. According to Geurts, Maassen and Van Vught in the coming years the rank of research associate will become increasingly attractive as a separate career track inside the universities. This trend is also seen in the United States and the United Kingdom (Kogan, 1994), where grants and contracts became more prominent in university resourcing in the 1960s and fully established during the 1970s. Consequently, universities appointed increasing numbers of non-tenured full-time researchers.
The occupational ranks and tasks

44. The rank or status within an institution or career ladder is also an important element of the heterogeneity of the academic profession (Teichler, 1996). While the titles vary, in all academic systems there are typically three of four specific ranks in the academic hierarchy with the most junior rank made without tenure, on a probationary basis (Altbach, 1991). In some systems, the probationary period is relatively short and the process of confirmation not very difficult. In other systems, i.e. the United States, it takes some more time to promote in rank and the process of tenure is of a competitive nature. On the other hand, in the American academic system it is assumed that most tenured academics will end their careers as full professors, the top academic rank. This in contrast to most other academic systems where typically there is one full professor in each department and an appointment to this rank is somewhat unusual.

Breakdowns of academic staff

45. In order to test the applicability of the Carnegie framework, breaking down academic staff by either the constituting tasks or by academic rank, and explore data availability at the system level in different OECD countries, two quick surveys on academic staff were administered to Network C members. The first quick survey focussed on data regarding academic staff. Two ways to breakdown academic staff were explored, i) by main tasks and ii) by academic rank. The results of the first survey showed that at the system level only data on ranks are available. Data on main tasks are not available from national statistics. To collect these data, an institution or faculty level questionnaire would be needed.

46. To further explore the types of tertiary staffing data available, it was agreed to collect additional detail on the characteristics underlying faculty ranks. The results of the second survey showed that faculty ranks, and the characteristics that underlie them, are quite different across countries, and in some cases within countries. In the Czech Republic, for example, there is a new Higher Education Law that only describes the positions of professor and senior lecturer. Each higher education institution is free to determine the other intermediate and lower positions (lecturer, assistant, lector and researcher). In the United Kingdom, higher education institutes have a considerable degree of freedom in determining the requirements for appointment and promotion. In addition in New Zealand, there are not national criteria for appointing and promoting academic staff. Each university is free to set its own guidelines.

47. Moreover, even for a subset of countries, a classification between junior staff, middle rank professors and full professors based on temporary and permanent contracts, as had been used in the Carnegie study, did not appear realistic. Tenure, or permanent appointment, as mentioned before, in itself is a complicated and multifaceted concept (Altbach, 1991) and is not always distinguished in national statistics.

48. Because the problems of comparability and the lack of data at the system level it was not possible to develop Tertiary staffing indicators in the short term. There remains, however, little comparable data on how tertiary education is organised in different countries, especially regarding the structure of a student’s learning programme, the instructional methods and types of technology utilised in the classroom, as well as how tertiary institutions are reacting to changes the composition of the student body that have resulted from the rapid expansion of tertiary education in recent year. Beyond broad policy statements, only a handful of countries collect data at the tertiary level that extend beyond the basics of how many students are enrolled and the number of personnel employed. In order to advance our understanding of how tertiary education is organised across countries, an International Survey of Tertiary-level Institutions might be considered. This would not be a trivial matter, however, as the diversity of institutional structures both within and between countries would require a relatively large sample of institutions, as well as a sample of faculty within institutions. Because of the large degree of heterogeneity across academic fields, it may be necessary to focus on a few that are of key interest to policy makers (e.g., physical science and engineering). This type of effort would require commitment beyond education Ministries, however, as the most of the burden would fall on educational institutions themselves.
Universities would need to be partners in this effort, with any data collected being relevant for university-level evaluation and management in addition to national benchmarking.

Conclusions

49. This chapter is intended to highlight both the conceptual and practical issues involved in developing staffing indicators that are both policy relevant and comparable across countries. Four broad categories of staffing indicators at the system and at school level have been distinguished: training, working conditions, staff characteristics and stability and mobility of education staff.

50. Currently available data form only a small sub-set of the data required to explore these topics in depth. Currently, the OECD/INES project is limited to basic descriptive information on teachers, including ‘the number of years of initial training’, ‘the type of initial training’, ‘the number of teaching hours per year’, ‘minimum and maximum salaries’, ‘years from minimum to maximum salary’, ‘additional benefits’ and ‘the age and gender distribution of teachers’. Lack of comparable data continues to limit the scope, and the policy relevance, of teacher indicators currently published in Education at a Glance. Indicators on personnel in tertiary education and indicators on professional development of teachers are still at the conceptual stage and further data development would require a substantial commitment from OECD Member countries. Given the lack of data available at the national level, however, further development in these areas are likely to be time consuming and costly. A teacher survey on pre-service training and professional development and a faculty survey on the conditions of academic’s work are two possible options. OECD Member countries should weigh the policy relevance of the types of data suggested in this chapter with the additional costs that mounting these surveys would entail.
REFERENCES


Ball, D.L. (March 1996). Teacher learning and the mathematics reform, what we think, what we know and what we need to learn". Phi Delta Kappan, 500-508.


APPENDIX A

Indicators relating to staffing that have been published since the General Assembly in 1995:

- teaching time (Network C teacher survey, yearly)
- teacher compensation (Network C teacher survey, yearly)
- student/teacher ratio (Technical Group, based on the yearly UOE data collection)
- staff employed in education (Technical Group, yearly)
- initial teacher education (Network C teacher survey, ‘95 and ‘00)
- teacher age and gender (Network C teacher survey, yearly, from ’00 onwards Technical Group through the UOE data collection)
- stability in primary school staff (Network C primary school survey ’95)
- school leadership at the primary level (Network C primary school survey ’95)
- utilisation of staff with different functions (Network C primary school survey ’95):
  - school level pupil/teacher ratio
  - professional support staff/teacher ratio
  - managerial support/classroom teacher ratio
  - managerial support/pupil ratio
MEASURING FUNCTIONAL (DE)CENTRALISATION

Joel Sherman, American Institutes for Research, United States
Jaap Scheerens, University of Twente, The Netherlands

Introduction

The chapter examines the conceptual basis of the “locus of decision making” questionnaire that has been developed in Network C of the INES-project in order to measure patterns of centralisation and decentralisation of educational systems. “Functional decentralisation” is seen as the most appropriate term to describe the concept underlying the measurement procedure. In several editions of Education at the Glance, indicators have been published which concentrate on the distribution of decision-making authority across administrative levels and, more particularly, on the autonomy of schools.

1. In this contribution the potential of the procedure to provide more specific information on decision-making with respect to educational sub-domains, like the curriculum, evaluation policies and others is illustrated by a more in-depth analyses concerning staffing.

2. In a final section the potential for future use of the locus of decision-making is discussed. A case is made for a more qualitative accompanying study that could provide further information on the actual policy goals and expectations concerning functional decentralisation.

Conceptual basis: functional decentralization

3. The OECD-INES procedure to measure the “locus of decision making” distinguishes three facets of decision-making:

   – the tier or administrative level where a decision is taken; this dimension is referred to as the locus of decision-making;
   – the amount of discretion, or the degree of autonomy of decision-making at a particular administrative level; this facet is called the mode of decision-making;
   – the particular element of educational administration which a decision concerns; this facet is referred to as the domain of decision-making.

4. These three facets can be related to existing categorisations in the relevant literature, although the use of central concepts is by no means consistent among authors and publications. Our three-dimensional conceptualisation is compared to the terminology as clarified by Bray (1994, p. 819) in an analysis of alternative meanings of centralisation and decentralisation.

5. The distinction between levels confirms to the concept of territorial decentralisation, defined as "the distribution of powers between different tiers of government”. In the operationalisation of this dimension we distinguished four tiers, to be further described in the section on methods.
6. Degrees of autonomy in decision making at a particular level are reflected in terms that refer to an increase in discretion. Again following Bray, deconcentration, delegation and devolution are modes of decision making in which an increased amount of decision-making authority resides at a lower level.

7. "Deconcentration is the process through which a central authority establishes field units, staffing them with its own officers".

8. "Delegation implies a stronger degree of decision making at the lower level. However, powers in a delegated system still basically rest with the central authority, which has chosen to "lend" them to a local one".

9. "Devolution is the most extreme form of decentralisation. Powers are formally held by local bodies, which do not need to seek approval for their actions" (ibid, p. 819).

10. In the operationalisation of this continuum of increasing autonomy, these abstract definitions were avoided and respondents were asked to indicate whether decisions could be taken within the framework determined by a higher level, in consultation with a higher level or in full autonomy.

11. In order to determine elements or domains of educational administration, many categorisation schemes are available in the literature (e.g. James, 1994; Winkler, 1989; Bacharach et al., 1990; Rideout and Ural, 1993). The common core of these categorisations is:
   a) an educational domain (goals, methods, curricula, evaluation procedures);
   b) an organisational, managerial and administrative domain (including human resource management);
   c) a dimension concerning finance and the way financial resources are applied.

12. In the operational classification that we chose four main categories were used, by subdividing area b (organisational) into two domains "planning structures" and "human resources", and including areas a and c.

13. The distinction between domains of decision-making in educational systems reflects in part Bray's use of the term "functional decentralisation" as cited from Rondinelli. "Functional decentralisation refers to the dispersal of control over particular activities" (Bray, 1994, p. 819). From the examples that he provides, however, it is not clear whether in functional decentralisation an exhaustive set of domains of educational decision-making is referred to, which is the purpose of the categorisation schemes cited above. There is a common denominator, however, namely the recognition that educational systems may be centralised in some domains of decision-making but not in others.

**Description of the locus-of-decision-making instrument and data-collection procedure**

14. To learn more about educational decision-making in OECD countries and to systematically compare decision-making processes across countries, Network C developed an instrument that examined the locus of decision making in four important domains. As stated above, these domains were: (1) the organisation of instruction; (2) personnel management; (3) planning and structures; and (4) resource allocation and use. Within each of these four domains, between seven and 15 decisions were examined. In the domain entitled, “organisation of instruction,” for example, the instrument focused on decisions about such matters as textbook selection, grouping of pupils for instruction, and assessment of pupils’ regular work. In “personnel management,” questions were asked about hiring and dismissal of teachers and other school staff, duties and conditions of service, and the setting of salary schedules. In “planning and structures,” the focus was on creation and abolition of schools and grade levels, the design and selection of programs of study, course content, and policies regarding credentials. Finally, in the area of “resource allocation and use,” the instrument focused on decisions about the allocation of resources for staff and materials, and the use of financial resources for these purposes.
15. Each of the questions in the instrument was designed to identify the level at which decisions are made in the governmental system (the “level” of decision making) and the way decisions are made (the “mode” of decision making). Six “levels” of were decision-making were set out in the instrument. These include the following: (1) central government; (2) state governments; (3) provincial/regional authorities or governments; (4) sub-regional or inter-municipal authorities or governments; (5) local authorities or governments; and (6) schools. Three “modes” of decision-making were examined in the instrument. Decision could be made by an authority (1) autonomously, (2) within a framework established by another level within the system, or (3) in consultation with other levels in the system. Based on the instrument, it was possible to determine how centralised or decentralised a decision was overall, in each of the four domains, and for individual education decisions.

16. Finally, it should be noted that the instrument included questions about decision-making at three different education levels: primary education, lower secondary education, and upper secondary education. Within upper secondary education, questions were asked separately for general education and vocational education.

17. The decision-making survey was administered in the spring of 1998 to panels of national experts. For each level of education, countries assembled two 3-person panels with representatives from each of the following government levels: (1) highest level (central government); (2) middle levels (state governments, provincial/regional authorities or governments, sub-regional or inter-municipal authorities or governments, local governments); and (3) schools.

18. The two panels constituted for each education level went through the instrument question by question and attempted to arrive at consensus on the “level” and “mode” of decision-making on each question. The responses of each panel were then reviewed by each country’s Network C representative to determine whether there was consistency in the panels’ responses to each question. In cases where the responses differed, the Network C representative used source documents and consultation with the National Co-ordinator of the INES Project to reconcile these differences.

19. Following the administration of the questionnaires by each country, completed instruments were sent to the survey co-ordinator, who entered countries’ responses into a database. These responses were then used to calculate the indicators on decision-making, which were published in the 1998 edition of Education at a Glance. The data-set is sufficiently rich to calculate additional indicators. Examining the locus of decision-making with respect to domains and sub-domains is one of the most interesting possibilities. As an illustration the discussion that follows examines patterns of decision-making about teachers and other school staff that were developed from the decision-making study. Other relevant sub-domains that could be similarly analysed are, for example, decision-making about financial aspects, curriculum issues (see the 2000 version of Education at a Glance) and assessment and evaluation in education.

**Decision-making on teachers and other staff in lower secondary education as an example of a more domain specific use of the data set**

**Policy context**

20. Teachers, principals, and other school staff are critical components of the educational process. As the primary point in the transmission of knowledge to students and the development of skills, teachers can greatly affect students’ educational performance. Other professional staff, such as guidance counselors and librarians, provide valuable services to students that can support their instructional programs. And principals, in their role as instructional leaders and school managers, provide an environment that can enhance or hinder student learning.

21. As the major resource in the education process, it is important to understand how decisions about teachers and other school staff are made in different countries. At what government level are decisions
about teachers made? Are all decisions about teachers made at the same level, or does the decision-making process vary across areas? Is the decision-making process similar for teachers and other school personnel, or do decision-making practices differ for different types of school staff?

22. This analysis of decision-making about teachers and other school staff focuses on five important policy questions: (1) hiring of staff for specific positions; (2) dismissal of staff from a specific position; (3) duties and conditions of service; (4) setting of staff salary scales; and (5) career options of staff within the school system. For each question, we have examined the level in the system at which decisions are made (national, intermediate, local, and school) and the mode of decision making (whether decisions are made autonomously, after consultation with other levels in the education system, or within a framework set by a higher authority in the education system). The information reported is for decisions in lower secondary education.

Centralization and decentralization in decision-making about teachers

23. The 22 OECD countries considered here differ substantially in the locus of decision-making about teachers and other staff. At one end of the continuum are countries where decision making is highly centralised. In such countries as Greece, Portugal and Turkey, all five decisions about teachers are made by the central government. And in France, the majority of decisions are made centrally. In another group of countries, Austria, Korea, and Germany, most decisions are made by a central government or an intermediate-level government. However, in Korea and Germany decisions about the career paths of teachers are determined at the school level.

24. Belgium, Italy and Spain are also countries with relatively centralised decision-making on issues concerning teachers, but decision-making authority rests with intermediate governments. The comunidades autónomas in Spain have the authority to make decisions in all five areas, although all of these decisions about teachers are made within a framework set by the national government.

25. For the most part, countries that centralise decision-making about teachers at the national or intermediate levels do not involve the local or school levels in the decision-making process. Central governments either make decisions autonomously or in consultation with an intermediate level government; intermediate-level governments make decisions autonomously or within a framework set at the national level. Germany, however, is an exception, with intermediate governments making decisions about the hiring and dismissal of teachers and teachers’ conditions of service after consultation with the school level.

26. Moving towards the middle of the decision-making continuum, we find countries like Denmark, Ireland, the Netherlands, New Zealand and Norway, which divide authority for decisions about teachers between the central government and local governments or the schools. In Denmark and Norway, decisions about the hiring and dismissal of teachers are made at the local level, while in Ireland, the Netherlands and New Zealand, these decisions are made at the school level. With the exception of New Zealand, the other countries all set salary schedules for teachers at the central level. However, in New Zealand, the schools set salaries within a framework set at the central level.

27. More localised decision-making about teachers is evident in seven countries, but in these countries decision-making authority is divided between local governments and schools. In the United States, all decisions about teachers are made at the local level by separate educational authorities (school districts), although decisions about hiring and dismissal of teachers and their conditions of service are generally made in consultation with individual schools. Finland splits authority for decisions about teachers between the local and school levels, with local authorities responsible for hiring and dismissal of teachers and their career paths and schools responsible for setting salary schedules and conditions of service.

28. Four countries – the Czech Republic, England, Hungary, and Sweden – give schools the authority to make all five decisions about teachers, but relatively few decisions are made autonomously at
the school level. Many decisions are made within a framework set by the central government and a few
decisions are made in consultation with the local government. While schools have a great deal of
decision-making authority in these countries, their decisions still involve higher levels of government in
the decision-making process.

**Consistency in levels of decision-making**

29. Overall, there is great consistency across issues (questionnaire items) within countries in
decision-making about teachers. The three areas of decision-making that tend to cluster most closely in
terms of the level at which decisions are made are hiring, dismissal and conditions of service (duties and
responsibilities) of teachers. In countries where the central government makes decisions about the hiring
of teachers, (e.g., France, Greece, Portugal and Turkey), it also tends to make decisions about dismissal of
teachers and conditions of service. In countries where the school makes decisions about hiring, (e.g., the
Czech Republic, England, Hungary, the Netherlands, New Zealand and Sweden), it also tends to make
decisions in the other two areas.

30. Decision-making about salary scales for teachers also tends to cluster strongly with decision-
making about teachers’ conditions of service and moderately with decision-making about hiring and
dismissal of teachers. However, decisions about career paths for teachers are more often made at different
government levels than other decisions about teachers, particularly decisions about teacher salary scales.
In several countries, including Korea, Denmark, Ireland, Scotland, Germany and the Netherlands, a central
or intermediate government makes decisions about teacher salary scales, but the school is responsible for
decisions affecting the career paths of teachers. The opposite situation is found in New Zealand, where
the school makes decisions about teachers’ salary scales, but the central government determines teachers’
career paths.

**Decision-making about other school personnel**

31. Countries’ education systems differ dramatically. In some countries, the exclusive responsibility
of the education system is to provide instruction to students. The system does not provide health and
social support services, transportation or food services; these functions are provided by other government
authorities or by private agencies. In other countries, these functions are provided by the education system
and the staff that supports these functions is employed by this system.

32. Even when these functions fall within the domain of the education system, they may be financed
or provided by different levels of government. Central or intermediate governments may be responsible
for classroom instruction and for making decisions about teachers and other educational personnel, while
local governments may be responsible for maintaining school buildings and making decisions about
personnel in this functional area.

33. An examination of decision-making about different types of school personnel helps to clarify the
structure of countries’ education systems. The focus here is on two areas of decision-making: (1) hiring of
teachers and other school staff; and (2) setting of salary scales for teachers and other school staff.

**Hiring of teachers and other school personnel**

34. Countries differ in their decision-making about the hiring of teachers and other school personnel.
Countries such as Turkey are very centralised, with the central government responsible for the hiring of all
school personnel. France and Greece are also quite centralised, dividing authority for hiring school
personnel between the central and intermediate governments. In France, the central government makes
hiring decisions about both teachers and principals, while intermediate governments make this decision for
other school personnel. And in Greece, the central government makes decisions about hiring teachers,
while decisions about promoting principals and hiring other school personnel are made by intermediate
governments (Prefecture). Countries like Belgium, Italy, and Korea also centralise decision-making about the hiring of school staff, but intermediate, rather than central, governments make these decisions.

35. In the middle of the centralisation/decentralisation continuum are countries like Austria, Portugal, Spain, and the Czech Republic, which split decisions about the hiring of teachers and other school personnel between a central or intermediate government and local governments or the schools. In Austria, intermediate governments make hiring decisions about principals and teachers, while local governments hire other school staff. In Portugal, the central government makes hiring decisions about teachers and other school staff, while schools hire the principals. In the Czech Republic, intermediate governments hire school principals and schools hire the teachers and other school staff. And in Spain, the intermediate governments hire the teachers and other school staff and schools hire the principals.

36. The 11 remaining countries are much more decentralised in their decision making about hiring school personnel. In Finland, Norway and the United States, all decisions are made at the local level, although the three countries differ in their decision-making process. In Finland, local authorities make all of these hiring decisions autonomously, while in Norway they make decisions about principals and other staff autonomously and decisions about teachers within a framework set at the central level. In the United States school districts typically make decisions about hiring principals autonomously and decisions about teachers and other school staff after consultation with individual schools.

37. Denmark, Hungary, Scotland and Sweden divide decision-making authority for the hiring of teachers and other school personnel between the local and school levels. In all four countries, local authorities hire school principals, in most cases in consultation within individual schools. In all four countries except Denmark, schools hire the teachers autonomously; in Denmark they are hired at the local level, after consultation with the school. Decisions about the hiring of other school personnel are made at the school level in Denmark, Hungary, and Sweden, but at the local level in Scotland; the mode of decision making differs across the four countries.

38. Decisions about the hiring of all school personnel are made at the school level in England, Ireland, the Netherlands and New Zealand. In England and the Netherlands, all hiring decisions are completely autonomous. In Ireland, decisions to hire teachers and other school personnel are made autonomously, but decisions to hire principals are made within a framework established by the central government. In New Zealand, hiring decisions about teachers and principals are autonomous, but decisions about other staff also take place within a framework.

**Salaries of teachers and other staff**

39. Decision-making about salary scales for teachers and other staff shows most of the same breakdowns presented above for the hiring of school staff. However, this area of decision-making tends to be more centralised than decision-making about hiring staff. In nine countries – Austria, Denmark, France, Greece, Ireland, Italy, Korea, Portugal and Turkey – decisions about salary scales for all personnel are made by the central government. Except for Korea and Portugal, all of these salary decisions are made autonomously. In Germany, decisions are divided between central and intermediate governments. And in Belgium and Spain, decisions about salary scales for all personnel are made at the intermediate level.

40. Norway, Scotland and the Czech Republic fall in the middle of the centralisation/decentralisation continuum, dividing decision-making authority between either the central government and local governments (Norway and Scotland) or between an intermediate government and the schools (Czech Republic). The modes of decision making on salaries are varied in Norway and Scotland, while in the Czech Republic all decisions about staff salaries are made within a framework set by the central government.

41. Decentralised decision-making about salary scales is found in six countries. In the United States, decisions about salary schedules for all personnel are made at the local level by school districts; in England and New Zealand, all of these salary decisions are made at the school level. Decisions about
salary schedules are divided between the local and school levels in Finland, Hungary and Sweden. In all three countries, schools make decisions about teachers’ salaries and local authorities make decisions about principals’ salaries. However, in Finland, local authorities make salary decisions about other staff, but schools make these decisions in Hungary and Sweden.

42. As to decision-making modes, in the United States, all decisions at the local level about salaries are made autonomously. In England and Hungary, all salary decisions are made within a framework set by the central government, and this is the case in New Zealand as well regarding salary decisions for teachers and principals. And in Sweden, school decisions about salaries for teachers and other school staff are made after consultation with the local level.

43. In summary, decisions about staff salaries are made in the context of diverse governance structures. In about half the countries in the group, these decisions are centralised at the national level or at an intermediate level; in a small group they are split between central and local governments; and in about a third of the countries, these decisions are decentralised either at the local or the school levels.

<table>
<thead>
<tr>
<th>DECISION-MAKING ABOUT TEACHERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hiring of Teachers</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td><strong>ALL DECISIONS ARE MADE AT THE CENTRAL (NATIONAL) LEVEL</strong></td>
</tr>
<tr>
<td>Greece</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Portugal</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Turkey</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>MAJORITY OF DECISIONS ARE MADE AT THE CENTRAL (NATIONAL) LEVEL</strong></td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>MOST DECISIONS ARE SPLIT BETWEEN THE CENTRAL (NATIONAL) AND INTERMEDIATE (REGIONAL OR SUB-REGIONAL) LEVEL</strong></td>
</tr>
<tr>
<td>Austria</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Country</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Korea</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Germany</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Denmark</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Norway</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Ireland</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
</tr>
<tr>
<td>Country</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Belgium (Fr)</td>
</tr>
<tr>
<td>Spain</td>
</tr>
<tr>
<td>Italy</td>
</tr>
<tr>
<td>United States</td>
</tr>
<tr>
<td>Country</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>Finland</td>
</tr>
<tr>
<td>Scotland</td>
</tr>
<tr>
<td>Czech Rep.</td>
</tr>
<tr>
<td>England</td>
</tr>
<tr>
<td>Hungary</td>
</tr>
<tr>
<td>Sweden</td>
</tr>
</tbody>
</table>
**CORRELATIONS BETWEEN DECISIONS ABOUT TEACHERS**

<table>
<thead>
<tr>
<th></th>
<th>Hiring of Teachers</th>
<th>Dismissal of Teachers</th>
<th>Conditions of Service of Teachers</th>
<th>Setting of Salary Scales of Teachers</th>
<th>Career path of teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hiring of Teachers</td>
<td>+1.00</td>
<td>+.93</td>
<td>+.75</td>
<td>+.52</td>
<td>+.54</td>
</tr>
<tr>
<td>Dismissal of Teachers</td>
<td></td>
<td></td>
<td>+1.00</td>
<td>+.76</td>
<td>+.61</td>
</tr>
<tr>
<td>Conditions of Service of Teachers</td>
<td></td>
<td></td>
<td>+1.00</td>
<td>+.83</td>
<td>+.48</td>
</tr>
<tr>
<td>Setting of Salary Scales of Teachers</td>
<td></td>
<td></td>
<td></td>
<td>+1.00</td>
<td>+.29</td>
</tr>
<tr>
<td>Hiring of Principals</td>
<td>+.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hiring of Other Staff</td>
<td>+.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditions of Service of Principals</td>
<td></td>
<td></td>
<td></td>
<td>+.81</td>
<td></td>
</tr>
<tr>
<td>Setting of Salary Scales of Principals</td>
<td></td>
<td></td>
<td></td>
<td>+.91</td>
<td></td>
</tr>
<tr>
<td>Setting of Salary Scales of Other Staff</td>
<td></td>
<td></td>
<td></td>
<td>+.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hiring Teachers</td>
<td>Hiring Principals</td>
<td>Hiring Other Staff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DECISION-MAKING ABOUT HIRING TEACHERS AND OTHER STAFF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ALL DECISIONS ARE MADE AT THE CENTRAL (NATIONAL) LEVEL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous*</td>
<td>Autonomous*</td>
<td>Autonomous*</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DECISIONS ARE SPLIT BETWEEN THE CENTRAL (NATIONAL) AND INTERMEDIATE (REGIONAL AND SUB-REGIONAL) LEVELS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td><strong>Intermediate</strong> (Prefecture)</td>
<td><strong>Intermediate (Prefecture)</strong>*</td>
<td>Central</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*Promotion of Principals</td>
<td>Consultation with Intermediate Level (Prefecture)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Intermediate (Province)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Framework at Central Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DECISIONS ARE SPLIT BETWEEN THE CENTRAL (NATIONAL) AND SCHOOL LEVELS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>Central</td>
<td>School</td>
<td>Central</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Framework at Central Level</td>
<td>Autonomous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Framework at Regional Level</td>
<td>Framework at Sub-Regional Level</td>
<td>Framework at Central Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------</td>
<td>---------------------------------</td>
<td>-----------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium (Fr)</td>
<td>Intermediate (Sub-regional) Framework at State Level</td>
<td>Intermediate (Sub-regional) Framework at State Level</td>
<td>Intermediate (Sub-regional) Framework at State Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Intermediate (Province) Framework at Central Level</td>
<td>Intermediate (Province) Framework at Central Level</td>
<td>Intermediate (Province) Framework at Central Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Intermediate (Province) Consultation with School Level</td>
<td>Intermediate (State) Framework at Central Level</td>
<td>Intermediate (Provincial) and Local Consultation with School Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korea</td>
<td>Intermediate (Province) Autonomous</td>
<td>Intermediate (Province) Autonomous</td>
<td>Intermediate (Province) Autonomous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>Intermediate (State) Framework at Central Level</td>
<td>Intermediate (State) Consultation with School Level</td>
<td>Local Autonomous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>State Framework at Central Level</td>
<td>School Framework at Central and State Levels</td>
<td>State Framework at Central Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>School Framework at Central Level</td>
<td>Intermediate (Sub-regional) Framework at Central Level</td>
<td>School Consultation with Sub-regional Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Local Level</td>
<td>School Level</td>
<td>Other Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>Local</td>
<td>Local</td>
<td>Local</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>Local</td>
<td>Local</td>
<td>Local</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Framework at Central Level</td>
<td>Autonomous</td>
<td>Autonomous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>Local</td>
<td>Local</td>
<td>Local</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consultation with School Level</td>
<td>Autonomous</td>
<td>Consultation with School Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>Local</td>
<td>Local</td>
<td>School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consultation with School Level</td>
<td>Consultation with School Level</td>
<td>Framework at Local Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>School</td>
<td>Local</td>
<td>School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Consultation with School Level</td>
<td>Autonomous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scotland</td>
<td>School</td>
<td>Local</td>
<td>Local</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Other</td>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>School</td>
<td>Local</td>
<td>School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Consultation with School Level</td>
<td>Autonomous</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### DECISION-MAKING ABOUT SALARIES OF TEACHERS AND OTHER STAFF

<table>
<thead>
<tr>
<th>Country</th>
<th>Salaries of Teachers</th>
<th>Salaries of Principals</th>
<th>Salaries of Other Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
</tr>
<tr>
<td>Denmark</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>Other</td>
<td>Other</td>
</tr>
<tr>
<td>France</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
</tr>
<tr>
<td>Greece</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
</tr>
<tr>
<td>Ireland</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
</tr>
<tr>
<td>Italy</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous*</td>
<td>Autonomous*</td>
<td>Autonomous*</td>
</tr>
<tr>
<td>Korea</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Other*</td>
<td>Other</td>
<td>Other</td>
</tr>
<tr>
<td>Country</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>Other</td>
<td>Other</td>
</tr>
<tr>
<td>Portugal</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous*</td>
<td>Autonomous*</td>
<td>Autonomous*</td>
</tr>
<tr>
<td>Turkey</td>
<td>Central</td>
<td>Central</td>
<td>Central</td>
</tr>
<tr>
<td></td>
<td>Autonomous*</td>
<td>Autonomous*</td>
<td>Autonomous*</td>
</tr>
<tr>
<td>Germany</td>
<td>Central</td>
<td>Central</td>
<td>Intermediate (State)</td>
</tr>
<tr>
<td></td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Framework at Central Level</td>
</tr>
</tbody>
</table>

**DECISIONS ARE SPLIT BETWEEN THE CENTRAL (NATIONAL) AND INTERMEDIATE (REGIONAL AND SUB-REGIONAL) LEVELS**
### DECISIONS ARE SPLIT BETWEEN THE CENTRAL (NATIONAL) AND LOCAL LEVELS

<table>
<thead>
<tr>
<th>Country</th>
<th>Central</th>
<th>Central</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norway</td>
<td>Other</td>
<td>Other</td>
<td>Autonomous</td>
</tr>
<tr>
<td>Scotland</td>
<td>Central</td>
<td>Central</td>
<td>Local</td>
</tr>
<tr>
<td></td>
<td>Consultation with Local Level</td>
<td>Consultation with Local Level</td>
<td>Framework at Central Level</td>
</tr>
</tbody>
</table>

### ALL DECISIONS ARE MADE AT THE INTERMEDIATE (REGIONAL AND SUB-REGIONAL) LEVELS

<table>
<thead>
<tr>
<th>Country</th>
<th>Intermediate (State)</th>
<th>Intermediate (State)</th>
<th>Intermediate (Sub-regional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium (Fr)</td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
</tr>
<tr>
<td>Spain</td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
</tr>
</tbody>
</table>

### DECISIONS ARE SPLIT BETWEEN INTERMEDIATE (REGIONAL AND SUB-REGIONAL) AND SCHOOL LEVELS

<table>
<thead>
<tr>
<th>Country</th>
<th>School</th>
<th>Intermediate</th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Republic</td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
</tr>
</tbody>
</table>

### ALL DECISIONS ARE MADE AT THE LOCAL LEVEL

<table>
<thead>
<tr>
<th>Country</th>
<th>Local</th>
<th>Local</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>Autonomous</td>
<td>Autonomous</td>
<td>Autonomous</td>
</tr>
</tbody>
</table>

### DECISIONS ARE SPLIT BETWEEN THE LOCAL AND SCHOOL LEVELS

<table>
<thead>
<tr>
<th>Country</th>
<th>School</th>
<th>Local</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>Other</td>
<td>Other</td>
<td>Other</td>
</tr>
<tr>
<td>Hungary</td>
<td>School</td>
<td>Local</td>
<td>School</td>
</tr>
<tr>
<td></td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
</tr>
<tr>
<td>Sweden</td>
<td>School</td>
<td>Local</td>
<td>School</td>
</tr>
<tr>
<td></td>
<td>Consultation with Local Level</td>
<td>Other</td>
<td>Consultation with Local Level</td>
</tr>
<tr>
<td></td>
<td>School Framework at Central Level</td>
<td>School Framework at Central Level</td>
<td>School Framework at Local Level</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>England</td>
<td>School</td>
<td>School</td>
<td>School</td>
</tr>
<tr>
<td>New Zealand</td>
<td>School</td>
<td>School</td>
<td>School</td>
</tr>
<tr>
<td></td>
<td>Framework at Central Level</td>
<td>Framework at Central Level</td>
<td>Autonomous</td>
</tr>
</tbody>
</table>
Conclusion: future applications

44. The decision-making survey provided important information about different aspects of educational decision-making in OECD countries. Based on the survey, it was possible to calculate for each country the percentage of decisions made at different levels, as well as the percent of decisions made at different levels in each of the four educational domains. Through this analysis, it was possible to assess the relative centralisation/decentralisation of educational decision-making across countries and to determine the extent to which countries have placed decision-making authority at the school level. It was also possible to examine the relationship between the financing of education and decision-making in education across countries and to examine differences in decision-making within educational domains (e.g., decisions about teachers and school staff, as presented above). The use of the material, concerning decision-making on staffing presented in the previous section illustrated more detailed use of the data set for specific domains and sub-domains. However, the work on decision-making conducted to date still represents only a small part of the potential work that can be done in this area.

45. Future work on educational decision-making could be pursued in several areas. These are:

- **Develop more in-depth understanding of decision-making in selected areas.** As a follow-up to the questionnaire, Network C conducted a quick survey to learn more about the decision-making processes in the areas of teaching and curriculum. These surveys provide rich information about the participants in the decision-making process at different levels and the decision-making processes. This information could be used to develop short monographs that compare decision-making processes in greater depth.

- **Reanalyse survey results to focus on the school level.** Although the decision-making surveys administered in 1992 and 1998 did not contain the same “levels” of decision-making, the “school” remained a decision-making level in both instruments. Since there is strong interest in decision-making at the school level in many countries, in the light of policies to decentralise decision-making, the development of indicators of changes in decision-making at the school level would enhance understanding of the decentralisation process.

- **Examine decision-making in upper secondary education.** The last edition of *Education at a Glance* focused on decision making in lower secondary education, since there was a high level of correspondence in decision-making patterns at different education levels. However, areas of difference in decision-making between lower and upper secondary could be explored and differences between the two levels described in future indicators.

- **Link decision-making in upper secondary education with results of the upper secondary school survey.** Although the results of the upper secondary school survey will not be available for some time, the results of the survey could be linked with information about decision-making when they become available. These linkages will permit an exploration of the relationship between decision-making and key school policies related to the transition to work and higher education, human resources, and information technology. Analysis of these relationships could be developed into future indicators for *Education at a Glance*.

- **Link decision-making in upper secondary education with school process and outcome variables.** The Programme for International Student Assessment (PISA) will provide information about student achievement and selected school processes. When data from PISA become available, they could be linked with information from the decision-making survey to explore the relationship between decision-making and selected school processes and outcomes. Again, analysis of these relationships could be developed into future indicators for *Education at a Glance*.

- **Develop different methodologies for obtaining information about decision-making.** The panel approach used in the two administrations of the decision-making questionnaire relied on a
small group of experts to characterise decision-making about different aspects of education in their countries. This method provided valuable information, but it may not have captured different views about how decisions are actually made. Future work on decision-making could employ a survey to obtain more representative views on levels and modes of decision-making and explore differences in perceptions about the decision-making process by respondents at different levels of the education system. The panel approach could also be supplemented by case studies that focus more specifically on decision-making processes in different areas of education.

Conduct periodic surveys to examine changes in decision-making over time. Network C has administered the decision-making survey twice – first in 1992 and again in 1998. The second administration was intended, in part, to examine changes in decision-making across countries over time. However, changes in the “levels” of decision-making and a shift from “formal” decision-making, as embodied in laws and regulations, to “actual” decision-making, precluded analysis of changes in decision-making. Periodic administration of the same instrument at regular intervals (e.g., every 3-5 years), would address this problem and provide information on changes in decision-making processes within and across countries. Indicators could be produced that present changes in the percent of decisions made at different levels, the percent of decisions made in different domains, and changes in decision-making in specific areas (e.g., designing school programs).

46. In short, Network C has only begun the process of examining decision-making in education. The opportunities for more in-depth use of current information and the development of new information about decision-making abound. A programme of work in decision-making should therefore remain a central component of future work of the Network on school structures and processes. The revitalised interest in conditions that determine success in education in terms of attained outcomes, evident in the agenda for the 2000 General Assembly, may be an inspiration to develop a framework for assessing the effects of changing patterns in nations’ functional (de)centralisation.
SCHOOL TO WORK TRANSITION

Jittie Brandsma, University of Twente, The Netherlands

Introduction

1. The transition from school to work and the match between (vocational) education and the labour market have been persistent themes in the debates with regard to designing educational as well as labour (market) policies in recent decades. During the sixties and seventies the focus of the debate was on the expected changes in occupational structures and skill requirements, due to industrial changes and technological innovations in the (western) highly industrialised countries (cf. Lutz, 1994). In the early eighties this focus appeared to shift to the actual match between the qualifications acquired through vocational education and training and the qualification demands of enterprises. One of the impediments to this shift was the growing youth unemployment in various countries. The foreseen impact of technological innovation on necessary skills and qualifications, however, remained on the agenda as well.

2. Concerning the latter, the debate had a twofold character. On the one hand, there were various signals that vocational education and training was not keeping up with changing technology applications within business and industry and was delivering young people with obsolete skills. On the other hand, the issue of equipping young people with future oriented skills, providing them with the competence to adapt fairly easily to new situations and new technologies, also was an element in the debate. The combination of falling birth rates, expected leaps in technology and increasing qualification demands, led many to forecast serious skill shortages, which would threaten the competitiveness of western industrialised economies. This was certainly so if the newly emerging Asian economies were taken into account. Until now, such serious skill shortages seem not to have appeared, at least not on an overall level (which does not mean that there are no serious skill shortages in certain branches). Nevertheless, the awareness of the importance of a highly qualified and skilled workforce for a competitive economy as well as the importance of qualitatively good vocational education and training to ensure such a workforce, has grown during the last decades (McFarland & Vickers, 1994). This general recognition of the importance of a skilled workforce for economic competitiveness is reflected in the structural reforms of (upper) secondary education that have been or are undertaken in various EU-Member States and non-European OECD-countries since the mid-eighties (Brandsma, Kessler & Münch, 1995; McFarland & Vickers, 1994).

3. Processes of school to work transition are closely linked to the education and training system or structures. However, how these transitions actually take place -including particular problems that (can) occur- does not depend only on the education and training structures. The overall economic and labour market situation also has an influence on transition processes, as do the actions of and interaction between various actors involved (such as: employers, unions, public authorities, education and training establishments and young people themselves). At the same time, these also mould existing education and training systems (Pair, 1998). This contribution however, focusses on the supply side, that is: (vocational) education and training structures and upper secondary (vocational) education institutions.

Transition and transition problems

4. Whereas the degree of policy attention paid to young people's transition from initial education to working life has remained constant over the last three decades, the nature of the transition process as such has changed substantially. School to work transition appears no longer (or only exceptionally) to refer to
the point of concluding compulsory education and entering working life. Most, if not all OECD-countries, have faced a considerable extension of the number of years young people stay in (full-time) education. Countries also experienced increased levels of educational attainment. School to work transition therefore, is no longer the simple transfer from full-time education to stable full-time employment. It has developed towards a long and complex process with highly differentiated patterns of combinations and sequences of activities (education, training, work) and transitions between activities (Clasquin, Gérardin & Torese, 1998; Hannan & Werquin, 1999).

5. In this respect, school to work transition cannot be perceived as a single-time event, but needs to be conceptualised as a process during which study, work, job seeking, unemployment and other activities can be alternated and combined. In accordance with the definition provided in the OECD's transition review, school to work transition can thus be framed as "the period between the end of compulsory education and the point at which young people's principal activity is work" (OECD, 1998). Capturing transition in this sense means -to a certain extent- that the perspective for analysing factors influencing the (un)successfulness of transition processes as well as transition problems needs to be broadened. In this context, the changed meaning of the term 'drop-out' may be used as an illustration. Where formerly drop-outs were merely perceived as those students that left the (full-time) education system before reaching the end of compulsory schooling, the term is nowadays also applied to those young people that do not continue in upper secondary education or leave upper secondary education without having obtained an occupational or labour market relevant qualification (Brandsma, 2000; Hannan et al., 1998).

6. The problems young people encounter in their transition from school to work can have various backgrounds. As indicated earlier, the overall economic and labour market situation plays a role. In situations of high unemployment levels young people might encounter difficulties in obtaining stable employment due to the general lack of work. Although this will also depend on the particular unemployment patterns (in which sectors and at which levels) in relation to the qualification with which young people enter the labour market.

7. With regard to problems occurring in the process of school to work transition, a distinction can be made between mismatches in terms of 'quantitative problems' and 'qualitative problems'. Quantitative mismatches refer to the discrepancies in the volume of supply and demand for particular qualifications. Qualitative mismatches refer to the discrepancies between the content and level of the qualifications acquired through (vocational) education and the content and level of the qualifications required for the performance within an occupation or job (Brandsma, 1993; 1999). Such discrepancies can manifest themselves on the level of the external labour market (e.g.: over-schooling) as well as the job level (e.g.: waste of qualifications), which refers to the distinction between ‘overt’ and ‘hidden’ discrepancies (De Grip, 1987).

8. Trying to grasp the different types of qualitative mismatches is important; not only the distinction between the location where these mismatches appear, but also the points of view from which the discrepancies are defined. It will be clear that employers’ perceptions of mismatches will differ from what employees perceive as being a mismatch. Figure 1 provides a typology of qualitative mismatches between vocational education and the labour market, in which the two dimensions, location and perspective, are taken into account.
9. From the point of view of the employers, qualitative mismatches consist in staffing problems and qualification problems. Distribution problems and utilisation problems can be perceived as their opposite, as being the mismatches from the perspective of the employees.

10. **Staffing problems** refer to the problems which employers can encounter in filling specific vacancies. This kind of qualitative discrepancy appears to be a typical cyclical problem. Normally, during economic recession staffing problems are rather small because labour supply actually exceeds labour demand in most economic sectors. However, currently staffing problems appear to persist (at least in some countries) especially for jobs in high-tech areas with specific and high qualification demands, but also in various other occupational sectors.

11. **Distribution problems** concern the problems that job-seekers encounter in finding a job that fits their education and training. Available jobs in terms of their job content, payment and status, are distributed unequally among different groups of potential employees. This inequality in distribution does not stem only from the differences in perspective between the various education and training programmes, but is reinforced, also by recruitment strategies and criteria which are used by employers (Glebbeek, 1993; Meesters, 1992; Van Hoof, 1987).

12. **Qualification problems** concern the mismatch between the qualifications an employee has and the qualifications that are required for the performance of a specific job. Complaints of employers concerning the obsolescence of (vocational) education and training programmes and the incapacity of (vocational) education and training systems to adapt quickly to changing working methods and production processes, have often been phrased in this context. At the same time it appears that young people that do not have any relevant occupational or labour market qualifications, encounter more difficulties in this area. That does not necessarily concern only the least qualified. It also afflicts the relatively better educated that only attended general education without transferring into (or completing) a tertiary education (Hannan et al., 1998).

13. **Utilisation problems** are the direct opposite of the qualification problems. Utilisation problems refer to discrepancies between the employee’s qualification and the actual use of these qualifications within the labour organisation. In this respect both an “over demand” and a “waste” is possible. In the former case, the job demands more/higher qualifications than the worker has, while in the latter the reverse is the case. Given the ongoing educational expansion a waste of qualifications appears to be more likely, although there also is evidence that young people start in jobs that are actually somewhat below
their qualification level, making up for this during the first years of working (Hannan et al., 1998; Hannan & Werquin, 1999; Hartog, 1999).

14. To what extent and how do these kind of qualitative mismatches affect upper secondary (vocational) education? Nieuwenhuis (1991) distinguishes three levels of discontinuity in the relation between education and labour from an educational point of view:

- At the structural-distributive level: the discontinuity represents the mismatches that become manifest in the labour market. This discontinuity refers to the question whether it is possible for vocational education to realise such programmes and specialisations that the imperfections in the match between various educational programmes on the one hand, and various jobs on the other hand, can be solved;

- At the qualification level: the discontinuity refers to the issue of qualifications to be taught in different vocational training programmes. More specifically this discontinuity refers to the crucial issue of adapting vocational curricula to changes in occupations. How should this be done and which changes should or should not be taken into account?

- At the didactical or instructional level: the discontinuity is related to the dilemma of best learning environments. School-based learning is often perceived as inadequate for teaching vocational skills and attitudes. Work-based learning, however, is not always a good alternative. Production as such gets the priority in labour organisations, and the extent to which jobs provide opportunities for (informal) learning will depend on the way in which production is organised (Münch, 1985; Nieuwenhuis, 1991).

15. Structural-distributive discontinuity encompasses both the quantitative and qualitative discrepancies in the labour market. Solving these discrepancies, in particular the quantitative ones, is a difficult task. Educational policy cannot create employment. At the same time, attempts to influence participation rates in particular programmes at particular levels have the tendency to lag behind the developments in the employment system. Therefore more pro-active strategies are needed, which however, have to deal with uncertainties and the limitations of forecasts. Though discrepancies at the qualification and the didactical or instructional level might be easier addressed by education and training policies, it should be taken into account that discontinuity at the structural-distributive level does influence the discontinuity at the other two levels. Attempts to solve discontinuity problems at the qualification level, should take the restrictions that are set at the structural-distributive level.

Developments in upper secondary (vocational) education and training systems

16. In various OECD-countries systems of (upper) secondary education have undergone considerable reform over the past 10 years. Though problems in the school to work transition have not been the only reason for these reforms, they have been part of the considerations underpinning changes like the extension of compulsory schooling, the elaboration of access routes into higher education or the establishment of new or thoroughly innovated (vocational) programmes (cf. Raffe, 1998). In this sense there is a certain convergence between upper secondary education systems (Pair, 1998). At the same time, the differences between OECD-countries concerning the characteristics and differentiation of the upper secondary education systems are considerable (see also tables 1 and 2) and grounded in systemic differences with regard to (Hannan et al., 1998; OECD, 1998):

- the extent of development of vocational education and training as a separate (sub)system with its own historical development;

- the horizontal and vertical differentiation within (upper) secondary education systems;

- the access to and different pathways into higher education;
– the extent of articulation of (formal) links between upper secondary education and the labour market/employment system, including, for instance, the involvement of employers in vocational education and training.

17. Table 1 provides insight in the student numbers in upper secondary education (absolute and as a percentage of the total ISCED 3 population with the percentages between brackets\(^\text{69}\)).

### Table 1: Student numbers in general, pre-vocational and vocational oriented upper secondary education programmes

<table>
<thead>
<tr>
<th>Country</th>
<th>General</th>
<th>Pre-vocational</th>
<th>Vocational</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>77,000</td>
<td>28,537</td>
<td>273,150</td>
<td>378,687</td>
</tr>
<tr>
<td>Belgium</td>
<td>116,613</td>
<td>80,102</td>
<td>274,284</td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>126,355</td>
<td>-</td>
<td>378,687</td>
<td>464,186</td>
</tr>
<tr>
<td>Denmark</td>
<td>107,429</td>
<td>144,592</td>
<td>257,641</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>126,000</td>
<td>62,888</td>
<td>188,888</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>1,484,132</td>
<td>533,844</td>
<td>2,117,976</td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>111,672</td>
<td>133,622</td>
<td>426,510</td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>126,374</td>
<td>-</td>
<td>153,929</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>918,048</td>
<td>1,829,514</td>
<td>2,747,562</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>1,730,659</td>
<td>410,154</td>
<td>2,140,813</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>194,500</td>
<td>100,376</td>
<td>294,876</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>93,617</td>
<td>66,808</td>
<td>160,425</td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>243,543</td>
<td>97,343</td>
<td>340,886</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>930,957</td>
<td>210,017</td>
<td>1,140,974</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>139,489</td>
<td>154,131</td>
<td>293,620</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>49,240</td>
<td>194,000</td>
<td>243,924</td>
<td></td>
</tr>
<tr>
<td>United Kingdom (NYA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>8,800,000</td>
<td>-</td>
<td>-</td>
<td>8,800,000</td>
</tr>
</tbody>
</table>

\(^{69}\): The figures are based on the work undertaken in the classification study upper secondary education, which was one of the preparatory stages for the upper secondary school survey. Student numbers and calculations are based on the definition of the target population for the survey, as developed within the framework in the classification study.
18. The relative share of the vocational oriented programmes varies considerably between countries. Here, it should be taken into account that in one case (USA) all upper secondary programmes are classified as ISCED 3A without the possibility to further breakdown student numbers according to particular programme orientation, while in other countries (Ireland) programmes that prepare young people for the labour market (that is: providing a full labour market qualification) are mainly situated at ISCED level 4. Taking this into account, it is clear that the relative share of vocational education and training in upper secondary education ranges from just under 20% (Spain, Mexico) up to well over two-third of the ISCED 3 population, especially in countries where traditionally apprenticeship systems have been a major route for young people to continue their studies after compulsory schooling.

19. All countries have gone through a general raising of educational attainment levels through extending the duration of schooling. Even though the proportions of young people affected differ between countries (Murray & Steedman, 1998; OECD, 1998). The extension of the duration of schooling is, however, often accounted for by the increasing participation rates in general upper secondary education. Pair (1998) concludes that only few countries have seen an actual and substantial growth in participation rates in upper secondary vocational education (cf. OECD, 1998). Moreover, he concludes that some of the countries with a traditionally strongly developed apprenticeship system have seen stagnating or even falling participation rates for the vocational oriented programmes. In particular, these slightly falling rates appear to be accounted for by declining participation rates in apprenticeship schemes, which could be explained by the development of school-based alternatives providing access to higher education (Lassnig, 1998).

20. Table 2 lists the various vocational programmes provided in the OECD-countries included in the classification study (across ISCED 3A, 3B and 3C). This listing indicates the destination for which the programmes prepare (or for which they provide access).
Table 2: Upper secondary education programmes with a vocational orientation according to destination

<table>
<thead>
<tr>
<th>Programmes</th>
<th>Destination</th>
<th>Student numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Austria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• secondary technical and vocational colleges</td>
<td>labour market/ ISCED 4C</td>
<td>100,000</td>
</tr>
<tr>
<td>• secondary academic schools with vocational training</td>
<td>ISCED 4C/ ISCED 5B</td>
<td>150</td>
</tr>
<tr>
<td>• apprenticeship training</td>
<td>labour market</td>
<td>125,000</td>
</tr>
<tr>
<td>• secondary technical and vocational schools</td>
<td>labour market</td>
<td>33,000</td>
</tr>
<tr>
<td>• vocational schools for agriculture and forestry</td>
<td>labour market</td>
<td>9,000</td>
</tr>
<tr>
<td>• training of auxiliary nurses</td>
<td>labour market</td>
<td>2,000</td>
</tr>
<tr>
<td>• training for sport instructors</td>
<td>labour market</td>
<td>4,000</td>
</tr>
<tr>
<td><strong>Belgium – Flemish Community</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• technical upper secondary (2nd stage and 1st and 2nd year 3rd stage) (TSO)</td>
<td>ISCED 5A/5B or labour market</td>
<td>89,180</td>
</tr>
<tr>
<td>• artistic upper secondary (2nd stage and 1st and 2nd year 3rd stage) (KSO)</td>
<td>ISCED 5A/5B or labour market</td>
<td>4,770</td>
</tr>
<tr>
<td>• vocational upper secondary education (2nd stage and 1st and 2nd year 3rd stage) (BSO)</td>
<td>labour market</td>
<td>72,432</td>
</tr>
<tr>
<td><strong>Belgium – French Community</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• technical secondary (qualification stream)</td>
<td>labour market</td>
<td>43,720</td>
</tr>
<tr>
<td>• technical type II</td>
<td>labour market/ ISCED 5B</td>
<td>697</td>
</tr>
<tr>
<td>• vocational upper secondary education</td>
<td>labour market</td>
<td>35,685</td>
</tr>
<tr>
<td>• dual vocational education and training</td>
<td>labour market</td>
<td>?</td>
</tr>
<tr>
<td>• apprenticeship training</td>
<td>labour market</td>
<td>?</td>
</tr>
<tr>
<td>Programmes</td>
<td>Destination</td>
<td>Student numbers</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>Czech republic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• secondary technical schools</td>
<td>ISCED 5A</td>
<td>180,259</td>
</tr>
<tr>
<td>• secondary vocational school programmes with maturita exam</td>
<td>ISCED 5A</td>
<td>25,084</td>
</tr>
<tr>
<td>• technical lyceum</td>
<td>ISCED 5A</td>
<td>?</td>
</tr>
<tr>
<td>• economic lyceum</td>
<td>ISCED 5A</td>
<td>?</td>
</tr>
<tr>
<td>• secondary technical without maturita exam</td>
<td>labour market</td>
<td>6,811</td>
</tr>
<tr>
<td>• secondary vocational without maturita exam(^2)</td>
<td>labour market</td>
<td>125,677</td>
</tr>
<tr>
<td><strong>Denmark</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• upper secondary vocational</td>
<td>labour market</td>
<td>129,592</td>
</tr>
<tr>
<td>• agricultural vocational</td>
<td>labour market</td>
<td>2,000</td>
</tr>
<tr>
<td>• maritime vocational</td>
<td>labour market</td>
<td>1,000</td>
</tr>
<tr>
<td>• social and health training</td>
<td>labour market</td>
<td>12,000</td>
</tr>
<tr>
<td><strong>Finland</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• three year vocational programme</td>
<td>ISCED 5A/5B</td>
<td>62,888</td>
</tr>
<tr>
<td><strong>France</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Bac Pro</td>
<td>ISCED 5B labour</td>
<td>160,247</td>
</tr>
<tr>
<td>• Brevet de Technicien</td>
<td>ISCED 5B/ labour</td>
<td>20,000</td>
</tr>
<tr>
<td>• BEP</td>
<td>ISCED 5B/ labour</td>
<td>463,152</td>
</tr>
<tr>
<td>• CAP</td>
<td>labour market/ ISCED 3B</td>
<td>70,692</td>
</tr>
<tr>
<td>• vocational training for NVQL certificate/programme</td>
<td>labour market/ other ISCED 3 programmes</td>
<td></td>
</tr>
<tr>
<td><strong>Italy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• technical institutes (1(^{st}) and 2(^{nd}) cycle)</td>
<td>ISCED 5A/5B / labour market</td>
<td>1,079,431</td>
</tr>
<tr>
<td>• artistic lyceum (+ integration course)</td>
<td>ISCED 5A/5B</td>
<td>31,000</td>
</tr>
<tr>
<td>• vocational and technical experimental courses</td>
<td>ISCED 5A/5B / labour market</td>
<td>?</td>
</tr>
<tr>
<td>• vocational and art institutes (1(^{st}) + 2(^{nd}) cycle)</td>
<td>ISCED 5A/5B / labour market</td>
<td>216,098</td>
</tr>
<tr>
<td></td>
<td>labour market/ ISCED</td>
<td></td>
</tr>
<tr>
<td>Programmes</td>
<td>Destination</td>
<td>Student numbers</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>• instituto magistrale (+ integration course)</td>
<td>5A</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• technical professional education</td>
<td>labour market</td>
<td>410,154</td>
</tr>
<tr>
<td>Netherlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• senior vocational education: middle management training</td>
<td>ISCED 5A</td>
<td>68,938</td>
</tr>
<tr>
<td>• basic vocational training</td>
<td>labour market/ other ISCED 3C programmes</td>
<td>18,318</td>
</tr>
<tr>
<td>• vocational training</td>
<td>labour market/ ISCED 3A vocational</td>
<td>13,120</td>
</tr>
<tr>
<td>Norway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• vocational upper secondary</td>
<td>labour market/ ISCED 3A</td>
<td>66,808</td>
</tr>
<tr>
<td>Portugal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• technological courses</td>
<td>ISCED 5B/ labour market</td>
<td>70,566</td>
</tr>
<tr>
<td>• school-based vocational training</td>
<td>labour market</td>
<td>26,777</td>
</tr>
<tr>
<td>• apprenticeship training</td>
<td>labour market</td>
<td>NA</td>
</tr>
<tr>
<td>• Casa Pia vocational training</td>
<td>labour market</td>
<td>NA</td>
</tr>
<tr>
<td>Spain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• vocational education intermediate level</td>
<td>labour market/ ISCED 3A</td>
<td>186,384</td>
</tr>
<tr>
<td>• vocational education for the plastic arts/applied arts and artistic crafts</td>
<td>labour market/ ISCED 4B</td>
<td>14,633</td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• vocational upper secondary</td>
<td>ISCED 5A</td>
<td>154,131</td>
</tr>
<tr>
<td>Switzerland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• vocational education</td>
<td>ISCED 5B/ labour market</td>
<td>170,800</td>
</tr>
<tr>
<td>• basic vocational education</td>
<td>labour market</td>
<td>3,200</td>
</tr>
<tr>
<td>• vocational education (1-2 years)</td>
<td>labour market</td>
<td>20,000</td>
</tr>
</tbody>
</table>
## Programmes

<table>
<thead>
<tr>
<th>Programmes</th>
<th>Destination</th>
<th>Student numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• NVQ/SVQ level 3</td>
<td>ISCED 5A/5B or labour market</td>
<td>NYA</td>
</tr>
<tr>
<td>• modern apprenticeship</td>
<td>ISCED 5B</td>
<td>NYA</td>
</tr>
<tr>
<td>• GNVQ/GSVQ advanced</td>
<td>ISCED 5A/5B or labour market</td>
<td>NYA</td>
</tr>
<tr>
<td>• GNVQ/GSVQ intermediate level</td>
<td>ISCED 3A or labour market</td>
<td>NA</td>
</tr>
<tr>
<td>• NVQ/SVQ level 1 and 2</td>
<td>labour market/ ISCED 3A</td>
<td>NA</td>
</tr>
<tr>
<td>• National Traineeship</td>
<td>labour market</td>
<td>NA</td>
</tr>
<tr>
<td>• traditional traineeship</td>
<td>labour market</td>
<td>NA</td>
</tr>
</tbody>
</table>

### Pathways

21. As becomes apparent, the extent to which vocational oriented programmes provide access to tertiary education differs. In a number of countries vocational pathways appear to cut off access to tertiary education. However, the extent to which this provides ‘dead ends’ for young people will also depend on the actual routes or pathways they take through the education system. There is, for instance, the example of young people in Germany first taking up a general upper secondary education programme providing a qualification that gives access to tertiary education, before entering a vocational training within the apprenticeship system. At the same time there are attempts to increase the possibilities for accessing higher education following apprenticeship training. It is often argued that in attempts to build up the parity of esteem between general and vocational education, it is important that the attractiveness of vocational education is increased by providing access to tertiary education programmes as well. Reforms undertaken in various countries over the last years, often have had this particular intention. The reforms of the upper secondary education system in the Nordic countries (in particular Sweden, Norway and Finland) are an example of this. The reform of the education system presently implemented in Spain provides a similar example. This reform gives young people enrolled in intermediate vocational training programmes the choice to pursue studies in more specialised advanced vocational programmes or to continue in programmes preparing for the upper secondary general diploma, providing access to university education.

22. Raffe (1998) defines 'pathways' in terms of programmes, which may be linked together in sequences. He indicates that educational policy makers can try to establish such links with the intention of building pathways, but that these do not necessarily correspond the pathways young people actually take. The 'real' pathways young people take, are sometimes considered inefficient by policy makers. Nevertheless, as Raffe indicates, such diverting pathways might be attractive for individual students given the opportunities, incentives, costs and constraints associated to these pathways in the perception of students. In relation to the further development of pathways, Pair (1998) concludes that there is a development towards upper secondary education systems with birdges between the different (general and vocational) programmes. Even though there are still substantial differences between countries the consequence of this development is that students can enter and leave the system at different points.

23. Young people most confronted with transition problems are those that do not continue in upper secondary education after the end of compulsory schooling. The same applies for those that fail to obtain an upper secondary qualification that provide them either with access to tertiary education or a good
starting position in the labour market. The labour market situation of young people with only lower secondary education or young people that did not obtain a vocational or professional qualification is relatively weak and often characterised by employment in less stable jobs with little career perspectives alternated with spells of unemployment. In various OECD countries programmes have been established or developed that attempt to cater for those young people that for various reasons do not (wish to) continue in upper secondary education or run the risk of dropping out.

24. Such programmes could be characterised as ‘safety nets’ that attempt to cater for those groups of young people, either to retain them in the education system or to reinsert them in the education system. Apart from specific labour market (training) schemes, this also concerns education and training programmes within the (upper) secondary education system that have been more or less specifically designed for this purpose. Concerning the latter, a distinction could be made between programmes postponing the choice for a particular pathway and programmes catering for students that would otherwise have left the educational system. Although the differences between these two are more gradual than fundamental. Examples of programmes postponing the choice for a particular pathway in upper secondary education and/or better preparing students for the continuation of the studies are the pre-vocational programmes in Austria and the transition year programme in Ireland. These programmes can be perceived as a bridge between lower and upper secondary education, providing the possibility of concluding lower secondary education without already opting for a particular track in upper secondary education (Austria), or providing the opportunity to utilise the full three years allowed in upper secondary education and become more mature before choosing a particular option (Ireland). Examples of programmes catering for students that otherwise might have dropped out of the educational system, are the vocational basic training and individual youth training programmes in Denmark, the individual upper secondary education programmes in Sweden, the social guarantee programme in Spain and the assistant training programme in the Netherlands (which is, however, classified as an ISCED 2 programme). These programmes attempt to cater for young people that, due to ability or aptitude, might encounter difficulties to proceed in upper secondary education. Overall, the programmes appear to fulfil a double function providing basic (vocational) training which results in basic skills for employment, as well as preparing students for a continuation of studies in other (vocational) ISCED 3 programmes.

Vocational education and training institutions; facilitating school to work transitions

25. The OECD’s thematic review of the transition from initial education to working life outlines a number of features that appear to contribute to successful transitions. These features are (OECD, 1998):

- clearly defined, well organised, open and coherent learning pathways and qualification frameworks;
- availability of opportunities for young people to participate and learn in real work settings while being a student;
- provision of a broad range of vocational and technical skills together with general education and personal skills, in particular for young people not continuing into higher education;
- existence of ‘youth friendly’ labour markets;
- tightly knit safety nets for young people most at risk of social and economic exclusion (with the aim of reintegrating those young people into mainstream education and training);
- attractive and accessible information, guidance and follow-up services for all young people;
- institutional frameworks that allow for organised and continuous involvement and co-operation of all relevant national, sectoral and local actors, in order to achieve policy coherence and effective programme implementation;
- well designed monitoring tools such as statistics, indicators and longitudinal surveys that reflect both developments in and interaction between education and employment systems.
26. These features appear to be mainly located at the system level, though their implementation will affect educational institutions as well. The question, however, is to what extent upper secondary educational institutions can and do contribute to facilitating the process of school to work transition. This issue can be highlighted from different perspectives.

**School effectiveness research in vocational education and training**

27. A first perspective can be derived from school effectiveness research. It is known that schools do matter where the achievements and educational success of students is concerned. However, school effectiveness research has until now mainly covered primary and lower secondary education. There are some indications concerning differences between vocational upper secondary education institutions with regard to their effectiveness. In the Netherlands for instance, it is clear that overall drop-out rates from upper secondary vocational education are relatively high on average. It is also known that differences between educational establishments and between different units within educational establishments are substantial (Inspectorate, 1999). School effectiveness research in this educational sector appears to be little developed until now. There are some examples of American studies (Raizen, 1989; Wardlow et al., 1992), but these are mainly qualitative studies focussing at exemplary or 'excellent' schools. However, these studies incorporate a substantial number of school process variables which inhibit the testing of hypotheses through replication in a more quantitative approach.

28. Studies in the Netherlands in which it has been attempted to adopt the school effectiveness approach in upper secondary vocational education, show that it is difficult to identify process variables that can explain the differences between educational establishments. On the one hand, differences in educational effectiveness appear to be mainly explained by student characteristics like previous educational attainment, age, ethnicity and -to a lesser extent- gender. On the other hand, empirical evidence of the influence of those school process characteristics which are thought to matter, appears to be small if not insignificant. The results of different studies are often not consistent. This is partly due to differences in measurement and operationalisation of process variables and to differences in the school process variables included in the studies. A more or less consistent result of these studies concerns the 'sector' effect. That is: the particular occupation sector for which programmes and institutions train (broadly distinguished as 'economic occupations', 'technical occupations' and 'health and social care occupations') does make a difference for the effectiveness of institutions and programmes. Various explanations have been given for this phenomenon. Among which: motivational differences between students, differential effects of the push and pull mechanisms on the labour market and the degree of real or occupational specific skills provided for in the programmes. Until now it has proven difficult to fully disentangle these varying explanations (cf. Van Batenburg, 1995; Brandsma, 1999; De Bruijn, 1997).

29. In these Dutch studies the effectiveness criterion was whether or not students obtained the final diploma or qualification. A distinction that is often made in the context of effectiveness of vocational education and training concerns the distinction between 'output' and 'outcomes' (or internal versus external success rates). Output then refers to the issue of whether students conclude the programme and obtain the diploma, whereas 'outcomes' refer to the issue whether students obtain gainful employment or continue studies in higher education (depending on the orientation of the programme) (cf. Brandsma et al., 1999). These are predominantly quantitative success indicators and one could argue that it matters as much what students actually learn during their participation in upper secondary (vocational) education. This more qualitative output indicator, however, brings up specific measurement problems that are not easily resolved, given the differentiation in various vocational specialisations with their different specialised subjects and goals. Concerning outcomes some argue that the question not only whether students obtain gainful employment, but also whether such employment matches both the direction and level of the education or training programmes taken. Apart from the question at which moment after transition to the labour market this 'match' should be measured, it also raises the issue of the meaning it might have if students end up outside the occupational area for which they have been trained. Concerning the former it has already been indicated that young people often start at a job level actually below their level of training, but catch up during the early years of their working career. Concerning the latter, it is known that
graduates of particular specialties end up outside the main occupational field for which they have been trained (early or later in their working career). Instead of perceiving this as a mismatch, it could also be concluded that particular pathways equip students with knowledge and skills, including adaptability and flexibility, which equip them for a range of opportunities in the labour market (cf. Heijke, 2000). Moreover, it should be acknowledged that educational institutions might be able to influence 'output', but that 'outcome' is much less easy to be influenced by these institutions. Outcomes will also - probably even more - be influenced by the labour market and the employment system.

30. The 'effectiveness criteria' as mentioned in the previous paragraph are of a rather quantitative nature. Some argue for alternative effectiveness indicators like job satisfaction of teaching staff, cohesion within educational institutions, capacity of raising (additional) resources (cf. Quinn & Rhorbaugh, 1983) or quick reinsertion of drop-outs and job satisfaction of graduates. As such these alternative effectiveness indicators might be interesting. However, with regard to these criteria for the organisation as being an effective organisation, the question could be raised whether it might not be more important to investigate how these 'intermediate' criteria finally contribute to the effectiveness of the primary process. Concerning the reinsertion of drop-outs and job satisfaction of graduates, primarily the first appears to be relevant, assuming that the latter can hardly be influenced by the educational institutions. Quick reinsertion would require good monitoring of students' progress and early signalling of problems. What schools should primarily be accountable for is whether or not they can help students to obtain a meaningful qualification. This as such is an important contribution towards facilitating the transition process.

Alternative perspectives

31. Even though knowledge with regard to the extent to which schools can contribute to facilitating the transition process and the way in which they can is still not very well developed, there are particular themes that attract continuing interest of both policy makers and researchers.

32. Practical training and work-based learning. There appears to be an increasing emphasis on the role of practical training and especially work-based learning in realistic working environments. Both from the perspective of facilitating the transitions process and from the perspective of acquiring specific work-related knowledge and skills. Even in this area there appears to be a certain controversy with regard to the effects of such work-based periods of learning (often stemming from differences in goals and effects to be measured). Raizen (1994) indicates that contextualised learning following inductive approaches might not only be more effective, but also more motivating for students. Providing that certain conditions are met (Onstenk, 1997), the working place can offer such a learning environment.

33. What became apparent from the classification study, carried out within the context of OECD/INES Network C as a pre-study in preparing a school survey at the upper secondary level, is that:

- practical training periods within labour organisations are mainly restricted to vocational oriented programmes (that is: those programmes that predominantly prepare for the labour market or for a 'double qualification' giving access to tertiary education and providing a full labour market qualification);
- the form in which such practical training takes place (actual duration as percentage of total curricular time, location within the curriculum) differs substantially between programmes and countries.

34. Concerning the latter it has also to be mentioned that the length of practical training periods was often indicated as being variable, differing between particular specialisation or between educational institutions. This occurs in particular if decision making about the actual implementation of curricula and instructional design is devolved to the schools (as is the case in the Netherlands).

35. Monitoring and counselling. As indicated in the context of quick reinsertion and preventing drop-out, monitoring and assessing students' progress and students' particular problems is important. If
particular problems (difficulties with the particular pathway chosen, motivation) are signalled in an early stage it will be easier to develop alternatives or find solutions. Monitoring or actually ‘tracking’ students is also important from the perspective of knowing what happens to students once they do drop-out. In the Netherlands, for instance, drop-out from the vocational training programmes appears to be relatively high, but if one investigates what happens to students it becomes clear that a substantial proportion of these drop-outs return to another training programme, often even within the same regional training centre. Due to inadequate monitoring systems, the establishments themselves are often not aware of this. Counselling is important not only in the context of providing alternatives for students facing difficulties. As Raffe (1998) indicates, the increasing number of alternative pathways and the complexity of pathways, requires that guidance and counselling is available for students to make a (rational) choice between different pathways. Early drop-out from vocational programmes for example, often appears to be caused by the incongruence between the perception young people have of the particular occupation or work area and reality. Guidance in choosing a particular programme could at least reduce such incongruences.

Conclusion: Prospective indicators from the upper secondary school survey

36. A school survey within upper secondary education is presently planned to be administered in 2001. As preparation for this survey two activities have been undertaken:

- a classification study of upper secondary school programmes and providers that should be the basis for the development of a sampling frame;
- the design of a questionnaire for written administration addressing school policies and practices to enhance transition, school functioning, human resources, information and communication technology and some administrative issues.

37. Given the relatively scarce and fragmented information available thus far on the issue of whether and how educational institutions can contribute to the facilitation of the school to work transition process, collecting information directly from upper secondary education institutions can provide better understanding of what schools actually attempt to do in this area. Until now, Education at a Glance mainly contained indicators dealing with the labour market outcomes of school to work transitions. The prospective indicators to be derived from the upper secondary school survey could increase the insight into the process; at least from the perspective of the supply side.
REFERENCES


CEREQ.


SECTION IV

CONTRIBUTION FROM THE TECHNICAL GROUP
INTRODUCTION

722. About 65 representatives from all OECD Member countries - statisticians, researchers and policy analysts - plus observers from various non-Member countries co-operate in the INES Technical Group. The majority of participants come from Ministries of Education but many countries are also represented through National Statistical Offices and some by research institutes or the Ministry of Finance. This reflects the interdisciplinary nature of the work of the Technical Group.

723. The primary focus of the work of the Technical Group is on costs and human resources in education; access to and participation in education and training; student stocks and flows; and the graduate output of educational institutions. In these domains the Technical Group has made a significant contribution both with respect to the establishment of a broad knowledge base at the OECD and with respect to the development of indicators, the methodology of their calculation and their placement in the policy context of national education systems.

724. Since Lahti, the Technical Group has become the main provider of education statistics and indicators for the INES project and, beyond this, for a range of related activities at the OECD. The development of a coherent data collection framework and the centralisation of the data management and indicator calculation at the Secretariat have provided an opportunity for the role of the Technical Group to evolve from one of supplying data into one characterised by an active involvement in data and indicator development that is of growing importance for the ongoing analytical work in INES.

725. Over the last years, the INES Technical Group focused much of its attention on the tertiary level of education. The overall policy context for this work was provided by OECD’s thematic review of the first years of tertiary education (OECD, 1998), which illuminated major shifts in policy and structural reforms in tertiary education in OECD Member countries. Efforts in the Technical Group were directed to:

- The further development of indicators on student access, choice, modes of participation and survival at the tertiary level of education, under the new paradigm of lifelong learning;
- Measuring progression, survival and drop-out in tertiary education;
- The expansion of the knowledge base on the financing of tertiary education - to capture new financing strategies for mobilising resources, promoting efficiency, financing change and encouraging responsiveness;
- Conceptual and data development in the area of public subsidies for education;
- The development of methods for the structural analysis of indicators on costs and resources.

726. The technical group also devoted a major, and ongoing, effort to the development of trend indicators on both participation and spending patterns and to improving the coverage of private educational provision and funding at all levels of education.

727. Most recently, the INES Technical Group started to extend the coverage of its work from a traditional focus on initial education towards developing a statistical framework for measuring progress towards lifelong learning. The following chapter, which was prepared by Walter Hörner from the German Federal Statistical Office, lays out the objectives and challenges for this work.
In addition to the various developmental activities, the INES Technical Group devotes considerable attention to the consolidation of the INES information base. Member countries continue to co-operate to gather relevant information and to develop and apply common definitions and criteria for the quality control of the data.

- A first step after Lahti was the development of the UNESCO/OECD/EUROSTAT (UOE) data collection on education statistics. The objective of this instrument, which was both developed and implemented through the Technical Group, was to regularly provide internationally comparable data on key aspects of education systems. Since 1996 this instrument is administered annually in electronic form, replacing various paper-and-pencil data collections that the three organisations had previously administered independently. The preparation of a manual for internationally comparative education statistics is one of the next steps envisaged by the Technical Group.

- As a second step, following OECD’s collaboration with UNESCO on the development of a new framework for the revised International Standard Classification of Education (ISCED), the Technical Group produced an operational manual for the implementation of ISCED-97 in OECD Member countries (OECD, 1998). This manual paved the way for considerable improvements in the comparability of the OECD education indicators. The approach taken was a practical and empirical one: the Technical Group carried out a special survey of the main educational pathways in OECD countries and then mapped each country’s system onto ISCED-97. Subsequent negotiations between countries, through the Technical Group, led to a better understanding of the similarities and differences of education systems and to an improved and more comparable allocation of national programmes to ISCED-97. The 2000 edition of Education at a Glance was the first publication based on ISCED-97. However, further progress will be needed to improve the alignment of the classification of educational programmes across countries.
TOWARDS A STATISTICAL FRAMEWORK FOR MONITORING PROGRESS TOWARDS LIFELONG LEARNING

Walter Hörner, Federal Statistical Office in Germany, INES Technical Group

Introduction

1. Many countries are seeking to implement the vision of lifelong learning and to measure progress towards making lifelong learning a reality for all citizens. At the same time, the question of what the concept of lifelong learning precisely means and how it should be operationalised so as to become amenable to quantitative assessment is still controversially debated. As early as 1996 did the OECD note the lack of quantitative evidence:

"Most importantly, there is a critical absence of quantitative information on lifelong learning and its impact on society and the economy. Since economies can no longer rely solely on a gradual expansion of initial schooling to meet the demands for new and high level skills, other indicators are needed that will help the policy makers to improve the foundations for lifelong learning. ...Understanding the factors that influence patterns of learning throughout the lifecycle will be a challenge." (OECD 1996, p. 11).

2. The aims and efforts of education policy concerning lifelong learning cannot be evaluated effectively without relevant empirical data and indicators. Monitoring lifelong learning in quantitative and statistical terms is necessary in order to determine what incentives and actions are needed in this field and whether they prove to be effective.

3. This paper was prepared by Walter Hörner following a first discussion of this matter by the INES Technical Group in May 2000.

The concept of lifelong learning

From the transmission of knowledge to “learning how to learn”

4. The enormous speed with which knowledge develops and ages today underlines at least two points:

– Learning must be conceptualised as a permanent process which starts at birth and continues throughout life;

70 See the overview regarding the contribution of formal education systems in the member countries of the European Union in Eurydice (2000). Among the multitude of relevant national initiatives are: Australia ("Learning for Life"), Finland ("Joy of Learning – National Lifelong Learning Strategy"), Great-Britain ("The Learning Age", "Learning Direct Initiative", "University for Industry"), the Netherlands ("Lifelong Learning – The Dutch Initiative") and Germany (BMBF action programme on lifelong learning for all, BIBB-project "L²-Lebenslanges Lernen").
Reflecting its diversity and complexity, learning needs to be understood in a wide societal and personal sense, beyond the context of job-related training.

5. As a result, learning cannot be characterised as a limited phase in childhood and youth dedicated only to the promotion of the individual’s vocational career. Rather, learning is a continuous task of the society and the individual that extends to all areas and phases of life. The perspective of learning "from cradle to grave" is a fundamental element in most approaches to lifelong learning.

6. Usually, data and analysis on educational activities and educational indicators have been focused on formal education and its institutional providers. Taking into account the educational and economic environment of today, it seems necessary to put lifelong learning in a wider perspective, covering also less formalised education programmes as well as informal learning activities. Generally, the latter take place outside the formal and non-formal educational institutions. Moreover, while traditional analyses see the educational programmes of the formal education system as their main subject of investigation, lifelong learning brings the individual and his or her educational activities into focus, with people often benefiting from different types of learning opportunities in different phases of their life.

7. Lifelong learning, however, is not just a simple summing-up or integration of traditional education programmes and modern learning opportunities. Instead, the approach to lifelong learning includes fundamental differences in educational content and perspectives: While traditional educational institutions have been (and still are) primarily concerned with transmitting knowledge, modern learning opportunities and the lifelong learning approach put the emphasis on the development of individual capabilities and personal learning competencies. At the heart of the lifelong learning concept is the idea of enabling and encouraging people "to learn how to learn".

8. Against this background, lifelong learning implies a paradigmatic shift from the dominance of traditional education institutions to a diverse field of traditional and modern learning opportunities that are more process-oriented and of a modular structure. At the same time responsibility for education and learning shifts from the government to a broader range of both public and private education providers and the individual him/herself. Political support for the implementation of lifelong learning (e.g., by providing an adequate institutional, financial and legal framework) presumes an analysis of current education and learning structures that is based on reliable and comparable data.

9. While formal barriers access to education might decrease in importance in the future, the volume of time and money that individuals are able and willing to invest for their own learning might become more important.

10. Another point in this context is how learning is organised by the society (public) and where it is located. While learning currently is often expected to take place in institutions of formal education, it is increasingly recognised that learning occurs in very different settings and locations, including schools, universities, the work place in its very different settings, the family etc.

11. An example is distance learning which offers opportunities which are neither restricted by physical distance nor by time. The same is true with respect to most of what is called informal learning. Especially digital networks allow for a high level of independence concerning time and space (e.g., Computer Based Training – CBT, Internet Based Training - IBT, Tele Learning). The growing flexibility of distance learning being possible due to advances in modern information and communication technologies leads to integrated patterns of non-formal and informal learning that become, statistically, increasingly difficult to distinguish.

The 'lifelong-lifewide framework'

12. In the past the time-dimension has often dominated the discussion about lifelong learning. It underlines the fact that learning activities occur in different phases over the lifecycle. However, there is also a lifewide-dimension to be recognised, which is of growing importance today. It takes into account
that learning increasingly takes place in a multitude of settings and situations in real life (see
Figure 1):
13. With respect to the time dimension of lifelong learning (vertical axis) the sequence of different learning processes within the life cycle is of primary concern (lifelong learning in a narrow sense). The lifewide dimension (horizontal axis) refers to the different social settings in which learning takes place (lifewide learning).  

14. The 'lifelong-lifewide framework' has now been widely accepted. The more traditional educational processes have been institutionalised (e.g., within formal education) and mostly embedded in a more or less strictly organised time schedule (e.g., within compulsory education). Others might be used flexibly in different phases of life by the individuals according to their own occupational status or situation at the workplace or according to their personal interest and motivation (e.g., within informal learning).

15. The focus of formal education is on young people. In

---

72 Cf. OECD Ad-hoc Group on Lifelong Learning (1999), Chapter 4.
Figure 1 this is localised in area 3. With respect to the implementation of lifelong learning there are essentially two main challenges:

- First, the existence and importance of education and training in the remaining areas of
Figure 1 must be recognised by political decision-makers and the society. While the extension of continuing education and training, which is mostly located in areas 2 and 1 in
Figure 1, is already underway, the activities in areas 4 and partly in 1 still need more emphasis. A statistical documentation of the situation of lifelong learning might help to succeed in these efforts.

- Second, the main focus of education policy will shift from formal, largely publicly provided education (area 3) in
Figure 1), which is highly regulated and to a large extent compulsory, to the other areas, representing more flexible learning activities.

16. As a result, a statistical concept for lifelong learning should try to capture those learning activities that are outside the formal sector and which require a high level of motivation and personal commitment. Due to fundamental differences among learning activities, it is not an easy task to collect statistical data on these types of learning activities. Therefore, in order to be able to draw a coherent and comprehensive picture of lifelong learning there is, as a first step, a need for a discussion on what educational activities should be included in a statistical approach to measuring lifelong learning. Especially with respect to informal learning new methodological instruments will be needed. Informal learning cannot be regarded as an unimportant side-effect. Estimates concerning the relative weight of self-directed learning, which constitutes a main component of informal learning, indicate that actually up to 90 per cent of all learning activities fall within this category (Candy / Crebert 1997, p. 88).

Scope of lifelong learning

17. In order to determine the specific learning situations that should be covered by a statistical framework on lifelong learning, it is important to differentiate between education and learning: The International Standard Classification of Education (ISCED) defines education as follows: "...the term education is [...] taken to comprise all deliberate and systematic activities designed to meet learning needs." (ISCED '97, Paragraph 7). In contrast, learning is characterised from the perspective of the individual as "...any improvement in behaviour, information, knowledge, understanding, attitude, values or skills." (ISCED, Paragraph 9).

18. Education consists of two sub-fields: formal and non-formal education. Within learning one can distinguish between intended and unintended learning. Unintended learning takes place in all situations and under all circumstances within the whole lifespan. Although unintended learning may be of great importance in real life, it must be excluded from the statistical considerations discussed in this paper since unintended learning can hardly be captured empirically. So, in our context, the focus is restricted to intended learning. As far as this takes place outside formal or non-formal educational institutions, it is labelled as informal learning. See Figure 2 for an illustration of this point.
19. The activities to be covered therefore comprise three main areas: formal and non-formal education as well as informal learning:

- **Formal education**: According to ISCED formal education refers to "...the system of schools, colleges, universities and other formal educational institutions that normally constitutes a continuous 'ladder' of full-time education for children and young people, generally beginning at age five to seven and continuing up to 20 or 25 years old." (UNESCO 1997, p. 41). In some countries, however, these age limits need to be extended.

- **Non-formal education** comprises "any organised and sustained educational activities that do not correspond exactly to the above definition of formal education. Non-formal education may therefore take place both within and outside educational institutions, and cater to persons of all ages." (UNESCO 1997, p. 41).

- **Informal learning** consists of all intended learning activities and/or situations that cannot be classified as formal or non-formal education. Informal learning activities are characterised by a relatively low level of organisation and may take place at the individual level (e.g., self-directed learning) as well as in groups of people (e.g., at the workplace or within the family). Some of these settings, such as computer based learning modules or other structured learning material may to a large extent be similar to non-formal programmes while others, such as learning within the family or by visiting cultural events, can hardly be captured statistically.

20. Unfortunately, the aforementioned definitions do not indicate how to collect data on (formal or non-formal) educational or (informal) learning activities. The classification unit of ISCED is the educational programme. This pragmatic approach taken by ISCED allows to map activities of formal education in a relatively comprehensive and internationally comparable way. However, there is evidence that it is difficult to proceed in the same way concerning non-formal education even if also in non-formal
education a lot of educational activities can be mapped by using educational programmes as a classification unit. As regarding informal learning the educational programme can generally not be used as a classification unit because informal learning largely takes place outside organised programmes.

**Draft definition and basic elements of lifelong learning**

21. Summarising the above considerations in order to get something like a working definition on lifelong learning, one might conclude that: lifelong learning comprises all learning processes and activities that an individual makes use of during his or her entire life and that are intended to acquire and extend knowledge and skills for personal, vocational and societal purposes.

22. This is not an operational definition, but it describes the idea of lifelong learning in a short and comprehensive way. A first step to make this definition more operational for the purpose of data collection could be to show how it can be used in different fields of educational activities:

   (a) **Early childhood education** captures organised educational programmes directed to toddlers before compulsory education starts (ISCED 0), e.g., kindergarten, pre-school, music and language teaching for 3-5 years olds.

   (b) **Compulsory education** comprises all compulsory educational programmes for children and youth, irrespectively whether they are provided by public or private institutions (essentially ISCED 1 and ISCED 2).

   (c) **Post-compulsory education** includes all educational programmes provided within the formal education sector that follow (full-time) compulsory education (ISCED 3-6).

   (d) **Continuing vocational education and training** refers to job-related educational programmes mainly outside the formal education sector, which are organised in a way similar as those in formal education. It may include programmes of initial and non-initial vocational education and training, which are not part of formal education. While these programmes are, to a large extent, provided by employers for their employees (e.g., at the work place), they may also be managed by external instructors or even organised and sometimes financed by other institutions (e.g., labour market institutions, such as ministries of labour, trade unions, or private providers of educational services).

   (e) **Non-formal general education** includes educational programmes outside the formal education sector that are not (directly) related to vocational objectives, e.g., adult evening classes, education programmes provided by the churches, open and distance learning. It also includes courses (in foreign language, computer skills, music etc.) or other remedial courses outside formal education institutions for young people (who are at the same time regular participants of the formal education sector) or young adults.

   (f) **Informal learning** captures all remaining learning processes that are planned (i.e. intended). It comprises learning activities of people of all ages and in all learning settings or situations. It includes learning within informal groups and families as well as self-directed learning (e.g., making use of different media, such as books or computers).

23. This classification deviates from ISCED which is classifying programmes according to levels. But it seems more appropriate in order to capture the lifelong-lifewide framework mentioned earlier.

24. Categories (a), (b) and (c) refer to the sequence of traditional educational stages as regards time and content. These categories mainly correlate with formal education. Categories (d) and (e) constitute the non-formal education sector. In contrast to this chronological and sectoral differentiation, category (f) captures a cross-section of activities in all settings and in all age-groups, comprising all planned learning activities outside formal and non-formal education.
**Coverage of lifelong learning**

25. Traditionally, formal education is the core of education policy and statistics, although it refers only to a relatively short period of time within the lifespan. By contrast, informal learning activities might take place during the whole lifespan. The intensity of learning and the extent of knowledge acquired may vary greatly, not only between different learning activities, but also between age-groups and individuals. In contrast to the phase of formal education in which educational activities are in the centre of the time use of an individual, in later phases of life only a minor share of time may be explicitly dedicated to learning activities.

26. Figure 3 highlights the fact that different categories of learning activities are related to different target groups. This implies that the mix of relevant learning activities will usually vary between different age-groups.

![Figure 3: Coverage of lifelong learning according to the 'lifelong-lifewide framework'](#)

<table>
<thead>
<tr>
<th>Age</th>
<th>Formal settings</th>
<th>Non-formal settings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early childhood education</td>
<td>Compulsory education</td>
</tr>
<tr>
<td>0-5</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>6-10</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>11-15</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>16-20</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>21-25</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>26-30</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>31-35</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>36-40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41-45</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>46-50</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>51-55</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>56-60</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>61-65</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>66-70</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>71-75</td>
<td>shaded</td>
<td></td>
</tr>
<tr>
<td>76-99</td>
<td>shaded</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Shaded areas indicate when the programmes and activities take place within the whole lifespan. This chart gives only a rough order of magnitude, since age-groups do not correspond exactly to the duration of specific programmes and activities.

27. Figure 3 also underlines that formal education consists of a sequence of a relatively small number of modules forming an educational ladder. While formal programmes are primarily directed to children, youth and young adults, i.e. a relatively small group of people in most countries, a large part of educational content is conveyed in this category. Non-formal educational programmes may be targeted at the whole population or at some sub-groups, such as, e.g. the working population, while informal learning activities generally take place in all settings and within all age-groups.

28. The faster knowledge ages the more the formal education system will have to focus on the development of personal knowledge and skills which are the prerequisites for continuous learning. Hence, in the future the acquisition of professional knowledge will, at least to some extent, be transferred to later phases of life.
29. Compulsory education as well as subsequent post-compulsory programmes are already well documented in traditional education statistics. The development of a statistical approach of lifelong learning should therefore have an emphasis on the remaining fields of learning, which are not sufficiently represented in present education statistics.

**Measuring lifelong learning**

**Indicators on lifelong learning**

30. In the context of lifelong learning, similar to traditional education statistics, three types of indicators can be distinguished:

- indicators on access and participation in learning activities,
- indicators on investment in learning, and
- indicators on the output and outcomes of learning activities.

31. Indicators on access and participation focus on the number of participants and on their socio-economic characteristics. Additionally, it would be desirable to have data concerning the time and intensity individuals invest in education and learning.

32. Indicators on investment in learning provide data on, for example, total expenditure, public and private expenditure as well as on public and private sources that are used for financing learning activities.

33. Indicators on the output or outcomes of learning activities show the returns on investment in learning. As regarding formal education, data can be captured by collecting data on graduates having some kind of formal certificate which documents a certain qualification. A promising alternative approach would be to collect data on the knowledge and skills of individuals, which is of particular relevance for assessing learning outcomes in educational activities that are difficult to connect to formal qualifications and certifications.

34. At present, data are available only for some of these categories, while for others they are missing or inadequate. Figure 4 illustrates this point.

**Figure 4: Indicators on lifelong learning and data availability**

<table>
<thead>
<tr>
<th></th>
<th>Participation</th>
<th>Financing / Expenditure</th>
<th>Output / Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Early childhood education</td>
<td>( )</td>
<td>( )</td>
<td>a</td>
</tr>
<tr>
<td>(b) Compulsory education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Post-compulsory education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) Continuing vocational education and training</td>
<td>( )</td>
<td>( )</td>
<td>m</td>
</tr>
<tr>
<td>(e) Non-formal general education</td>
<td>( )</td>
<td>( )</td>
<td>m</td>
</tr>
<tr>
<td>(f) Informal learning</td>
<td></td>
<td>m</td>
<td>m</td>
</tr>
</tbody>
</table>

**Notes:** The degree of data availability is given by the intensity of the shaded areas:

- ( ) = data available
- a = data only partially or poorly available
- category does not apply
- m = data actually not available or not directly available
35. The importance of education and learning may vary between different phases of life. Hence, in order to be able to calculate age-specific participation rates, it would be desirable not only to have complete data on the number of participants, but also detailed information on their age. Data on the age of participants could be used to establish a comprehensive learning profile of the population that encompasses all fields of lifelong learning as well as the whole life cycle. This would make it possible to analyse the importance of different learning activities in different phases of life. Hitherto, such information is limited, mainly referring to programmes of formal education.

36. Often, it is difficult to measure participation in informal learning directly. It can be expected that the variation in volume and intensity of the utilisation of informal learning opportunities is much higher than that of educational programmes in formal and non-formal education as defined above.

37. As mentioned, the collection of data on informal learning would require an agreement on the learning activities that should be captured. It would be desirable to establish a specific classification scheme for informal learning that fulfils a similar function as ISCED does with respect to organised educational programmes. Such a classification scheme would need to be related to learning activities, rather than to educational programmes as it is the case in ISCED. In the following, some examples of those activities are given:

- the extent of the utilisation of print media for learning (e.g., specialist books or magazines),
- the extent of the utilisation of modern information and communication technologies – ICT (e.g., computer based training, internet based training, educational broadcasting),
- the extent of the utilisation of libraries (e.g., frequency and duration of attendance, number of books lent),
- the frequency and duration of attendance at museums, (scientific) exhibitions and other cultural events with educational content.

38. As long as such survey data is not available, one could try to make use of proxy variables, such as the distribution of relevant goods or equipment that are used for informal learning activities, as well as related services in society. Examples are data on purchased computers, books, magazines and newspapers as well as aggregate data on the extent households make use of internet access, educational broadcasting, educational journeys or cultural events with educational content. Figure 5 provides an overview of possible variables that might be used in order to get data on informal learning indirectly.

**Figure 5: Proxy indicators on informal learning**
Extent of "utilisation" of instruments for informal learning

<table>
<thead>
<tr>
<th>Category</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Share of households with at least one computer</td>
<td></td>
</tr>
<tr>
<td>• Share of households with internet access</td>
<td></td>
</tr>
<tr>
<td>• Using computers beyond work (average hours per day)</td>
<td></td>
</tr>
<tr>
<td>• Using the internet beyond work (average hours per day)</td>
<td></td>
</tr>
<tr>
<td>• Expenditures of private households on computer equipment (p.a.)</td>
<td></td>
</tr>
<tr>
<td>• Expenditures of private households on internet access (p.a.)</td>
<td></td>
</tr>
<tr>
<td>• Expenditures of private households on books, magazines and newspapers (p.a.)</td>
<td></td>
</tr>
<tr>
<td>• Number of users of libraries and/or number of books lent out (p.a.)</td>
<td></td>
</tr>
<tr>
<td>• Number of visitors of museums, exhibitions and other cultural events (p.a.)</td>
<td></td>
</tr>
<tr>
<td>• Number of people participating in education oriented travelling (p.a.)</td>
<td></td>
</tr>
<tr>
<td>• Extent of using educational broadcasting (average hours per day)</td>
<td></td>
</tr>
<tr>
<td>• ...</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context indicators</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Socio-economic characteristics of users and their families (age, sex, marital status, educational attainment, income, labour force status, etc)</td>
<td></td>
</tr>
<tr>
<td>• ...</td>
<td></td>
</tr>
</tbody>
</table>

39. Additionally, it would be desirable to have data on the socio-economic structure of each user or learning group, at least on the age structure of the participants, as well as on the time spent on each of the related learning activities.

Obstacles in statistical measurement and comparability: the aggregation problem

40. Lifelong learning consists of a number of very different educational and learning activities that are often closely interconnected as regarding both time and content. These activities take place in different phases of the life cycle and differ with respect to their learning intensity and extent.

41. In an attempt to quantify learning activities, two main problems arise:

- Some of the activities to be taken into account cannot be quantified and are therefore hard to capture. Especially with respect to informal learning proxy variables, at least in part (see Figure 5), may be needed as substitutes.

- In some cases the aggregation of variables into an overall indicator is not possible due to differences in the dimensions of these variables. For example, it is difficult to aggregate data on participation in a university-based seminar or lecture with data on a visit to a scientific library or the reading of a book.

42. As far as formal and non-formal education is concerned, the number of participants in educational programmes might be aggregated quite meaningfully. However, a direct aggregation of the extent of informal learning activities imposes problems. This is not only due to data shortages, but also due to the fact that an aggregation of data does not seem sensible if different dimensions are involved.

43. In many cases informal learning activities differ significantly from an organisational point of view, while they are complementary as regards content (e.g., self-directed learning aimed at preparing oneself for attending a formal lecture or passing an exam). In other cases, informal learning activities which seem to be quite different at first sight, may have a similar objective with respect to content and output. So these activities might be substitutes for one another (e.g., reading a professional book instead of attending a formal lecture, both covering the same subject matter). Hence, with respect to a comprehensive mapping of lifelong learning it would be helpful not to present data on the different learning activities in...
isolation. If possible, aggregate figures should be constructed which might also be better instruments for international comparisons.

**Data collection**

*Strategies for data collection*

44. There already exist a number of methods and strategies for collecting data on learning activities. Traditionally, data have been based on information provided by educational institutions. In this case the focus was on the provision of education, with data collected through educational programmes. Besides these traditional data sources, there are others, such as household surveys, which focus on households or individuals as consumers or participants in education activities. Both methods can contribute to a mapping of lifelong learning.

45. Generally, each of the existing sources refers only to a small range of learning activities and cannot cover all the activities which have to be taken into account in the context of lifelong learning. Despite these limitations, the (potential) wealth of information incorporated in the existing sources allows for significant progress in mapping lifelong learning. Figure 6 provides an overview of existing data sources on education and learning and proposes some further strategies for data collection in the field of lifelong learning.
## Figure 6: Strategies for data collection on lifelong learning

<table>
<thead>
<tr>
<th>Method/Source</th>
<th>Contents / scope</th>
<th>Survey unit</th>
<th>Countries covered</th>
<th>Strengths and weaknesses (with respect to monitoring lifelong learning)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing sources</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UOE data base / INES project</td>
<td>Participation in education at country level (students enrolled, repeaters, graduates, organisations, personnel, finance); demographic, social, financial indicators</td>
<td>Programmes</td>
<td>OECD countries, others</td>
<td>International comparable information on participation, completion, expenditure and types of resources dedicated to education; covers mainly public providers; outcomes in terms of graduates; includes data on the years of age of the participants Not included: vocational training at the work place</td>
</tr>
<tr>
<td>European Labour Force Survey (LFS)</td>
<td>Participation in different kinds of education and training; educational attainment as a proxy for skills; link with employment status and background variables; ad-hoc module on lifelong learning probably in 2002</td>
<td>Households</td>
<td>Member States of the EU</td>
<td>Intends to cover the whole resident population Refers to a particular survey week in spring</td>
</tr>
<tr>
<td>Continuing Vocational Training Survey (CVTS I+II)</td>
<td>Participation in continuing education and training in enterprises; costs and financing of training; type of activity, inside vs. outside training, gender</td>
<td>Enterprises</td>
<td>Member States of the EU</td>
<td>Includes parts of self-directed learning (at the work place); certification Data collected only from employers; outcome in terms of occupational skills</td>
</tr>
<tr>
<td>European Community Household Panel (ECHP)</td>
<td>Participation in education; languages used New survey under development</td>
<td>Households</td>
<td>Member States of the EU</td>
<td>Information collected from the individual; linked to working and living conditions; covers only parts of relevant activities</td>
</tr>
<tr>
<td>OECD Programme for International Student Assessment (PISA)</td>
<td>Assessing effectiveness of educational systems at national levels Refers to reading literacy, mathematical literacy, scientific literacy First assessment in 2000 (results in 2001), thereafter: three-year cycles</td>
<td>Individuals (only 15 year olds)</td>
<td>32 countries, of which 28 are OECD members</td>
<td>Internationally comparable data on learning outcomes of &quot;completing&quot; compulsory schooling Includes assessment of cross curriculum competencies Limited age group (15 year olds)</td>
</tr>
<tr>
<td>International Adult Literacy Survey (IALS)</td>
<td>Assessment of literacy and numeracy skills of the adult population (prose literacy, document literacy, quantitative literacy)</td>
<td>Individuals (only adults)</td>
<td>Canada, Germany, Netherlands, Poland, Sweden, Switzerland, United States, Australia, Belgium (Flanders), Ireland, New Zealand, United Kingdom</td>
<td>Internationally comparable data on literacy and numeracy skills Refers only to adults</td>
</tr>
<tr>
<td>International Life Skills Survey (ILSS)</td>
<td>Direct assessment of life skills, such as prose and document literacy, numeracy, problem solving, practical cognition, teamwork, ICT literacy</td>
<td>Individuals (only adults)</td>
<td>Australia, Belgium, Canada, Czech Republic, Denmark, Finland, France, Hungary, Italy, Korea, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States</td>
<td>Going far beyond adult literacy skills, ILSS intends to measure a broader set of basic life skills Refers only to adults</td>
</tr>
<tr>
<td>TIMSS-R (Third International Mathematics and Science Study - Repeat) 1999</td>
<td>Direct assessment of mathematics and science skills</td>
<td>Individuals</td>
<td>Australia, Iran, New Zealand, Belgium (Flemish), Israel, Philippines, Bulgaria, Italy, Romania, Canada, Japan, Russian Federation, Chile, Jordan, Singapore, Chinese Taipei, Republic of Korea, Slovak Republic, Cyprus, Latvia, Slovenia, Czech Republic, Lithuania, South Africa, England, Republic of Macedonia, Thailand, Cross-country and time comparisons (1994/95 with 1999 data)</td>
<td></td>
</tr>
</tbody>
</table>
46. Basically, lifelong learning not only calls for an extension of the scope of learning activities to be taken into account, but also for changes in content. As mentioned before, the transmission of knowledge is no longer sufficient in the information society. There is also a need for developing personal skills and competencies. A key point of departure concerning monitoring lifelong learning consists of shifting attention from counting inputs, such as resources dedicated to education and learning (including time), to measuring output or outcomes. The best way to do this would be to introduce separate surveys specifically designed to meet data needs with respect to lifelong learning. Such surveys would open the opportunity to cover all relevant learning activities, and they would also allow for the collection of specific data on basic skills needed for personal and professional life.

47. As demonstrated by the International Adult Literacy Survey, direct measures of skills are better adapted to predict successful participation in society than traditional indicators such as education credentials. The same may be true with respect to statistical measures of progress in the realisation of lifelong learning goals. While national skill surveys, putting their emphasis on different learning outcomes, already exist in some countries, a coherent and comprehensive mapping of learning according to the approach of lifelong learning must still be developed at the international level.

48. A promising departure concerning this issue might be a survey referring to the use of personal time. A focus on the time individuals spend on learning (and education) would bring the individual into the centre of the discussion. With such a statistical instrument it would be possible to collect coherent and comparable data on a broad range of learning activities (including formal/non-formal education and informal learning as well as socio-economic characteristics of the learners themselves, e.g. their year of
Since all data would be based on the common dimension "time", the results could be presented in a comparable form (see Figure 7):

**Figure 7: Volume of learning based on time use data (in hours per day)***

<table>
<thead>
<tr>
<th>Age</th>
<th>Formal education</th>
<th>Non-formal education</th>
<th>Informal learning</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-2</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>3-5</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>6-10</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>11-15</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>16-20</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>21-25</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>26-30</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
<tr>
<td>...-99</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
<td>![Bar graph]</td>
</tr>
</tbody>
</table>

1 Hypothetical data.

49. The time use approach has the advantage, that the time used for learning could serve as a "common denominator". This would offer the opportunity to construct aggregated indicators at the national level, but would also allow for cross-country comparisons on the volume of learning. Hence, based on data on time use a coherent and comprehensive picture of the volume of lifelong learning could be drawn. Moreover, it should be possible to link these data with the more detailed and less aggregated traditional data based on educational programmes of formal education.

50. At the same time, the time use approach considers only a single dimension of lifelong learning related to the volume of participation (broken down into fields of educational activities etc.). The volume may not strictly correlate with the quality of learning. A time-based instrument also cannot provide information on financing or on learning outcomes. A household survey specifically designed for collecting information on lifelong learning would be a better instrument to collect data on educational and learning activities that are comparable across countries. 73

51. An alternative to a new separate survey on lifelong learning could be to supplement existing surveys by ad-hoc modules on lifelong learning specifically designed to meet key information requirements with respect to monitoring lifelong learning. In fact, such a module is planned with respect to the European Labour Force Survey (LFS) for 2002.

**Recent approaches to monitoring lifelong learning**

52. The demand for data on lifelong learning has been growing rapidly. Annex 2 contains an example of a very recent initiative of the European Commission to establish a set of indicators on lifelong learning. This initiative is intended to monitor the current situation on lifelong learning in the EU countries with respect to the EU employment guidelines decided on in March 2000 at the Lisbon Summit.

73 Taking into account the constraints in the public budget and the restricted capabilities for new surveys in many countries it can be expected that there is not much room for new international surveys which are specifically designed for monitoring lifelong learning.
of the European Council. Moreover, in May 2000 a task force on measuring lifelong learning has been launched by Eurostat.\textsuperscript{74} While the EU employment guidelines refer only to employment issues, which cover just a small section of lifelong learning only, the objective of the Eurostat task force is to cover a broader range of learning activities including an evaluation of existing data sources on lifelong learning.

53. In addition, several countries have implemented national action programmes on lifelong learning in the recent past, that include measures, specific tasks and responsibilities which contribute to the process of instituting lifelong learning.

54. These examples underline that some initiatives and some work on the statistical measurement of lifelong learning are already underway. At the same time, in most cases the approaches emphasise specific aspects of learning. An overall, comprehensive concept or approach which could be the base for the production of internationally comparable data on lifelong learning is still lacking.

\textit{Conclusion}

55. The paper underlines that lifelong learning encompasses a number of very different aspects. Integrating the “time dimension” into a statistical approach to measuring lifelong learning would allow to capture the volume of formal and non-formal education as well as of informal learning in a consistent way. This could be a first step towards an international benchmarking and improve the data base for national decision-makers. Although the time dedicated to learning is an important indicator it is, nevertheless, only one of the elements necessary for monitoring progress towards lifelong learning. To capture other aspects, including the quality of learning, investments in learning, and the output and outcomes of learning activities in a comparable way a special survey on lifelong learning would be needed. To develop such a survey instrument could be an important goal in the long-term. In the medium-term, the development of modules on lifelong learning which could be integrated into existing national surveys in a harmonised way should be considered.

\textit{Annex 1: Policy dimensions of lifelong learning}

Due to its broad conception, lifelong learning is not related to educational policy alone, but also to a number of other fields, e.g. economy, labour, social affairs and finance, to mention just the most important ones. Hence, the implementation of lifelong learning is not a task of the ministry of education alone but of a much wider range of governmental and non-governmental stakeholders.

The question, to which extent the different ministries could contribute to the implementation of a national action programme on lifelong learning is of high importance. Figure 8 provides an overview on possible starting points for different public stakeholders.

\textit{Figure 8: Contribution of political units (ministries etc) to the implementation of lifelong learning}

<table>
<thead>
<tr>
<th>Department</th>
<th>Formal settings</th>
<th>Non-formal settings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early childhood education</td>
<td>Compulsory education</td>
</tr>
<tr>
<td>Education</td>
<td>Educational reform</td>
<td>Financial support to suppliers and consumers of learning activities</td>
</tr>
<tr>
<td>Economy</td>
<td></td>
<td>Certification schemes</td>
</tr>
<tr>
<td></td>
<td>Financial support</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{74} Within the OECD-INES project a task force on lifelong learning was established already in 1998.
### Annex 2: Monitoring the implementation of EU employment guidelines on lifelong learning

Education and training policy constitutes a major element in the European employment strategy. Therefore, EU employment guidelines take into account the contribution lifelong learning can make to fulfil the guidelines. In order to monitor the implementation of these guidelines of the European Commission (EC) proposes the following empirical indicators.

**LLL** = Educational attainment of secondary level by age groups and employment status  
**LLL** = Participation of adults in education and training by gender and age groups  
**LLL** = Proportion of workforce in job-related training resulting from collective agreements by employment status and age groups (15-24 and 25-59 year olds)  
**LLL** = Proportion of early school leavers (drop-outs) by gender  
**LLL** = Proportion of 18-29 years olds by employment status one year after leaving education by gender and education level  
**LLL** = Number of students per computer by education level (primary, secondary)  
**LLL** = The number of Internet connections per school by education level (primary, secondary)  
**LLL** = Share of teaching hours computer supported

<table>
<thead>
<tr>
<th>Department</th>
<th>Formal settings</th>
<th>Non-formal settings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early childhood education</td>
<td>Compulsory education</td>
</tr>
<tr>
<td>Certification</td>
<td>Recognition of learning efforts</td>
<td></td>
</tr>
<tr>
<td>Improving knowledge about the requirements of the economy concerning staff qualifications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labour</td>
<td>Financial support for vocational training, especially in favour of the unemployed</td>
<td>Financial support for vocational learning activities, especially in favour of the unemployed</td>
</tr>
<tr>
<td>Social affairs</td>
<td>“Equity”: improving access of disadvantaged, vulnerable and marginalised groups to education, training and learning</td>
<td></td>
</tr>
<tr>
<td>Make education and learning more compatible to the working life and to family-related obligations (e.g. child care, care of the elderly)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finance</td>
<td>Tax relief for suppliers of educational services and/or of complementary goods</td>
<td></td>
</tr>
<tr>
<td>Tax relief for consumers of educational services and/or of complementary goods</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The points mentioned above indicate possible points of departure concerning the implementation of lifelong learning. Ultimately, progress in this field cannot be achieved without the co-operation of different political units and private organisations, such as firms and economic associations.
One has to bear in mind that the focus of the EU employment guidelines is on employment issues rather than on the concept or measurement of lifelong learning in general. Conceptually, the indicators are based on a broad definition of lifelong learning provided by EU. On the one hand, it is ascertained that a number of learning activities of lifelong learning take place across the whole lifecycle. On the other, it is acknowledged that these activities cannot be summarised in a single aggregate measure. Additionally, it is pointed out that in most member countries the scope of lifelong learning is more narrowly defined, mainly referring to continuing vocational education and training.

Modern information and communication technologies (ICT) are widely used in formal and non-formal education at the present time and they will also become an indispensable instrument for informal learning in the future. The fact that these technologies play an important role in lifelong learning is reflected in the above mentioned set of indicators.

Annex 3: Inclusion of non-formal education and informal learning in selected national education surveys

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey</th>
<th>Non-formal education</th>
<th>Informal learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>Canadian Adult Education and Training Survey</td>
<td>• Classroom instruction (continuing education) • Seminars (continuing education) • Workshops • Correspondence courses</td>
<td>• Training on the job • Educational software • Radio or TV broadcasting • Audio/video tapes or discs • Reading material • Internet based learning</td>
</tr>
<tr>
<td>Finland</td>
<td>Adult Education Survey</td>
<td>• Evening school • Apprenticeship training / training for adults (at vocational adult education centres or vocational education institutes) • Employment promoting courses (financed by labour authorities) • Further training at the university / continuing studies centres • Summer university courses • Folk high school / folk academy • Studies at the music school adult department / college</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Reporting System on Continuing Training (Berichtssystem Weiterbildung)</td>
<td>Questions on non-vocational courses (practical skills, arts, literature, sports etc) Questions on vocationally oriented courses (for another occupation, new challenges in the job, career advancement etc)</td>
<td>Questions on participation of employed persons in job-related informal learning (instruction at the work place by colleagues, self-directed learning, work shops, quality circles etc)</td>
</tr>
<tr>
<td>Sweden</td>
<td>Swedish Surveys of Staff Training</td>
<td>• Participation in job-related training paid for totally or partly by the employer (courses, seminars etc)</td>
<td>• Self-instruction • Conferences</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Swiss Survey on CET (module of regular LFS)</td>
<td>• Job-related continuing education courses (in-house training, leisure-time courses etc) • Non job-related continuing education courses</td>
<td>• Audio/video tapes • Computer aided learning • Radio/TV courses • Conferences • Special literature • Training at the work place</td>
</tr>
<tr>
<td>UK</td>
<td>National Adult Learning Survey</td>
<td>• Courses leading to job-related qualifications</td>
<td>• Self-directed learning based on materials provided by a training provider</td>
</tr>
<tr>
<td>Country</td>
<td>Survey</td>
<td>Non-formal education</td>
<td>Informal learning</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>----------------------</td>
<td>-------------------</td>
</tr>
</tbody>
</table>
| USA     | National Household Education Survey (NHES) | • Courses for adults leading to non job-related qualifications  
• Instructions in playing an instrument, in an art or crafts, in sport or other skills  
• Evening classes | • Participation in adult education (basic education, English as a foreign language, job-related courses etc) |
|         | National Education Longitudinal Survey (NELS) | • Continuing education or non-credit courses  
• Private instruction or tutoring  
• Job-related training (required / encouraged by employer) | • Self-directed learning (courses by mail, radio, TV, computer based training etc) |

**References**


SECTION V

CONTRIBUTION FROM THE AD-HOC GROUP ON EQUITY
TOWARDS A SYSTEM OF EQUALITY AND EQUITY INDICATORS

Walo Hutmacher, Switzerland

1. This chapter focuses on the conceptual debate and presents proposals for a system of equality and equity indicators within the general framework of INES education indicators. It reflects the results of the work of an ad hoc group on equity issues mandated by the INES steering group. Both a forthcoming volume and this chapter intend to help clarify the issues of equality and equity in education and to invite further discussion and work on these complex questions.

2. Indicators are referred to here as statistics with a strong policy message. They describe attributes of populations or groups, using measures of central tendency and/or variation, or something similar (a mean, a rate, a proportion, a probability, or another statistic). In very large, diverse, and fragmented societies, statistics and indicators are becoming a major knowledge tool for citizens, policy makers and other stakeholders to describe the situation in a given field and assess the diversity and the trends. As Gary Orfield puts it "Virtually all theories of democratic governance include a very high value and strong protections for the discussion and debate that helps to dismiss error and discover truth and help to legitimise the imposition on the minority of the decision of the majority. The production of appropriate indicators on basic dimensions of the society, especially on issues of equality and equity, is of particular importance because the more stratified a society is, the less likely there is to be widespread personal contact and understanding of problems and the more important it is to have reliable data that shows relationships and trends." (EQR 2000)

3. On the other hand, the decision to collect statistics reflects a judgement now or in the past that the information collected is important (ibid.), and a decision about how it is to be defined, a conceptual point of view. Since 1988, within INES many such decisions have been taken, referring to a very broad conceptual framework that identifies inputs, processes and outputs of education systems in relation with their social and economic context. Recently, in close connection with the shift of education policies and management from an input to more of an output orientation, the question of outcomes has been given very particular attention and has culminated in the implementation of the PISA project.

4. Meanwhile, issues of equality and equity of opportunities and outcomes have frequently been declared of high policy relevance, and the INES General Assembly in Lahti considered them part of the top priorities for future work. However, these issues have not been treated as such within a comprehensive conceptual framework. Making sure that the most important aspects of equality and equity in education systems have been identified and are part of a common understanding is a first but major step in the process of agreement building among the various stakeholders.

75 The Steering Group of the OECD INES project decided in April 1998 to create an ad hoc group on equity issues. The focus of its mandate was to develop a conceptual and methodological framework for a system of equality and equity indicators in the field of education. This effort was supported by the Education Research Unit of the Canton of Geneva (Service de la recherche en éducation, SRED), in co-operation with the Swiss Authorities and by the NCES, Washington D.C. Starting in September 1998 the group held 4 meetings in Geneva, organised by Norberto Bottani, head of the SRED. About a dozen experts and researchers from 9 OECD member countries and from the OECD/CERI Secretariat participated. The results of this effort will be published as a separate volume by a private publisher. The volume will be referred to here as “EQR 2000”.
5. This chapter is devoted essentially to this first step, taking into account the results of recent research, as well as the major policy concerns. A first section will try to define the problem and show its policy relevance. The second section will look at past INES work in order to assess what has been done in the perspective of equality and equity and what remains to be done. The third section will present a first version of an organisational framework for a system of education equity indicators. Finally, the last section will submit proposals for further developments in this field.

Defining the problem

6. In modern societies all human beings are declared to be free, to have equal legal and political rights and equal dignity. This fundamental principle of democratic societies is embedded in the Constitutions of Western democracies since the 19th century and is part of the Universal Declaration of Human Rights adopted more than 50 years ago. The right to education is also proclaimed in this declaration which, although it is not enforceable, provides internationally accepted minimum standards. More recently, the Convention on the Rights of the Child acquired the status of international law (1990). Through this most widely ratified international rights instrument, states of signatory countries are bound to take responsibility for education provision.

Differences and disparities

7. The contribution by Vittoria Cavicchioni and Albert Motivans from UNESCO (EQR 2000) recalls these aspects of formal rights and the international standard-setting instruments that are being elaborated and discussed. The authors also show the considerable disparities between developed and developing countries as well as between different groups within developing countries. For example:

- Between 1970 and 2000, net enrolment ratios of primary-age children in primary schools have increased from 63 to 85 percent in less developed regions. Yet, some 90-100 million primary school-age children are not enrolled in school, representing at least 15 percent of all children in these regions (12 percent of boys and 19 percent of girls). (ibid.).

- Considering only the younger cohorts (age group 15-24), the overall literacy rate in less developed regions is about 85 percent. In least developed countries the rate is close to 64 percent. Gender disparities are declining, with a female/male literacy ratio of 0.91 as compared with 0.81 for the adult age group. Disparities between rural and urban areas vary from country to country but remain significant, as well as persistent illiteracy pockets in urban areas (ibid.).

- With regard to socio-economic stratification, a study for Latin American countries shows that by the end of the 1990s, there will have been a slight narrowing of the gaps across different groups. Nonetheless children from the poorest 25 percent of all households will still be at a considerable disadvantage in relation to the average and higher-income households (ibid.).

8. In the industrialised world, not only is education compulsory for 8-9 years in every child’s life, but the enforcement of this obligation is highly effective and most children would go to school even without this obligation. Disparities of educational opportunities, attainment and achievement, however, exist at compulsory level as well as post-compulsory secondary and tertiary levels. For example:

- Research repeatedly shows that children from different social origins achieve differing results and scholarly success in primary school and are selected into different tracks at lower secondary level and later (Hutmacher, 1993, Crahay, 1996)

- Overall school expectancy varies between 15 and 19 years among OECD member countries, with an average of 16.5 for men and 16.8 for women. For tertiary education, the range of variation is between 1.3 and 4.1 years (EAG, 1998).
The distribution of the adult population (25-64 years) by highest completed level of education varies considerably. In some OECD countries (Turkey, Portugal, Luxembourg, Spain) more than 70% of the population have left school after completing compulsory education, while in others (Czech Republic, Germany, Norway, United States) this proportion lies below 20%. Within countries, older generations and women show lower rates of completion of upper secondary education than younger generations and men. Generally, the lower the general level of education of the adult population, the larger the differences (EAG, 1998).

The recent international literacy study (IALS) showed that reading competencies of the adult population (16-64 years) vary significantly both across and within 13 participating countries. Average literacy scores vary from 224 (Poland) to 306 (Sweden), on a scale ranging from 0 to 500. Countries also differ in their patterns of dispersion around these average values. Apart from Poland that shows a particularly low level (85), literacy scores at the 5th percentile range from around 120 (German speaking Switzerland, USA) to more than 200 (Netherlands, Sweden). At the other end of the spectrum, the scores of the best performers are rather homogeneous; scores at the 95th percentile vary between 330 in Poland, 345 in Ireland and 387 in Sweden. The distance between these extremes—what might be called “the literacy divide” - ranges from 154 (Sweden, Netherlands) to about 245 (Canada, Poland, German speaking Switzerland, United States). The proportion of adults who have to cope with very severe functional illiteracy problems (literacy level 1) ranges from less than 10% (Sweden, Germany) to more than 20% (Ireland, United Kingdom, United States) and even 45% in Poland.

Empirical research consistently shows a positive relationship between educational attainment and earnings: better-educated persons earn higher incomes in all countries, but the return to higher educational attainment differs considerably. Among OECD countries, the earnings premium for persons with a university degree compared to persons who achieved “only” upper secondary education, ranges from about 35% (Denmark, Netherlands) to about 80% (Finland, France, Ireland, Portugal, New Zealand, United Kingdom, United States). In all countries, average earnings of women are consistently lower than those of men of the same age and educational attainment (EAG, 1998).

Inequalities - a matter of agreed upon criteria of preference

9. Across regions and countries and within them, among groups and individuals, there is wide variation in access to education, in progress and completion of education as well as in outcomes and consequences. Most of these differences are to be considered as inequalities because they are socially neither neutral nor indifferent; they make a difference in terms of widely agreed upon criteria of preference. Such preference criteria need not be universal throughout all human societies; it is sufficient if they are widely shared in a given society. In modern societies most people would agree that it is preferable to be wealthy rather than poor, to be healthy rather than ailing or sick, to occupy stable rather than precarious jobs, to be socially prestigious and influential rather than ignored or excluded, to share a wide network of influential social relationships and loyalties rather than to be isolated, etc.. Disparities in education will be defined as inequalities whenever the knowledge, competencies and credentials acquired through formal education are directly or indirectly associated with advantages or criteria such as the amount and nature of available material resources, power and authority, social prestige and influence, social capital, reward allocation, life conditions, risks and opportunities, life expectancy, etc. among individuals as well as, at an aggregate level, between regions and national societies. The general trend towards convergence of national societies that lies behind these observations should be underlined. National societies converge not only in terms of formally declared and adopted principles of individual freedom, equality and dignity, but also in terms of the major criteria and factors that determine inequalities of real social conditions and positions, as well as in the relationship between these inequalities and educational opportunities, outcomes and consequences. Despite large disparities between countries and with differing impact, factors like gender, wealth, socio-economic status, ethnicity, geographical region, etc., are at work all over the world and are considered important in relation to inequality and inequity.
10. Education is part of a public service under democratic control, and at the same time lies at the heart of a fundamental tension between formal civic and legal equality and inequality of real social conditions and positions. In all democratic countries, public education was established to give all children and youngsters equal access to and opportunity in education regardless of gender, place of residence or social, cultural, ethnic or economic background. The option for a "common school for all" is basic in modern democracies. And the goal to achieve "high quality education for all" has recently (1991) been stressed again by the conference of education ministers of the OECD countries.

11. Why and how do educational inequalities exist and persist in democratic societies? It is no surprise for sociologists, as Luciano Benadusi (EQR 2000) shows, that the main cleavages and inequalities in society are also more or less directly related to inequality in education. Clearly social inequalities existing outside of the education system contribute to educational inequalities in terms of access, opportunity, process and outcomes as well as in terms of the consequences of achievement and attainment. The impact of students’ gender on access and opportunities seems to be diminishing and indeed even reversing itself, a consequence in part of the feminist movement. On the other hand, the socio-economic status and wealth of students’ families, the level of education of their parents, their belonging to so-called majority or minority groups identified on the basis of ethnic, linguistic, racial, or other differences, still play a major role. As Benadusi also shows, sociologists have produced a number of answers to the question of how and why these social cleavages significantly determine the outcome of the education process. Although none of them can be considered conclusive, this research has 1) contributed to the general recognition of these factors as major determinants of the success of the educational process; 2) increased political awareness and relevance of the problem of educational inequality and 3) indicated tentative actions to counter the effects of social determinants.

12. A still not uncommon representation tends to consider that schools and teachers are (unfortunately) powerless to counterbalance the effects of pre-existing inequalities. This rather fatalistic interpretation places the locus of control outside of the education system; it excludes the schools and the teachers from the problem of inequality, but at the same time it prevents them from being part of the solution. There is now a rising awareness, however, in schools and in public opinion that the education system does not play a passive or neutral role in the process by which social, cultural and economic inequalities translate into scholastic and academic inequalities, which in turn translate into socio-economic inequalities. Increasingly the question rises about how schools and the education systems handle differences between students. Such differences are inevitable in a diversified, pluralistic and unequal society where conditions of life vary considerably. What do schools and education systems do about these differences and inequalities? Do they ignore, confirm, re-enforce or reduce them? And which systems do what, with what kind of effects and consequences? An improved set of international equality indicators should help to shed light on such issues.

13. It may be useful to recall very briefly some basic characteristics of education processes and systems. In the discourse and in the minds of many, education systems aim to transmit knowledge, skills, beliefs and values. This view is not wrong, but it is oversimplified.

– Firstly, education is project-oriented action. No educational relationship is symmetric, education implies the exercise of some form of authority and power by educators on the educated, be it within families or within schools. In education systems, the asymmetry expresses itself in the institutional structure where teachers, in the last resort, act on behalf and under the command of a political authority. Democratic control of this educational authority is no doubt a major factor contributing to its legitimacy, but this does not prevent it from being an authority. The asymmetry also expresses itself in the very choice of a definition of human excellence and of the rules of the game in formal educational settings, be such choices explicit or tacit. Education systems do not intend to transmit any or all available knowledge, skills, beliefs or values, but a selection that is considered worthy of specialised, organised and professional transmission under the more or less direct authority and control of the State. In segmented and pluralistic societies, the authorised and authoritative standard culture of reference which schools try to transmit includes components of the knowledge, beliefs, values, skills and dispositions that make up
cultures and subcultures existing among and within different groups. It inevitably excludes other components. However, the selection cannot be considered as a representative random sample of existing cultures or subcultures; it is a biased sample, reflecting the standards of human excellence that prevail in the dominating fractions of society, the ones precisely that set the standards, whatever they may be. In other words, no education is culturally neutral, and no education can be considered socially neutral either in families or in schools and formal education systems. The culture that is considered worthwhile at school may well be legitimised by democratic decision and it may appear desirable for various reasons to a majority. Arguments explaining the necessity of such choices are numerous: the sake of cultural and linguistic homogeneity within the country, of social integration and cohesion, of economic benefit and even of the democratic process itself. This does not change the fact that the culture of the school is a biased selection.

And inequality begins there. Different groups and families with different cultures will be positioned at more or less distance from the schools' reference culture, in terms of knowledge, beliefs, values and skills taught at school but more generally in terms of standards of human excellence, including the system of dispositions that guides individual perception, thinking, feeling, evaluation and action, i.e. of social habitus in Bourdieu’s (1980) terms. Due to the cultural bias some students will find at school a very familiar cultural environment with familiar representations, beliefs and expectations, if not in terms of knowledge, certainly in terms of rules and standards of behaviour, relationship and communication. Other students, with different experience and cultural background, will have to learn more and change or adapt more to the school’s cultural and social environment, because it is less familiar to them. Since the very existence of schools entails a division of educational work and authority between families and the school, such students may also have to cope with the potential tensions and contradictions between the cultures and identities valued by the familial educational authority and the school educational authority.

- Secondly, unlike newspapers, radio or television, the education system not only transmits knowledge, skills, beliefs or values. It also assesses and certifies the amount of students’ knowledge and competence, i.e. the distance that separates their performance from the goals to be attained within a given social setting and a limited time frame. Actually, assessment and certification heavily determine the rules of the scholarly game and constitute the fuel of the learning process for most students. But schools always assess more than they teach (Perrenoud, 1984). The assessment inevitably also bears on the distance between students’ attitudes and dispositions and the prevailing reference standards of human excellence, that is, partly on the distance between these standards and the family culture in terms of beliefs, representations, aspirations, dispositions or attitudes regarding school, authority, learning, peers, physical violence, etc.

- Thirdly, the impact of education systems, schools and teachers, in terms of actual transmission and learning is often seen as limited by the students’ unequal capacities and ability to learn. However, it may just as well be limited by the students’ inalienable liberty to construct their own interpretation of what happens to them and to draw their own consequences from the meaning they attach to their experience in their families and at school, in relation with subject matters, teachers, peers, etc. In simple words: students can be forced to some extent to go to school and sit quietly, they cannot be forced to prefer mathematics to football or simply to like school. Persuasion or seduction can help, but these are precisely means to influence their construction of meaning in relation with their willingness to learn. Such constructions of meaning will of course take place on the background of the students’ total experience, including the more or less intense influences and tensions within and between the social environments they are experiencing. On the other hand, this liberty of the students is the main reason why teachers and education systems, like doctors and hospitals, ultimately cannot be held accountable for results, but only for process and method. This does not mean that teachers, schools or education systems cannot be held accountable at all, but rather that their responsibility can only lie in the systematic and persevering quest for the processes, methods and practices which bring the best results, given the students they have and the meaning these students construct of their experience. In pluralistic and unequal societies precisely, students' backgrounds, prospects and meanings differ widely.
and one of the stakes of the equality debate is certainly the way in which schools handle such differences.\(^{77}\)

14. Recently, in most OECD countries, education policies have stressed the issues of efficiency and effectiveness, especially in terms of level of achievement, as the strong interest in the results of the IEA / TIMSS project and the launching of the PISA project underline. But governments of all OECD countries are also preoccupied by the problem of educational inequalities. Many have engaged in policy measures aiming at their reduction and ranging from structural measures such as delaying the moment of the first major options in the students’ career, to increasing resources, compensatory measures, bussing, positive discrimination, affirmative action, etc. As Bloom asserted: an efficient education system should raise mean achievement levels, reduce achievement variance and decrease the correlation between the student’s performance and social background. With its accent on “Quality education for all”, the OECD conference of education ministers (1991) clearly underlined that effectiveness and equality are two major policy themes which cannot be played one against the other.

15. Rather than expressing an egalitarian ideology as in the sixties and seventies, the quest for education equality nowadays seems to stem increasingly from a realistic policy appraisal. For instance, reducing the literacy gap between the most and the least educated while raising the average literacy level at the same time has become a major issue in realistic and effective education policies for the future. It appears indispensable for social cohesion and for democratic policy debate and decision in modern, fragmented, pluralistic and complex societies. It is also a major factor for competitiveness for all countries in the rapidly developing knowledge economy, where the willingness and ability of all to learn throughout the life span are expected to be among the major outcomes of initial education. There is wide agreement now about the duty of governments to give each person the necessary minimal basic skills for employability and further learning throughout life. And governments recognise this as being their duty towards all citizens. Reducing inequality of access to education, of learning opportunities and of outcomes pertains to the fundamental and major goals of education systems. In our current unequal societies, equality will remain a goal.

16. These are precisely the major arguments in favour of a comprehensive set of indicators on educational equality at international level. Educational equality and inequality are relative dimensions over time and space. An effort on equality indicators should help answer complex questions such as: How (un)equal are education systems of OECD countries when compared with each other and taking into account the existing social, cultural and economic inequalities? How does inequality of access, opportunities and outcomes evolve over time, given the increase or decline of other dimensions of social and economic inequality? What is the added value of schools and the education system in terms of reduction of inequality? How does inequality evolve in relation to changes in structural characteristics, institutional arrangements and policy measures, including those that are not geared directly at reducing inequalities? How do recent changes in management structures and competency devolution of education systems affect the equality issue? Is there a trade off between effectiveness and equality?

17. The potential benefits of statistics and indicators are illustrated by some evidence-based contributions to EQR 2000. For instance, analysis of TIMSS data suggest that there is no necessary trade-off between high level average achievement and equality; countries like Sweden, Belgium and Norway show both high effectiveness and low socio-economic status inequality levels (EQR 2000: Wildt-Persson & Rosengren; Crahay; Vandenberghe ). On the other hand, in relation with changes in management, observations reported by Wildt-Persson & Rosengren show that in Sweden, where the education system has been recently decentralised, the variance in resources and results between municipalities and schools is increasing. No final conclusions can be drawn from such observations at the present stage. However, they do illustrate the high potential value of internationally comparable equality indicators for policy debate and decision.

---

\(^{77}\) Bourdieu for instance (1966) suggested that schools and education systems produce inequality precisely through their indifference to differences, by treating all students equally, i.e. uniformly.
Equity – a matter of judgement on fairness with rising policy relevance

18. While equality is a question of balance or imbalance in terms of socio-economic advantage or disadvantage (access to resources, power, social status and prestige, education, etc.), the concept of equity focuses on the question of fairness or justice, which is a matter of values, judgement, feeling, opinion and debate. In the public discussion the two notions are often used as synonyms, equity being more often used in English, equality (égalité) more often in French. This confusion may signify either that inequalities are most often considered as inequitable, i.e. unjust or unfair, or that the debate most often focuses on inequalities that are considered unfair.

19. However, the question of fairness of inequalities is clearly of a specific nature. Justice only exists as a social definition commanded by the conditions and ways of life of human groups and societies. Fairness or equity cannot be observed as such nor measured; but it can be agreed or disagreed upon. A clear distinction between the concepts of equality and equity is important because it raises important questions especially for our purpose: For example, are all inequalities considered unfair or unjust? Are there differences between groups in this respect? What are the criteria of such judgements? But also, from a more ethical point of view, for what reasons should which inequalities be considered (in)equitable under which conditions?

20. Equity issues arise in all phases of the organised education process, as well as in outcomes and consequences, because inequality is observable in all of these issues. For example, the definition of the culture and of the standards of human excellence considered worthwhile at school are socially biased; the assessment of performance and mastery encompasses more than what is taught at school; some social dispositions and familial cultural heritages pay back better at school than others; some students have to learn and change more than others for the same rewards; students get unequal grades or show unequal performance; with regard to consequences, better education determines higher incomes, etc. But differences in rewards and social conditions are inequitable only if actors define and/or experience them as such. For instance, fundamentally, in modern societies, the positive relationship between level of education and level of incomes is generally considered as fair (although the fairness of prevailing differentials is under constant debate).

21. The policy relevance of the equity question is significant because equity judgements and feelings have a social reality and impact in and of themselves on the political stage, since they determine citizens’ judgement of policy, their political opinions and options. There is wide agreement that at individual level, human beings are different in many respects and therefore unequal. But there are many signs that inequalities in education related to social categories and groups (gender, SES, racial, ethnic or linguistic minorities) are increasingly considered unjust. Citizens tend to hold their government accountable for inequalities in society and in education, and these are increasingly inter-dependent.

22. Moreover, the relationship between the quantum of inequality and the quantum of inequity experience does not seem to be linear. As differences narrow, Tocqueville suggested, inequalities are more deeply experienced as injustices; feelings of injustice and envy will occupy a growing place in the hearts of men and render them more and more difficult to govern [cited in Meuret b), EQR 2000]. Recent developments in the field of gender inequalities seem to confirm this tendency. In many domains, gender inequalities have clearly diminished during the last 50 years. But the closer women’s condition comes to men’s, the less tolerable residual gender inequalities seem to become and the more active, imperative and comprehensive become protest and struggle against it.

23. Similar developments can be observed in the education field. In European countries, as long as access to upper secondary education was restrained to the top of the social hierarchy and open only to some particularly brilliant children from the lower classes, it was perceived as an instrument of social promotion and equality. Nowadays, nearly all youngsters are enrolled at this level of the education systems; only a small portion of youngsters do not enter it after completion of compulsory schooling and another small portion drops out during secondary education. By the age of 20, about 80% of a cohort have completed some form of upper secondary education. Yet, by now upper secondary education appears more
as an instrument of exclusion and inequality. Having become more generally accessible, it receives less recognition and results in more resentment (Dubet, 1992).

24. A reduction in inequalities may thus translate into an increase rather than a decline in frustration (Kellerhals et al., 1997). As a consequence, the perception of injustice, like envy, can engender violence and disruption of the social fabric (Dupuy, 1997).

25. In conclusion, the consequences of equity concerns for society become more significant and will interest policy more and more, as:

- the relationship between level of education and performance on the labour market becomes tighter;
- individuals’ stakes in education are growing;
- education is seen increasingly as a service which the State owes its citizens;
- the State itself increasingly recognises this service as one of its major responsibilities; it must therefore expect to be judged, more and more, on the equity with which it renders the service;
- equity judgements generate consequences at individual and social level, which increase the violence to be contained on one hand but make it important to allow for political expression and debate on the other (cited from Meuret b), EQR 2000).

26. It seems obvious now that a policy relevant system of indicators will have to distinguish equality and equity issues and to cover both; it will therefore have two main inter-dependent components:

- The first component will document the state and change of “objective” or structural inequalities and their relationships; it will make education systems accountable to public opinion, citizens, parents, students and other stakeholders; it will help citizens form their opinion and policy makers diagnose problems and define actions.

- The second component is geared at the more “subjective” or cultural aspects. It will inform policy makers, teachers and other personnel about the state and change, on the part of citizens, parents, students, teachers and other stakeholders, of the representations, attitudes, opinions and feelings on fairness related to education, about the prevailing criteria of justice and the judgements about the equity of the education system. Related indicators will certainly guide decision-making. But as information about the prevailing and confronting feelings and criteria they will also feed directly into the ethico-political controversy, debate and opinion building.

27. For the development of a conceptual framework regarding the second component, one option is to take stock of the conceptions of justice emerging from the current debate in political philosophy. This may surprise some readers, but there are several arguments in favour of this option. Firstly, indicators must be grounded in the contemporary debate about equity and social justice. Secondly, political philosophy contributes to a more profound analysis and to a more rational and balanced debate about these issues where opinion and polemics tend to dominate the scene. Thirdly, political philosophy will tend to broaden the scope beyond education systems as such and also to clarify the possible contradictions between the pursuit of equality and other major values like liberty, responsibility or effectiveness and efficiency. Fourthly, at a more instrumental level, theories of justice figure among the best sources to generate the questions to be highlighted by equity indicators. Denis Meuret’s contribution [a] EQR 2000] deals more in depth with the present debate and its transposition to education, drawing especially from John Rawls’ theory of justice (1971) and its focus on the least favoured fractions of society.

28. As they distribute a vital utility unequally, education systems cannot actually work without some explicit or implicit definition of basic principles of justice. They teach, assess, classify and select the
students and then certify their achievements on the basis of criteria and by resorting to more or less visible procedures, which teachers, students, parents and other stakeholders can observe and judge, and the (un)fairness of which they will find more or less acceptable. In other words, schools and education systems cannot operate without some shared convictions about which inequalities are equitable; this is simply one of the conditions of their relatively smooth day-to-day operation. Schools, principals and teachers not only inform stakeholders about their criteria and procedures for admission, assessment, selection and certification, they must also convince them and be themselves convinced of their fairness. Much of this “conviction work” is actually a result of schooling itself over generations. For instance, generations of students have learned and learn at school that rewards are allocated according to individuals’ performance, not according to their needs. They also learn that success of some and failure of others in a fair exam cannot be contested. For many, the school system represents a first significant contact with “failure” and “success”, with the standards that determine them and the prevailing explanations, causal attributions and loci of control. Such learning, however, does not result primarily from explicit teaching as is the case for a subject matter. It results from continued every day experience and interaction within education institutions. Assessment is particularly important in this respect, as the social relation through which students learn the prevailing standards and the equity rules and criteria. This learning is part of tacit knowledge, however, of which one is normally not aware without a special effort.

29. In a much broader sense, education systems also produce the conviction of the fairness of the unequal distribution of social positions and destinies in society. Alongside with the relative homogenisation of cultural references and standards, this is an essential outcome of the education process and a major contribution of education systems to the acceptability and legitimacy of inequality in society at large. Statistically, the level of education is more and more tightly related to income, prestige, authority, influence and power, as well as access to further education and life long learning. But education systems can only perform this fundamental legitimating contribution if their criteria, procedures and practice explicitly or implicitly enact a credible and shared “theory” of justice. Reversing this proposition, Denis Meuret (a, EQR, 2000) suggests that, although this may not always appear to the consciousness of the actors, the traditional discourse, organisation and practice of schools and education systems match best with the equity criteria of the traditional utilitarian and meritocratic justice theories. However, as society changes, the representations, opinions and legitimate criteria seem to be changing as well in this matter, as has been indicated before. This is another argument in favour of indicators that permit the documentation of the trends and the pace of the change in mentalities and sensitivities.

30. The current discussion on social justice interestingly confers more attention to education than the traditional utilitarian approach, signalling thereby that formal education and certification have become major resources and major stakes for individuals and for societies. The equity debate may therefore help clarify possible directions for indicator construction.

**Major principles for a system of indicators on educational equality and equity**

31. Summing up the ad hoc group’s debate and reflections, Meuret (EQR 2000) comes to the following conclusion:

- There is a renewed debate on social justice, but “no single theory has yet proposed a definitive response to the question of equity in education. It is therefore the terms of the debate that must inspire the nature of the variables to measure, in order to fuel it.” In other words, a set of equity indicators must allow debate within the framework of diverse existing principles of justice, not just one.

- The debate draws attention to some emerging new components in the equality realm. It is not enough for the indicator set to measure the independence of scholastic process and outcomes with respect to gender, social origin or other social categories. Increasingly important are measures such as the proportion of individuals situated below a minimum threshold of competence, gaps in knowledge between the least and best educated, and especially, indicators
to help understand the effect of equality in educational systems on the situation of adults, particularly for the least advantaged among them.

32. The present endeavour is based on a fourfold conviction:

- there is a genuine policy interest to confront the problem of educational inequality, for the sake of social justice and cohesion as well as in the interest of economic competitiveness;
- international and intersystem comparison makes it possible to learn from the experience gathered among countries with differing systemic, institutional and organisational arrangements; it is one of the most promising approaches to understanding the factors of inequality in education and the trends in the evolution of mentalities and sensitivities on questions of equity;
- research has produced enough evidence as well as relevant conceptual foundations and methodological know how (Benadusi, EQR, 2000) to lay the groundwork for the development of a comprehensive programme of equality and equity indicators that may contribute to such international learning;
- the development of such a programme and the analysis of its results will in turn stimulate new research efforts.

33. Taking stock of existing research on and experience with education indicators in general and equality and equity indicators in particular, and considering the high policy relevance of these issues, nine main principles were retained that should guide the choice and the construction of a system of equality and equity indicators in the education field. A systematic presentation of these principles is reported in Meuret [b), EQR, 2000]. Here they are only listed; the ordering should not be interpreted as a priority ranking:

- The system of indicators must measure inequalities to help citizens and policy makers to judge the equality of the system; but it must also identify, for the sake of those who govern, the citizens’ equity judgements on the education system and the criteria that form the basis for these judgements.
- The indicators must measure not only inequality in educational results (knowledge, school careers, social utility of academic degrees . . .) but also more immediate inequalities related to life in school and to the way students are treated by the institution and its agents.
- The indicators must permit debate within the framework of diverse existing principles of justice, and not restrict themselves to only one of them.
- The educational inequalities relevant to most of the goods distributed in the framework of educational systems can be assembled into three large families: 1) inequalities among groups and categories, 2) disparities between individuals and 3) the proportion of individuals falling below a minimum threshold. All three must be taken into account.
- It is important to measure not only inequalities in educational outcomes, but also the inequalities upstream of the educational system and those which affect the teaching-learning process itself.
- Among the goods distributed by the educational system, priority must be given to those whose distribution is most important to individuals or for the democratic life of the country.
- From the equity point of view, among the groups and categories to which individuals belong, the most important are those from which they cannot escape.
- Indicators must permit discussion of the balance between the concern for equality and the concern for other values to which it may be opposed.
- Indicators must bear not only on educational inequalities, but also on the consequences of these on social inequalities in general.
34. The issue of equality and equity is a very complex one, even from a purely descriptive-analytical point of view, because of the number of interdependent factors intervening in a systemic context-input-process-output approach to education systems. The issue is also very controversial, because it inevitably has to do with values. Moreover, both the inequality factors and the equity criteria may differ between countries and change over time. A conceptual framework for internationally comparable equality and equity indicators should guide the construction of indicators to help describe the state and the change of social cleavages that structure modern societies and their education systems on one hand, and of the ethical background that informs value judgements about observed disparities on the other.

35. The core message of this report is that this necessary step is possible, as will be shown later.

**Recent results in INES**

36. The OECD education indicators project (INES) has already produced a number of indicators of educational equality and disparity since its inception in the late 1980s. The Policy Review and advisory group (PRAG)\(^78\) had recommended in 1995 that equality issues be taken into account whenever feasible, and that whenever relevant, statistics and indicators produced within INES should show breakdowns by gender and age as well as any other relevant social cleavages. For gender and age, this has become more or less regular practice since, mainly on the grounds of data availability. This section is a slightly adapted summary excerpt from the survey of existing data and indicators which Tom Healy and David Istance present in their contribution to the forthcoming volume (EQR 2000). Since 1992 various editions of *Education at a Glance, OECD Indicators* (EAG) have shown indicators related to differences and inequalities and which can be broadly categorised in two ways:

- Indicators of dispersion among individuals (intra-group);
- Indicators of inter-group disparities and inequalities (age, gender, educational attainment of parents, occupation, spatial units, special learning needs, etc.).

37. This distinction is relevant from the point of view of equity as defined earlier. Differences and inequalities among individuals can be seen and are often seen as reflecting the inevitable inter-individual variability of ability, motivation, physical condition, or other personal characteristics. They may have important consequences, but they are not necessarily considered unjust because of the inevitability of inter-individual variability. According to this perspective, equity would require, however, that such differences and inequalities be not related to any factors that would determine a discrimination between social groups or categories (gender, age, type of family background, social, linguistic, ethnic or racial origin, etc.), that is, evidence for equity and equality of opportunity would be found when individual differences are statistically independent of belonging to one of these groups. Individual differences and inequalities should ultimately be considered only as residual intra-group differences, once all relevant social factors are controlled in a multivariate approach.

**Inter-individual (intra-group) measures of dispersion**

38. At the level of global indicators, this first category typically covers measures of dispersion or variation for skills of individual students or adults. Examples include measures of disparity among students in achievement in mathematics and science (based on results of the *Third International Mathematics and Science Survey*, TIMSS)\(^79\). Other examples relate to disparities in adult literacy based on the OECD *International Adult Literacy Survey* (IALS). As has already been mentioned, countries vary considerably in the degree of disparity in literacy and student achievement. The comparison of results for

---

\(^78\) Former Steering Group of INES

\(^79\) TIMSS was conducted in 38 countries in 1995 by the International Association for the Evaluation of Educational Achievement (IEA).
different age groups (or grade-years in the case of the student population) also suggests varying patterns of progression (and attrition) in skill acquisition across countries.

**Inter-group or inter-category measures of inequality**

39. Whilst age and gender have become regular categories in data and indicators within INES, other relevant categories or groups (spatial, socio-economic, ethnic or special learning needs) have been taken into account on an occasional basis.

**Age**

40. Differences across age groups in levels of educational attainment (as measured by the highest completed level of initial education) have been well documented in EAG for the adult population. The rapid evolution in educational attainment in recent decades has initially opened up a sharp divergence in levels of attainment across different age groups. However, as higher levels of graduation spread to successive generations, such age gaps tend to narrow, except for “late starting” countries. Published data on attainment are available for the 1990s, but there are few internationally comparable data on attainment going back before this, except proxy measures comparing generations at a given time.

41. Comparing educational attainment of adults with their parents’ makes it possible to measure the mobility of attainment across generations. The International Adult Literacy Survey provided data for such inter-generational comparisons (e.g. the probability of having studied at the tertiary level, based on the level of attainment of one’s parents). The contribution by Noël and de Broucker to the forthcoming volume presents a comparative analysis of the influence of parents’ educational background on length of schooling and literacy skills (EQR 2000).

**Gender**

42. Differences in access, participation, graduation and literacy or achievement levels between boys and girls, men and women, have become a regular focus of analysis. Breakdown by gender of many indicators has been a regular practice in EAG. Despite the recent narrowing of the gender gap in terms of school retention and graduation, significant gaps persist between men and women in rates of labour market participation and in earnings. Differences in choice of fields of study also remain, but cannot be interpreted in terms of inequality or inequity without further examination. The same observation may hold in relation to labour market participation, which may also reflect individual choices and value judgements between family and work.

**Regional or spatial subdivisions**

43. National averages often hide considerable sub-national variations. In many countries, responsibility for education policy is delegated to sub-national authorities. Therefore differences by geographical sub-units may well be of interest, e.g. for funding, school processes, outcomes, etc., also in relation with other characteristics of sub-national units. The main challenge encountered in making such comparisons across spatial units is to define the appropriate and internationally comparable sub-national units. Indicators have been published occasionally on differences among sub-national jurisdictions within countries (states, Länder, provinces, cantons, counties, local school districts, académies etc.).

44. An increasingly relevant approach to spatial differentiation will be to compare metropolitan, urban and rural areas, for example, with respect to in conditions of teaching, resources, class size, access to initial and to adult education etc. The contribution by Doug Cochrane to the forthcoming volume (EQR 2000) will illustrate this perspective in more depth with data for the United States.
Socio-economic background

45. Socio-economic status (SES) may be defined as the relative position of a family or individual in an hierarchical social structure, based on their access to, or control over, wealth, prestige, and power. It is indexed by occupation, since in modern societies, most important aspects of individual and family life, such as income, social capital, life conditions, risks and opportunities, social status, prestige and influence, as well as access to information, further education and other valuable cultural goods, health and life expectancy, etc. are more or less directly related to the positions individuals and/or relatives hold in the field of work. Research consistently underlines the significant influence of the socio-economic status of parents and the quality of the home environment on their children’s access to education, expectations and opportunities, careers and learning outcomes. The link between socio-economic background and educational attainment and achievement can henceforth be considered as a fact, although its strength varies across countries and regions. Results also show that disadvantages at the entrance to formal education due to social background are exacerbated by the end.

46. Differences by socio-economic status remain a major focus of attention in the education field. The development of related statistics and indicators should be a priority in the future, also to permit the sorting out of the specific influences of other social cleavages that are often correlated with SES (income, level of education, racial or ethnic minority, etc.) through a multivariate approach to inequalities. But arriving at a standard for comparing different social classes or socio-economic groups at the international level is not an easy task. Usually, such comparisons require the specification of an international index of socio-economic status based on occupational data. Frequently, recourse is made to data exclusively on income or on the highest level of educational attainment as proxies for socio-economic status.

47. Some recent key indicators in this area include: student achievement in mathematics in relation to educational resources (whether students had a dictionary, a study desk or a computer at home) using IEA TIMSS data; student achievement in mathematics at 8th grade in relation to educational attainment of parents using IEA TIMSS data; participation in job-related continuing education and training in the previous year by employed adults, by highest level of educational attainment and IALS literacy level, 1994–95

Students with special learning needs

48. There is considerable international variability in conceptions and categories describing students with special educational needs as well in the scope of special education. Recent work in OECD/CERI shows that the term “special education” is used in some countries only to refer to disabled students with sensory, physical or cognitive impairment, whilst in others it also includes the disadvantaged (e.g. immigrant or ethnic minorities) and even highly gifted students. Hopefully, this work may contribute to some international conceptual convergence by focusing on the needs of these groups and on the amount and nature of additional educational resources to meet them. The contribution by Peter Evans to the forthcoming volume deals more in depth with these questions and also discusses the equity issues raised by the unequal allocation of resources to students with special needs (EQR 2000).

49. Meanwhile, some comparisons have been shown in EAG with respect to groups of such students.

80 Such a composite index results (International Socio-Economic Index) is proposed by Ganzeboom, de Graaf, and Treiman, (1992). It is based on components such as occupational prestige, educational attainment and/or income to assign an SES score to each occupational title. Knowing a person’s occupation then allows the assignment of an SES score to the individual even in the absence of data on income. This is the basic approach adopted in the OECD Program for International Student Assessment (PISA) project.
Ethnic or minority- language groups

50. The growing cultural and ethnic diversity across OECD societies raises important issues for education policy makers concerned about the integration of different groups and the provision of opportunities for learning and advancement for those who may be at a disadvantage. It remains difficult to arrive at comparable definitions of racial, ethnic or immigrant minority groups. Indicators that have been published in the past are mainly driven by data availability, depending on international surveys. They compare citizens of the country with non citizens, native with non-native language speakers, persons whose parents were born in the country and whose parents were immigrants.
### TABLE 1: AN OVERVIEW OF EQUALITY INDICATORS IN EDUCATION AT A GLANCE

<table>
<thead>
<tr>
<th>Social and economic context</th>
<th>Gender</th>
<th>Age</th>
<th>Region / spatial</th>
<th>Special learning need</th>
<th>Socio-economic status (SES)</th>
<th>Ethnic / linguistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of educational attainment in the adult population</td>
<td>Regular</td>
<td>Regular</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor force participation rate by educational attainment</td>
<td>Regular</td>
<td>Regular</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### FINANCE, TEACHING RESOURCES

| | | | | | |
| Spending on education as proportion of GDP | | | | | |

| Expenditure per student at primary and secondary levels | EAG1996 | EAG2000 | (under special learning needs) | (under special needs) | |
| Teaching and other personnel employed in educational institutions | Regular | Regular | | | |
| Eligibility for public student aid at the tertiary level | | | EAG2000 | | |
| Ratio of students to teaching staff | EAG1996 | EAG1998 | (under special learning needs) | (under special needs) | |

#### ACCESS TO EDUCATION AND TRAINING

| | | | |
| Rates of enrolment in various levels of primary and secondary level education | Regular | Regular | EAG1998 | |
| Expected educational attainment (years) | Regular | | | |
| Access to early childhood education and care | Regular | Regular | | |
| Access to tertiary education | Regular | Regular | | |
| Participation in continuing education and training | Regular | Regular | | |

#### SCHOOL AND LEARNING ENVIRONMENT

| | | | |
| Instruction time at primary and secondary level | | | | |
| Teacher salaries | Regular | | |
| Grouping of students by ability at lower secondary level | | | | |
| Level of differentiation by school and program type | | | | |
| Grade repetition at primary level | | | | |
| Participation by parents in school activities by parents’ educational attainment level | | | | |

#### GRADUATION RATES

| | | | |
| Upper-secondary graduation rates | Regular | Regular | | |
| Upper-secondary graduation rates -- vocational/general | Regular | Regular | | |
| Tertiary graduation rates | Regular | Regular | | |
| Tertiary graduation rates by field of study/level | Regular | Regular | | |

#### STUDENT LEARNING OUTCOMES AND ADULT LITERACY

| | | |
| Student achievement in mathematics and science | EAG1998 | | |
| Progression in student achievement between 7th and 8th grades | EAG1998 | | |
| Attitudes towards science among 4th graders | EAG2000 | | |
| Prose, document and quantitative literacy of adults | | | |

#### LABOR MARKET OUTCOMES

| | | | |
| Rate of unemployment by level of initial education (adult population) | Regular | Regular | | |
| Earnings by level of initial education | Regular | Regular | | |
| Rate of unemployment of recent upper-secondary and tertiary graduates | Regular | Regular | | |

Source: Healy & Istance, EQR 2000

- Feasible and desirable new area
- Not applicable, not feasible or not necessary in the near future

**Regular** indicates presentation in EAG every year, or in most years since the first edition in 1992. EAG2000 or EAG1996 denotes the latest year for which data were supplied with respect to a given category (data may not be available for other years)
51. Table 1 summarises the development of indicators in the various areas and signals possible and desirable areas for further development. It is not exhaustive but it can stimulate reflection on possible priority areas.

52. This overview shows that equality is already an issue within INES work as reflected by EAG. At least some data and indicators have been provided for all the main relevant social cleavages as reflected by the columns of Table 1, with the exception of socio-economic status (SES), for which data will be available for the first time through PISA.

53. Data availability however seems to be the main driving force, more than a systematic approach to equality and equity issues. This is also visible considering the significant gaps in the different domains of education indicators as reflected in the lines of the table.

54. Moreover, neither data nor indicators have been provided until now on the issue of equity judgements as defined in section A.

55. Finally, the need for international trend data as well as cohort data based on longitudinal data sets cannot be overemphasised. Such data permit countries to benchmark their own performance and progress towards more equal and equitable outcomes in education against “international best practices” or to describe patterns of inequality over time -- a useful diagnostic in national policy evaluations.

Towards an organisational framework for equity indicators

Taking into account existing debate and expertise as well as results from research and recent indicator work the ad hoc group discussed the organisational framework that might guide a system of equality and equity indicators. The following presentation is a large excerpt from a paper written by Denis Meuret (EQR b) 2000) which reflects these discussions and the consensus that has been reached.

General outline of a system of indicators on equity in education

56. This outline is presented as a basis for further deliberation on an appropriate system of indicators. It rests on yet imprecise knowledge of all the factors and consequences of educational inequalities. Multiple conceptual and technical problems remain to be resolved, without forgetting the need for a discussion and a political consultation for a system of this type. The outline presented here may also be improved by a more systematic analysis of studies and available data than the ad hoc group could accomplish. On the other hand, once agreement has been reached on the general principles, the establishment and regular calculation of appropriate indicators should start without waiting for consensus on every detailed indicator. The availability of new data and indicators will further our understanding of the causes and consequences of scholastic inequalities, and lead to a more refined discussion of equity. It will in any case improve on the present state of analysis, where inequalities are studied in too disparate a fashion among countries for any conclusion about equity of education systems to be drawn.

57. Some preliminary remarks may facilitate understanding:

1. Some of the proposed indicators already figure in international systems of indicators, for example, in Education at a Glance. Others are found in specific studies for international comparisons; still others have been calculated within the framework of national studies but seem relevant in an international perspective; and finally, there are others which do not yet exist, to our knowledge, except in the planning stage.81

2. Depending on the principle of justice that guides the reader, the relation to an equity indicator may change. At the same time, *grosso modo*, the outline is designed to be read in

81 Complete references to the different sources will be given in Meuret b) EQR 2000
the following manner: inequalities affecting internal processes and results will signal an inequity of the educational system to the extent that: their consequences on the future life of students are significant (external results); they can be attributed to the operation of the educational system (process) rather than to social inequalities themselves (social and cultural context); they strongly affect the opinion which citizens or users hold about the justice of the system (social and political context).

3. Inequalities of internal results are to be measured by three types of indicators: inequalities among groups and categories, disparities between individuals, proportion of individuals falling below a minimum threshold. It is useful therefore to echo these types for the context and process indicators. This will permit, for example, the linking of inequalities in scholastic success related to social origin (internal results) with differences in income or in diplomas among social categories (context). Similarly, it should be possible to relate individual inequalities in school careers to inter-individual inequalities in income, or parents’ education, or to link the proportion of individuals whose skills are below a certain threshold to the proportion of the population who live below the poverty level.

4. Since some of the proposed indicators may be unduly influenced by the specificity of some national educational systems, it may be necessary to design international equivalents.

58. The tentative outline is presented in its skeleton form in table 2.

Table 2: General outline of a system of indicators on equity in education

<table>
<thead>
<tr>
<th>1. Context</th>
<th>1.1 Social and cultural context</th>
<th>1.1.1 Inequalities in social resources</th>
<th>1.1.2 Inequalities of cultural resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 Political context</td>
<td>1.2.1 Equity criteria</td>
<td>1.2.2 Judgements about equity of the system</td>
<td></td>
</tr>
<tr>
<td>2. Process</td>
<td>2.1 Quantity of education received</td>
<td>2.1.1 Length of schooling</td>
<td>2.1.2 pending for education</td>
</tr>
<tr>
<td>2.2 Quality of education received</td>
<td>2.2.1 Inequalities in conditions of learning</td>
<td>2.2.2 Inequalities in quality of life</td>
<td></td>
</tr>
<tr>
<td>3. Internal results</td>
<td>3.1 Inequalities among individuals</td>
<td>3.1.1 Disparities in competencies</td>
<td>3.1.2 Proportion below the justice threshold</td>
</tr>
<tr>
<td>3.2 Inequalities among categories</td>
<td>3.2.1 Social origin (SES)</td>
<td>3.2.2 Educational attainment of parents</td>
<td>3.2.3 Gender</td>
</tr>
<tr>
<td></td>
<td>3.2.4 Racial, ethnic minorities and majority</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. External results</td>
<td>4.1 Individual consequences of educational inequalities</td>
<td>4.1.1 Economic and social consequences</td>
<td>4.1.2 Non-monetary consequences</td>
</tr>
<tr>
<td>4.2 Collective consequences of educational inequalities</td>
<td>4.2.1 Inequalities to the advantage of everyone or of the least favoured</td>
<td>4.2.2 Institutional consequences</td>
<td></td>
</tr>
</tbody>
</table>
1. Context

59. The social and cultural context, on one the hand, and the political context, on the other, should be distinguished.

   - The social and cultural context matters to explain inequalities of internal results, but also to compare systems’ equity among countries. For example, if two education systems present the same inequalities in results, the more equitable of the two is in the country which presents the highest social inequalities.

   - The political context is measured to inform decision-makers and sharpen their awareness of the criteria by which the public judges equity in the educational system, and of the opinions people hold on its equity. These indicators can serve to rank the different inequalities measured from the political vantage point, in particular those which more directly measure the equity of the system (external results, process).

Social and cultural context

60. There is ample evidence of the relationship of social and cultural inequalities to inequalities in school. Benadusi’s synthesis ([a], in EQR 2000) can inform and guide the choice of significant indicators of social inequalities. Examples of factors which contribute to scholastic success are: expectations (Bourdieu), mode of expression (Bernstein), income (Boudon), cultural capital (Bourdieu) and social capital (Coleman). The more unequal the distribution of these goods (among social categories or among individuals) in a given country, the stronger the influence of these factors on inequalities in school results, the longer the road will be for that country to achieve equality of educational opportunities.

61. It is not certain, however that the relationship between social and scholastic inequalities will be linear. There may be threshold effects: for instance, beyond a certain income and material security, differences may be weak.

62. Two large categories of resources should be further differentiated:

   - inequalities in the amount of social resources (e.g. income inequalities before and after social transfers among social categories and between individuals, proportion of individuals below the poverty threshold) and inequalities in security of their availability over time (life expectancy, unemployment rate, etc).

   - inequalities in cultural resources include education (e.g. education level of parents, adult literacy, proportion of adults below the minimum threshold of literacy) but, in the anthropologic sense of culture, also norms and expectations, social capital, representations and attitudes in relation with education and educability, etc.

Political context

63. Unlike the indicators just discussed, these serve less to explain the inequalities of results than to interpret them from the political point of view.

   - Equity criteria help to appreciate which inequalities are the most important pointers of equity in the education system in the minds of citizens, school personnel, students, parents, etc. Surveys can provide valuable data about the distribution of different principles of justice among and between such populations, about the responses to moral dilemmas or about the nature of inequalities that people find most disturbing.

   - Judgements bear principally on the equity of the system, but may bear also on the manner in which the concern for equity is considered (or not) to be in conflict with concerns about other
values. This type of indicator is also intended to measure the political effect over time of actions undertaken to improve equity. These dimensions should be captured through an international survey.

2. Process

64. Measuring equity in the educational process will help to understand better the origin of inequalities in internal results. It is important also because numerous inequalities affecting the educational process – inequalities in treatment, well-being, for example – can be inequitable by themselves, independent of their potential relation with inequalities of results. To apprehend inequalities in the scholastic process, it is useful to distinguish between quantity (2.1) and quality (2.2)

**Quantity of education received**

65. For individuals, one can judge the quantity by the length of schooling received and by its cost, the latter also depending on length; furthermore the intensity of the instruction might be captured through variables like class size, available material, etc.

- **Length of schooling**: No author, to our knowledge, has called for equality among individuals in the length of schooling, but some do propose such an equality in terms of training over the life span. They find it inequitable that the less apt receive less schooling, that is to say fewer resources to develop their abilities, especially because most students who quit school do so because they are forced to. Authors who accept drop-outs argue that these students would not learn anything if they stayed on. Debates of this kind are frequent and they are of high relevance to policy makers. They could be served by indicators such as: inequalities in length of early schooling according to social origin (Noel and de Broucker, EQR, 2000), inequalities accumulated in early and continuing education and training, proportion of students who leave school early against their will.

- **Spending for education**: The issue here is the distribution of public spending for education among different social categories. To what extent is public spending for education redistributive? What proportion of their income must families of differing social categories spend for the education of their children, at different scholastic levels? Do various social categories benefit equally from public investment in education? Are there geographic inequalities in annual cost per student? In countries where financing is centralised, the geographic financial inequalities of education are small. When financing is decentralised, as in the USA (Cochrane, EQR 2000), inequalities can be very pronounced. There is a debate on the consequences of such inequalities on results. It is therefore necessary to measure them.

**Quality of the education received**

66. On the side of quality, one can distinguish two aspects also: inequalities in conditions of learning (2.2.1) which cover everything which, over a given period, enables a student to learn or develop; and inequality in quality of life (2.2.2.) which bears on the scholastic experience. The importance of the quality of life in school stems from the fact that education is not only an investment for the future, but also an immediate experience that occupies about one-fourth of an individual's life. The way in which it is experienced, and in which one feels treated by others at school, should therefore be considered as a good in itself, the distribution of which must also be equitable.

- **Inequalities with respect to conditions of learning**: Education is considered here as an individual investment. Indicators intend to verify if the resources which the education system puts at the disposition of investors – students - are of a quality corresponding to what is required by fairness. Numerous studies have shown that equality depends not only on the quantity of resources, but also on the quality of schools and of teachers. There is also a relationship between
the quality of instruction and the quality of students enrolled: good students bring with them higher expectations, better discipline and thus better conditions for learning. In Vandenberghe’s contribution to EQR 2000, the “peer effect” appears to have a direct origin (influence of students on one another) and an indirect one through the interactions between teachers and students.

Many studies also show a strong link between scholastic segregation and inequality of teaching-learning conditions. The absence of segregation, be it scholastic (good students are in the same classes and schools) or social (students from low-income families are in the same classes and schools) appears as one of the best guarantees that all children are placed in equal conditions for learning.

- **Inequalities with respect to quality of life:** Jencks (1972) in his famous inquiry into inequalities in education, concluded that, given the time children spend in school, equalising their well-being in different schools would do more for the equality of conditions in America, than any other effort to equalise the effectiveness of schools. Inequalities of well-being cannot all be considered as unfair, but it is reasonable to think that the responsibility for proving that they are unavoidable or necessary to learning belongs to the educational system. Two issues are to be distinguished:
  - **Inequalities in well-being** can be measured by surveys, some of which may have international validity, such as: gaps in well-being at school between categories (girls vs. boys, wealthy vs. poor); gaps in self-esteem, perceived quality of schools and teachers, perceived legitimacy of the causes of inequalities in well-being; inequalities in well-being among schools; inequalities in feelings of insecurity, etc.
  - One of the principal factors of well-being of students certainly is the quality of their relationship with their teachers, in particular the feeling of being treated fairly. Such feelings can be measured through surveys asking students about the degree to which they feel they are treated fairly themselves and to which students in general are treated fairly or on the contrary the degree to which some are favoured in every day school-life or in assessment, grading, discipline, distribution of rewards and punishments, advancement in degrees, etc.

**Internal Results**

67. Results are defined as the effects of the education system's action on the dimensions of its direct goals: knowledge, skills, attitudes and diplomas that attest to students’ mastery. Indicators measuring these results are separated into individual inequalities and inequalities by category. Some argue that only inequalities between groups or categories are inequitable. This position however flows from the meritocratic conception of justice. Renouncing inequalities among individuals would be contrary to the project of treating equity from the point of view of several principles of justice.

68. When measuring results it is most relevant to confront equity to effectiveness. Vandenberghe, (EQR, 2000) shows, that in TIMSS there is no relationship between the effectiveness of the education system in mathematics achievement and the social origin of pupils. In other words, there is no trade-off between equality and effectiveness. There may even be a favourable relationship, if one notes that, among the four most effective countries in mathematics achievement considered by this author, three are also among the least unequal.

**Inequalities among individuals**

69. Two types of measurements for inter-individual equality should be considered: the dispersion and the proportion of individuals situated below a threshold.

- **Disparities and gaps:** A good deal of data concerning inter-individual differences is available in IEA studies and more will be provided by the PISA project. Disparity measures are provided in the form of: standard deviation, gap between strong and weak students; inter-quartile range, etc. The weakness of the studies used for measuring disparities lies in their limitation to academic
subjects. The development of appropriate measures of competencies, such as cross-curricular competencies, or non-cognitive or meta-cognitive competencies, will be a major challenge for the future in this domain.

- The proportion of individuals below the justice threshold is particularly important from the point of view of social exclusion. Examples of measures to develop in this domain could be: the proportion of students whose reading skills are below a standard deviation from the international average (EAG 1993) or the proportion of students leaving compulsory schooling without mastering the basic skills established as necessary to life in society.

**Inequalities among categories**

70. Social inequalities fall into multiple categories. There is agreement, however, to consider that most social hierarchies in which individuals are situated (income, prestige, power) are organised around their professional situation (Levy et al, 1997). Indicators should in any case retain those categories from which one cannot escape: gender and nationality or country of birth or ethnicity and social origin.

- **Social origin (SES)**: This remains the major category in relation with equality of opportunity and academic success of the children, since work and professional status are the major determinants of social position and condition in modern societies.

- **The educational attainment of the parents** has often been used as a proxy for SES and sometimes has shown a stronger influence on internal results of the education system than SES. Its influence on the school career, the skills at the end of compulsory education of children, as well as on adult literacy has been frequently underlined.

- **Inequality between genders**: Education careers of girls are now much closer to those of boys; girls often surpass them. However, TIMSS has recently shown unequal competencies in mathematics. Differences remain also in terms of the careers to which girls and women have access (higher education in science, top levels of the educational system).

- **Inequality between minorities and the majority**: The very definition of minorities differs from one country to the other, and this is a major methodological difficulty. These inequalities generally result from relations of domination and rejection to which specific groups are or have been subjected. Many factors enter into play, among them their history in the country, immigration, the relationship with other social categories (Benadusi, Orfield, EQR 2000). In Europe, social status being held constant, children of immigrants do as well as native children, but such is not the case in the United States between blacks, hispanic and whites.

71. Gender is relatively independent from the other categories, whilst all the others are correlated among themselves to a degree that may vary across countries and over time.

**4. External results**

72. The greater the negative social consequences of educational inequalities, the more unjust they are. Such consequences are of two types:

**Individual consequences of educational inequalities**

73. Individual consequences concern the relationship between the position in the educational hierarchy, on the one hand, and the position in the social hierarchy, income and the possession of other non-monetary goods or advantages, on the other. Certain skills or knowledge related to these effects are supposed to be acquired at school, the mastery of which is not guaranteed even by a lengthy education (even basic literacy).
Economic and social consequences: The relationship between education achievement or attainment and income or unemployment is well known, as well as the rate of return on education by level of schooling and gender (EAG, 1997, 1998).

Non-monetary consequences: Here the relationship between level of education and risk of imprisonment, life expectancy, situation in the marriage market, adult literacy skills (EAG, 1998), or access to further training, etc. come into play.

Collective consequences of educational inequalities

74. Collective consequences are of two types. The first concerns the extent to which educational inequalities affect the least privileged. One could call these inter-individual externalities. The second addresses the consequences of education and education inequality on the community as a whole. For example, if confidence in institutions is affected by educational injustice, all can experience the political consequences. One could call these global or institutional externalities.

Examples of inter-individual consequences are: income gap between the least and most educated, before and after taking into account the effects of financial transfers from public or other origins; the distribution of “professional resources” (doctors, teachers, lawyers, academics, etc) and of resources for local economic and social development, between poor and rich areas of a country; inequalities in the use of professional services between social categories; the rate of fiscal return on education, etc.

Examples of global and institutional consequences are: delinquency and crime rates, civic participation, confidence in the education system, in institutions in general, in public services, professionals; degree of violence towards institutions and their representatives; feeling of belonging to the society or the community which delivered the education, etc.

Conclusion and Recommendations

75. In all OECD member countries, social, cultural and economic inequality is more and more intimately connected with education. As countries evolve towards knowledge-based economies and societies, inequality also becomes more and more knowledge and competence-based. There is ample evidence for the policy relevance of the issues of educational inequality in highly differentiated and pluralistic democratic societies. The development of a comprehensive set of indicators in this domain therefore appears as an indispensable component in the process of building the knowledge base for education policy. Such a set should allow the measurement and comparison, across countries and over time, of the degree of inequality in access to educational resources, in progress and completion of education and in the outcomes and consequences of education in relation with prevailing social, cultural and economic inequalities. It should help the policy process in assessing and monitoring both the degree of inequalities and the degree to which they are judged (in)equitable in a more systematic fashion.

76. The issue of equality has not been ignored in recent work on education indicators, and in particular within INES. This is, however, the first systematic attempt to clarify the concepts involved, to formulate the guiding principles and some methodological guidelines for the development of a coherent system of indicators on the issues of equality and equity. Both are highly controversial because they inevitably raise questions of values, attitudes, conflicting judgements and opinions.

77. The most original contribution of the ad hoc group’s work is probably, at the conceptual level, the proposal to clearly distinguish between issues of equality and of equity. The proposed system of indicators is intended to reflect both the observable “objective” inequalities and the observable, more “subjective” representations, attitudes and value judgements concerning their equity and/or acceptability in the minds of citizens, parents, students, teachers and other stakeholders. This latter aspect is of particular relevance to policy making in democratic societies since it is aimed at informing the policy
process about the state and the change in opinions, sensitivities and options regarding the equity of prevailing or changing arrangements in the education system.

78. The ad hoc group is aware that the development of a coherent and comprehensive system of indicators on equality and equity issues in education, indispensable as it is, will be a very demanding long-term task. Development should therefore start immediately. Some lines of this work can be sketched out on the basis of the ad hoc group’s experience. They are not meant to be exhaustive, but they give an idea of the range of things that need to be done.

- The ad hoc group agrees that a conceptual and organisational framework is necessary for developing in a coherent way a set of indicators on education equity. The framework presented here and the conceptual and empirical contributions contained in the forthcoming volume (EQR 2000) aim to encourage debate on this issue. The theoretical and methodological groundwork that has been accomplished will need further elucidation, discussion, consolidation and deepening. Clearly also, it does not cover everything and to some extent reflects the composition and interests of the ad hoc group. For instance, further work will need to extend the scope from initial education to lifelong learning, keeping in mind their connections, precisely from the viewpoint of equality.

- The ad hoc group considers that clear policy support is needed to ensure a rapid and sustainable development and production of equality and equity data and indicators.

- The ad hoc group underlines the fact that equity issues are a permanent component of the policy agenda in most OECD countries. There exists, however, no exhaustive overview of the different ways these issues are addressed by policy across countries, of the measures implemented and their results, or of recent research projects and results or relevant data sources and data sets. The OECD secretariat could sustain the effort of constructing such an inventory of policy measures, research and data availability to the benefit of the knowledge base of all countries.

- The ad hoc group draws attention to statistical data already available at OECD which includes a great deal of information necessary to calculate some of the proposed equality indicators. Further development work along the lines of the present conceptual framework should be based on a systematic assessment of existing data.

- The ad hoc group recommends that a major effort be made to establish an international standard classification of social categories or groups to be distinguished for equality and equity indicators. The ad hoc group noted that

  - The lack of a coherent internationally comparable standardised SES classification is particularly crucial. The ad hoc group strongly advocates that OECD countries adopt such a classification for use in all future data collection and surveys. PISA lays excellent groundwork in this area (cf. Willms, 1999), and its first cycle will test such a classification. The results of this test should be examined also from the general point of view of equality and equity indicators.

  - Intra-national regional sub-units used up to now most often consider socio-political entities such as provinces, cantons, etc. These may themselves be of very heterogeneous composition. Comparisons between rural and urban areas should be developed further. In most OECD countries a major fraction of the population lives in cities: urbanisation continues, cities become ever larger (metropolisation) with increasing social segregation of metropolitan sub-areas. The consequences of urbanisation and education policies are closely linked at the level of schools and classrooms, although in most countries their upstream co-ordination remains uncertain and informal. The effort on educational equality and equity indicators might be an opportunity to co-ordinate with urban research and policy and to foster international comparisons focusing on urban or metropolitan areas across OECD countries. Co-ordination with the OECD/PEB project should also be explored.

  - Migration movements from the developing to developed countries and regions will undoubtedly increase over the next decades. Immigrant minorities with growing cultural
differences already constitute and will continue to constitute a major issue in education systems and in education policy, particularly in respect to equality and equity. The development of relevant indicators for this area may greatly benefit from closer links between the OECD/INES project and the OECD’s network of migration specialists (SOPEMI), as well as other organisations dealing with migration issues, in order to co-ordinate definitions and categories.

- The CERI work on students with special needs has already contributed to some definitional co-ordination in the domain of special education. This work should be continued.

- The populations of most OECD countries include some kind of socio-cultural minorities, be they defined as racial, ethnic or linguistic. Children and youngsters cannot escape membership in these groups, which tends to be associated with specific educational problems and inequalities. International comparisons are particularly difficult in this domain because historical and social origins, as well as the definitions and status of such minority groups differ so widely among countries. Challenging tasks for the future will include not only defining a common conceptual denominator across countries of such groups from the viewpoint of educational equality but also a common conceptual framework for equity policies and measures aimed at them.

- The ad hoc group notes that data concerning the issue of representations, attitudes and judgements about the equity of education arrangements and systems are practically non-existent. Important and highly specialised development work will be necessary in this area, partly on an experimental basis in the early stages. For example, in future PISA work, a small set of questions about the students’ feeling of being treated fairly in school should become an integral part of the context questionnaire, where climate and quality of life in school are addressed.

- Ultimately, considering the vast amount and range of developmental work that lies ahead, as well as the controversial nature of the task, all members of the ad hoc group agree on the need for solid institutional arrangements within which to pursue this effort. In other words, the ad hoc group recommends that a specific “Network on equality and equity issues” be created within INES, which countries join on a voluntary basis, signalling thus their support.

79. The exact shape of such a network within existing INES institutions will have to be further clarified, considering also the overall structure of the INES project. It may be a network per se like the existing ones while focusing on the development of equality and equity issues. At the same time the solution to be looked for should take into account the fact that equality and equity issues, although of a specific nature, arise in all domains as table 1 shows. The new network might therefore be instituted across the existing networks and the technical group. Shared participation between the existing NWs (and the TG) and the new one would also favour cross-fertilisation and co-ordination and allow combining human, conceptual, methodological and institutional resources.
REFERENCES


