

Please cite this paper as:

Nusche, D. (2008), "Assessment of Learning Outcomes in Higher Education: a comparative review of selected practices", *OECD Education Working Papers*, No. 15, OECD Publishing.
<http://dx.doi.org/10.1787/244257272573>



OECD Education Working Papers
No. 15

Assessment of Learning Outcomes in Higher Education

A COMPARATIVE REVIEW OF SELECTED
PRACTICES

Deborah Nusche

DIRECTORATE FOR EDUCATION

Cancels & replaces the same document of 20 February 2008

**ASSESSMENT OF LEARNING OUTCOMES IN HIGHER EDUCATION: A COMPARATIVE
REVIEW OF SELECTED PRACTICES**

OECD Education Working Paper No. 15

by Deborah Nusche

About the author

Deborah Nusche prepared this working paper as a consultant to the Indicators and Analysis Division of the OECD Directorate for Education in April 2007. She is now a Policy Analyst in the Education and Training Policy Division of the same Directorate.

Contact: Deborah Nusche, Directorate for Education
[Tel: +33 1 4524 7801; Email: Deborah.Nusche@oecd.org]

JT03241528

OECD DIRECTORATE FOR EDUCATION

OECD EDUCATION WORKING PAPERS SERIES

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

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

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

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

Copyright OECD 2008

ABSTRACT

Higher education institutions (HEIs) have experienced increasing pressures to provide accountability data and consumer information on the quality of teaching and learning. But existing ratings and rankings of HEIs tend to neglect information on student learning outcomes. Instead, they focus on inputs, activities and research outputs, such as resources used, classes taught, and articles published. Such indicators provide no indication of the degree to which HEIs actually develop the knowledge and skills of their students. In most countries, hardly any comparable information is available on the educational quality of different programmes and institutions.

In some countries, approaches to assess higher education learning outcomes have been developed, but little cross-country information is available on the characteristics of the instruments used. This paper provides an overview of experience gained in this domain across OECD and partner countries. Based on illustrative evidence collected for 18 assessment instruments, it examines conceptual, organizational and methodological aspects of existing assessments. It proposes a typology of higher education learning outcomes and reviews the ways in which these have been assessed across countries. Examples are drawn from Australia, Brazil, Mexico, the United Kingdom and the United States.

RÉSUMÉ

Les institutions d'enseignement supérieur sont de plus en plus amenées à rendre des comptes sur la qualité de leurs enseignements et les résultats de leurs étudiants. Mais les méthodologies de notation et de classement des universités considèrent rarement dans leurs critères l'information sur les « résultats de l'enseignement », à savoir ce que les étudiants ont vraiment appris au sein de ces institutions. Elles se concentrent plutôt sur les inputs, activités, et outputs, tels que les ressources mobilisées, les cours enseignés et le nombre d'articles publiés. Cependant, ces indicateurs ne permettent pas de déterminer dans quelle mesure les universités contribuent au développement des connaissances et des compétences de leurs étudiants. Dans la plupart des pays, il y a peu d'information disponible pour comparer la qualité éducative des différents programmes et institutions.

Dans certains pays, des approches ont été développées pour mesurer la qualité de l'enseignement dans les universités, mais peu d'études offrent une comparaison internationale des différents instruments utilisés. Ce papier présente un aperçu des expériences dans ce domaine au sein de l'OCDE et des pays partenaires. A partir de données illustratives concernant 18 tests, il examine des aspects conceptuels, organisationnels et méthodologiques des instruments d'évaluation existants. Le papier identifie différents types de résultats de l'enseignement et étudie la façon dont ceux-ci sont évalués dans les différents pays. Cette étude s'appuie sur des exemples provenant de l'Australie, du Brésil, du Mexique, du Royaume-Uni et des États-Unis.

TABLE OF CONTENTS

ABSTRACT	3
INTRODUCTION	5
1 SCOPE AND LIMITATIONS OF THIS PAPER	6
1.1 Types of assessments considered	6
1.2 Limitations	6
1.3 Sources	7
2 A TYPOLOGY OF HIGHER EDUCATION LEARNING OUTCOMES	7
2.1 The concept of learning outcomes.....	7
2.2 Selecting learning outcomes for assessment	8
2.2.1 Cognitive outcomes	8
2.2.2 Non-cognitive outcomes	10
3 CURRENT PRACTICES OF LEARNING OUTCOMES ASSESSMENT	13
3.1 Outcomes assessed	13
3.2 Assessment process	16
3.2.1 Organizations involved and funding of assessment	16
3.2.2 Frequency of assessment	16
3.3 Assessment Design	22
3.3.1 Type of instrument (skill assessment vs. self-reporting)	22
3.3.2 Format	22
3.4 Participants	26
3.4.1 Target populations	26
3.4.2 Selection of participants	26
3.4.3 Incentives	27
3.5 Assessment purpose and results	33
3.5.1 Focus of assessment.....	33
3.5.2 Use of assessment findings	34
REFERENCES	39
ANNEX: DATA SOURCES (BY COUNTRY).....	44

INTRODUCTION

Over the past two decades, higher education institutions (HEIs) have experienced increasing pressures to provide accountability data and consumer information on the quality of teaching and learning. Demands for comparable information on student learning outcomes stem from several stakeholders including prospective students, employers in search of qualified graduates, taxpayers concerned about the efficient use of public funding, and policy makers deciding on accreditation and resource allocation. Society at large has an interest to know in how far HEIs effectively prepare their students to participate in increasingly knowledge-based economies. HEIs themselves can benefit from comparing their students' learning outcomes, for purposes of instructional improvement, public accountability and consumer information. Evidence of high quality learning outcomes may also serve to attract public funding and fee-paying students.

Existing ratings and rankings of educational quality, such as the *USA News & World Report* and the *Times Higher Education Supplement*, tend to neglect information on learning outcomes.¹ Instead, they focus on inputs, activities and research outputs, such as resources used, classes taught, and articles published. Such performance indicators provide no measurement of the degree to which HEIs actually develop the knowledge and skills of their students. Hence, these ratings and rankings are ill-suited to inform governments, students and the general public about teaching and learning quality. But in the absence of comparable learning outcomes assessment across HEIs, ratings and rankings are widely used as proxies for relative educational quality. They have attracted extensive media attention and they clearly influence public perceptions of HEIs and their graduates.

Policy makers and stakeholders in a range of countries have emphasised the need to develop instruments to obtain comparable information on what students actually learn across HEIs. In a few countries, approaches have been developed to assess and compare higher education learning outcomes between institutions, but little cross-country information is available on the characteristics of these instruments.

This paper aims to provide an international perspective on current practices in standardized learning outcomes assessment in higher education. Section one explains the scope and limitations of this paper. Section two proposes a typology of different types of higher education learning outcomes, and comments on the advantages and drawbacks of using different types of outcomes as indicators of learning quality. Finally, section three describes the ways in which different types of outcomes have been assessed across countries.

¹ For a review of the indicators and methodologies used by noteworthy international rankings, see for example Salmi and Saroyan, 2006; IHEP, 2007; Sadlak and Liu, 2007.

1 SCOPE AND LIMITATIONS OF THIS PAPER

This paper proposes a typology of different types of higher education learning outcomes, and summarises the characteristics of 18 assessment instruments designed to measure such outcomes on a regional or national level. Examples are drawn from Australia, Brazil, Canada, Mexico, the UK and the USA. The examples were chosen on the basis that their results may be used to compare the quality of higher education learning outcomes across courses, programmes and institutions.

1.1 Types of assessments considered

First and foremost, this paper is interested in large-scale direct assessments that focus explicitly on the degree to which HEIs develop learning outcomes in their students. Brazil is currently the only country where testing is mandatory for students from all HEIs and takes place on the national level. But large-scale direct assessments also exist in Australia, Mexico and the USA. In Australia and Mexico, institutions may voluntarily subscribe to nation-wide standardized tests for graduating students. In the USA, non-governmental assessment agencies offer a vast array of different tests that are being used by hundreds of HEIs every year. Although not covering all HEIs, such widely administered assessment instruments also allow for comparisons of results across programmes and HEIs.

As large-scale standardized assessments are not widely applied in higher education institutions, the paper also looks at graduate admission tests that may be used as performance indicators for undergraduate HEIs. Graduate school entry tests measure students' mastery of the knowledge and skills required to undertake graduate study. Such tests exist in many countries. This paper includes only two examples, because there are important limitations to using the results of graduate admission tests to draw conclusions about undergraduate HEIs. Graduate admission tests may be of limited relevance to the learning goals of undergraduate HEIs, and there is a selection bias in who prepares for and takes these tests (Klein et al., 2005).

Finally, the paper considers measurements of learning outcomes that do not involve direct testing but rely instead on student reporting. Some surveys, such as the CEQ (Australia) and the NSSE (USA), ask students to rate the degree to which the HEI experience had an impact on their learning outcomes. Others, such as the DLHE (UK) and the GDS (Australia), are concerned with labour market outcomes and further study. While such surveys and questionnaires do not provide direct evidence of learning outcomes, they may serve as secondary indicators of some aspects of quality in higher education.

1.2 Limitations

This paper does not intend to provide a comprehensive overview of all existing instruments. The tables presented serve to illustrate different types, designs, and uses of learning outcomes assessment at the national levels.

The paper does not include any of the various learning outcomes assessments undertaken internally by individual HEIs. It only intends to discuss large-scale standardized assessment instruments that are designed and evaluated externally and thus offer comparability across programmes and institutions.

The paper does not include vocational qualifications, or assessment instruments designed by professional organizations or professional schools (such as law schools or medical schools).

1.3 Sources

The data collected stems from web-based research. Information on assessment instruments designed by private agencies is largely drawn from the testing companies' websites. Information on assessment instruments initiated by government bodies was drawn from the websites of various public agencies, such as the Ministries of Education and national statistics agencies. Websites of non-governmental organizations and individual HEIs involved in assessment were also reviewed to gather information on the concrete procedures of test administration and use of results. One limitation in the data collection was that web-based information on national assessments is often available in the national language only. The examples presented are based on information available in English, French, Spanish and Portuguese. It should be noted that some characteristics of assessment instruments are subject to change. All data was collected in April 2007. A list of sources used for each test is provided in the Annex.

2 A TYPOLOGY OF HIGHER EDUCATION LEARNING OUTCOMES

2.1 The concept of learning outcomes

Learning outcomes refer to the personal changes or benefits that follow as a result of learning. Such changes or benefits can be measured in terms of abilities or achievements. Otter (1992, p.i) defined learning outcomes as "what a learner knows or can do as a result of learning." As this paper focuses on learning outcomes in higher education, it will primarily consider those outcomes that are believed to be attributable to the higher education experience, rather than to normal individual development, social maturation and other influences beyond the reach of HEIs. This paper looks at those learning outcomes that follow as a result of students' engagement in the learning opportunities offered by HEIs.

Focussing on student learning requires an approach that clearly distinguishes between outcomes and other frequently used performance indicators of educational quality, namely inputs, activities, outputs.² The misuse of these terms can lead to much confusion, and it is therefore important to establish a coherent terminology. *Inputs* are the financial, human and material resources used, such as funding and endowments, faculty and administration, buildings and equipment. *Activities* are actions taken or work performed through which inputs are mobilized to produce specific outputs. Examples of higher education activities include curriculum design and teaching. *Outputs* are anything that an institution or system produces. HEI outputs can be measured in terms of articles published, classes taught, educational material distributed, and degrees awarded. Inputs, activities and outputs have little intrinsic value in terms of student learning. They are only the intermediate steps that may or may not lead to outcomes or benefits.

Outcomes describe what the student actually achieves, as opposed to what the institution intends to teach (Allan, 1996). According to Eisner (1979, p.103), outcomes "are essentially what one ends up with, intended or not, after some form of engagement". Many authors define learning outcomes as something that can be *observed, demonstrated and measured* (Spady, 1988; Melton, 1996). The statement of learning outcomes by educational institutions often implies that assessment and evaluation of their quality can be achieved (Melton, 1996).

² The following definitions of inputs, activities and outputs are based on terminology used in the OECD Glossary for Results-based management (2000).

The term learning outcomes has its origins in *outcomes-based education*, a model of educational structuring that involves the clear and explicit identification, statement and assessment of student learning (Spady, 1988; Allan, 1996; Andrich, 2002; Adam, 2004). Outcomes-based education systems organise curricula around explicit and detailed student outcome statements. Such statements describe what the learner is expected to know, understand and/or be able to demonstrate at the end of a period of learning (Adam, 2004). Outcomes-based approaches are most frequently used in secondary schooling (Ewell, 2005). In higher education, outcomes-based approaches were first introduced in the USA, Australia, New Zealand and the United Kingdom, but more recently also in other OECD countries (Adam, 2004). Defining curricula in terms of expected outcomes is an important step in the direction of learning outcomes assessment. Once HEIs have specified expected student outcomes explicitly and in a measurable way, comparative assessment of learning outcomes becomes feasible.

Outcomes of higher education are not limited to learning outcomes. Students can benefit from their HEI experience in many different ways, such as better social status, higher employment rates, civic engagement, opportunities to pursue further studies, or simply leading a more fulfilled life (Ewell, 2005). While such outcomes are related to learning, they should not be confused with the actual mastery of knowledge, abilities, and skills that result from students' engagement in HEI learning experiences (Ewell, 2005). Such long-term social and economic benefits of the HEI experience can serve as secondary proxies for learning outcomes, but they are not direct outcomes of learning.

2.2 Selecting learning outcomes for assessment

While it is relatively straightforward to define the meaning of outcomes, there is little consensus concerning the scope and content of learning. Learning has many dimensions some of which are easier to measure than others. Institutions pursue different missions and objectives and thus greatly differ in their teaching and learning priorities. Some programmes do best at transmitting domain-specific knowledge, whereas others excel in teaching generic skills, or competencies. No single assessment can comprehensively measure all the important outcomes of higher education. But those involved in assessment may select the outcomes that are most relevant to their purpose and that correspond to the missions and objectives of the HEIs under assessment.

This section proposes a typology of higher education learning outcomes. It aims to define, describe and classify different higher education learning outcomes, and it comments on the advantages and disadvantages of using different types of outcomes as indicators of learning quality in higher education. The typology is based on the classic separation between cognitive and non-cognitive learning outcomes and then presents the concept of "competencies" that goes beyond the cognitive/non-cognitive separation still underlying most assessment activities.

It should be noted that any categorization of learning outcomes involves some artificiality, as the different components of learning are interdependent and overlapping (Bowen, 1977). The typology presented should simply be considered as a heuristic framework aiming to identify different dimensions of learning and the achievement tests that might be applied to assess them.

2.2.1 Cognitive outcomes

Most of the assessment instruments included in this study focus on the assessment of cognitive learning outcomes. Cognitive learning refers to the recall or recognition of knowledge and to the development of intellectual abilities and skills (Posner, 1992). Broadly defined, cognitive learning outcomes "range from domain-specific knowledge to the most general of reasoning and problem-solving skills" (Shavelson and Huang, 2003, p.13).

Various classifications of cognitive learning outcomes exist (see for example Gagné, 1977; Kolb, 1981; Klemp, 1977; Eraut, 1990; Marzano, 2001). Many of them are inspired by Bloom's (1956) taxonomy of educational objectives, which provides a framework of outcomes that education intends to develop, going beyond factual knowledge and comprehension to include academic skills like application, synthesis, analysis, and evaluation. Though remaining a standard reference, Bloom's taxonomy has since been critiqued, developed and modified by many authors (for a review, see Anderson and Sosniak, 1994).

For the purposes of this paper it will suffice to rely on a simple distinction between knowledge and skills outcomes. This basic distinction has been widely adopted by the assessment literature.

Knowledge outcomes

Knowledge acquisition involves the "remembering, either by recognition or recall, of ideas, materials or phenomena" (Bloom, 1956, p.62). Assessments of knowledge outcomes may focus either on general content knowledge or on more domain-specific knowledge.

General content knowledge refers to the knowledge of a certain core curriculum whose content is considered "essential learning" (Maeroff, 2006). The Brazilian national exam for the assessment of student performance (ENADE), for example, has a section on general content knowledge including questions on social and biological diversity, public policy, and topics regarding Brazilian and international issues (Verhine and Dantas, 2005). Some stakeholders sustain that all undergraduate students should acquire a body of such essential knowledge, independently of their major field of study (AAC&U, 2004). Such core curricula remain an important part of undergraduate education in some countries. Assessing general content knowledge across courses, programmes and institutions can be useful to compare how different institutions promote certain "common" learning contents that are judged essential for all students. Yet, as general content knowledge constitutes often only a small part of what is taught in higher education programmes, it may be considered insufficient to use general education outcomes *alone* as indicators of educational quality.

Domain-specific, or subject-specific, knowledge outcomes refer to acquired knowledge in a particular field, such as biology or literature. Assessments focussing on domain-specific knowledge outcomes are particularly useful to compare learning quality in a particular field across different institutions. The opportunity for concentrated coursework within an academic field of specialization is arguably the most essential part of higher education and HEIs are specifically mandated to make students advance in their major fields of specialization (Volkwein, 2003). Pascarella and Terenzini (2005) found that while studying at a HEI, undergraduate students make the greatest gains in those domains that are consistent with their major area of studies. In contrast to other outcomes, improvement of highly domain-specific knowledge is quite clearly related to what is being taught and learned at HEIs (Allan, 1996). It can be argued that HEIs should thus primarily be held accountable for the specialist knowledge that their students acquire in their major field.

Skills outcomes

Cognitive *skills* are based on complex processes of thinking, such as verbal and quantitative reasoning, information processing, comprehension, analytic operations, critical thinking, problem-solving and evaluation of new ideas. There is some disagreement as to whether such thinking processes are generic (following general patterns) as opposed to being field-specific.

Assessments aiming to compare learning outcomes across different courses often focus on *generic skills outcomes*. The common characteristic of all generic skills outcomes is that they transcend disciplines. They are transferable between different subject areas and contextual situations. Such skills are not directly

tied to particular courses. They relate to any and all disciplines and they allow students to be operational in a number of new contextual situations (Pascarella and Terenzini, 2005). Generic skills outcomes can be assessed using tests that are based on application rather than on knowledge, thus focussing on students' ability to solve intellectual problems. Usually, students are asked to provide constructed answers that also give evidence of writing skills. Focusing on outcomes in terms of skills may allow comparing how well programmes and institutions with diverging missions and ways of teaching achieve to develop certain common skill dimensions in students.

Yet, there are some doubts as to whether such outcomes can really be connected to the university experience. Although HEIs often emphasise the transmission of generic skills in their overall missions and objectives, such skills are rarely an explicit part of particular course curricula. The question is, then, where such generic skills are actually acquired and which role universities play in developing them. Some studies have revealed that student development in terms of academic skills is correlated with the age of students (Banta and Pike in Ewell, 1991). Such an age/performance correlation suggests that some skills outcomes may be more a result of social maturation than of studying at a HEI (Ewell, 1991). There is a risk that generic skills assessment may end up measuring students' intelligence and results of prior schooling more than revealing the impact that HEIs have made on their learning.

Domain-specific skills are the thinking patterns used within a broad disciplinary domain, such as natural sciences or humanities. They are stated in terms of methods of enquiry, ways of evaluating evidence, and patterns of procedure necessary to confront new contextual situations in specific fields of study. They involve an understanding of how, why, and when certain knowledge applies (Shavelson and Huang, 2003). Domain-specific skills are not entirely transferable throughout subject areas. For example, the ability outcome "excellent writing" takes different forms and requires different skills in the various disciplines (AAC&U, 2004). Some argue that although academic skills are a general outcome of higher education, they can hardly be tested independently of disciplinary subject matters. In line with this view, domain-specific skills may be assessed by providing students with new documentation taken from their domain of expertise, and asking them to assess the quality of the evidence and make use of it to write complex answers (AAC&U, 2004).

2.2.2 Non-cognitive outcomes

Non-cognitive development refers to changes in beliefs or the development of certain values (Ewell, 2005). Mission statements of HEIs often include non-cognitive elements, which shows that their role goes beyond the acquisition of knowledge and skills. Non-cognitive outcomes may be developed both through classroom instruction and out-of-class activities that are organised by HEIs to supplement the curriculum. Such activities or "co-curricula" include advising, tutoring, counselling, student-faculty relations, clubs, athletics and other activities. The existence of such services indicates that HEIs value the development of non-cognitive learning as a way of complementing the learning that occurs during classroom teaching (Middle States Commission on Higher Education, 1996).

Many attempts have been made to establish theoretical and empirical taxonomies of non-cognitive higher education outcomes. Studies on non-cognitive learning outcomes often focus on the presence or absence of certain theorized stages of identity development (Pascarella and Terenzini, 2005). Pascarella and Terenzini (2005) provide a synthesis of more than 2,500 studies on the impact of American colleges on their students. Among the most frequently assessed variables are outcomes related to psychosocial development, attitudes and values.

Psychosocial development includes aspects of self-development such as identity development and self-esteem, as well as relational developments such as students' relationships with people, institutions and conditions. Relational outcomes include interpersonal and intercultural skills, as well as autonomy and

maturity. *Attitudes and values* are closely interrelated and often confounded. A distinction can be made between the two, as attitudes are beliefs focused on a specific object, whereas values are generalized standards that transcend attitudes (Rokeach, in Pascarella and Terenzini, 2005). Attitudinal and value outcomes may include social responsibility, motivation for learning and understanding of diversity (Volkwein, 2003).

According to Astin (1984), learning outcomes are not simply the consequence of an institution's educational quality, but rather a function of students' active engagement with the learning opportunities that the HEI presents. In line with this rationale, the National Survey for Student Engagement (USA), for example, aims to measure the extent to which HEIs actively encourage high levels of engagement.

The definition of desirable non-cognitive learning outcomes for the purpose of assessment may be somewhat controversial. Judgements as to which attitudes and values are desirable or "right" are not always shared by all stakeholders. Moreover, the definition of desirable non-cognitive outcomes may involve quite some variation between cultural contexts. There are also some doubts as to whether non-cognitive developments can really be attributed to the university experience. Not all students participate in extracurricular campus activities and not all their out-of-class experiences are related to campus life. Some studies suggest that non-cognitive outcomes are rather related to social maturation, generational effects (Pascarella and Terenzini, 2005) or "significant life events" (Glenn in Pascarella and Terenzini, 2005, p.272). These factors can hardly be controlled for when measuring the impact of HEIs on student development.

The study of non-cognitive outcomes of higher education is more complicated than that of cognitive outcomes. The links between values and beliefs on the one hand and observable activities and behaviours on the other are not clearly established (Pascarella and Terenzini, 2005). There is little evidence that values and beliefs can be demonstrated and measured by using behavioural data. Therefore, non-cognitive outcomes are generally measured indirectly, through questionnaires and surveys, including student self-reports and faculty and employer surveys. Such indirect measures are based on individual perceptions as much as on facts. The results may be less objective indicators of student learning than direct measurements of knowledge and skills.

Going beyond the cognitive/non-cognitive separation of outcomes, Rycher (2004, p.7) proposes a more general model of learning outcomes "in which competence is defined as the ability to meet demands or carry out a task successfully, and consists of *both cognitive and non-cognitive dimensions*". According to Winch and Foreman-Peck (2004, p.4), "competence lies in a mix of action, knowledge, values and goals in changing settings." Such concepts of student outcomes are based on the idea that outcome typologies may not adequately capture the essence of integrated outcomes that unite different skills into real expertise (Ewell, 2005). Along this rationale, student outcomes should be defined and observed in the context of actual performance tasks.

Assessment instruments focusing on *competencies* involve testing complex combinations of cognitive, affective and behaviour traits. Student portfolios, for example may represent direct evidence of student work, such as written assignments, field performances, interviews, laboratory and internship reports (Ewell, 2005). Portfolios may also include indirect evidence on such outcomes, namely surveys and questionnaires asking students to rate their own development in terms of competencies. Evaluating multiple student products allows integrating a wide range of learning outcomes (Ewell, 2005).

Otter (1992) emphasizes the difference between general competence, which is a broad set of abilities that can be applied in a range of settings, and *occupational competence*, which is a narrower description of the abilities needed for employment in a specific occupation. Occupational competencies may also be referred to as *employability*. Preparing students for competence in the workplace is a major goal of higher

education (Bowen, 1977). HEIs have come under increasing pressures to prepare their students' ability to meet the needs of industry, commerce and other service organizations (Seagraves et al, 1996). This includes the development of skills that are valued in employment (Dearing, 1997).

However, the definition of learning outcomes in terms of occupational competence may raise several issues of concern. Important dimensions of higher education, such as providing opportunities for in-depth study and helping students to develop their potential, may be undervalued if HEIs focus solely on occupational competencies (Otter, 1992; Melton, 1996). Also, it is not always possible to define clear occupational objectives for each subject domain, because graduates may take up employment in a wide range of occupations (Melton, 1996). For some subjects it may not be feasible to define clearly related occupational roles. Finally, employment outcomes statements tend to focus on immediate employment needs, whereas students may be more interested in developing intellectual skills that would enable them for lifelong and unpredictable future labour markets, rather than just for an initial job (Melton, 1996; AAC&U, 2004).

In some countries, desired occupational competencies are defined within national qualifications systems. Higher level qualifications may be offered at HEIs, either stand-alone or in combination with other degrees, as for example in the UK (Morgan, 2002). According to England's Qualifications and Curriculum Authority (QCA), such higher level vocational qualifications should be "a statement of competence clearly relevant to work and intended to facilitate entry into, or progression in, employment, further education and training... incorporating the assessment of skills to specified standards, relevant knowledge and understanding, and the ability to use skills and to apply knowledge and understanding to relevant tasks" (QCA Website, 2007). Assessment of occupational competencies may include the constitution of portfolios that reflect work-place performance of the candidate. Sources of evidence can include direct observation of the candidate, audio, video or electronic recording of his/her activities, and products of his/her work (QCA Website, 2007).

Another frequently used way to measure the extent to which HEIs develop their students' occupational competencies is to look at employment rates or graduate destinations. Graduate surveys may provide important information on the ways in which students perceive the usefulness of their HEI education in terms of occupational outcomes.

It should be noted, however, that labour market outcomes are not always accurate reflections of the actual competencies gained. A body of research suggests that employment rates depend not merely on higher education learning outcomes, but on a variety of factors including socio-economic factors, prior learning (see for example Bratti et al., 2001), the subject studied (see for example McGuinness, 2003), social networks (see for example Montgomery, 1991), and HEI cost (see for example Pascarella et al., 1992). According to Spence's (1973) job-market signalling model, employers seek graduates from more selective HEIs, because they assume that the proportion of students with high abilities is higher among them, as it is less costly for them to acquire education than it is for students with low abilities. In this model, the educational quality of the HEI is less relevant to the employer than the fact that the HEI fulfils a sorting function and conveys information about the applicants' abilities.

3 CURRENT PRACTICES OF LEARNING OUTCOMES ASSESSMENT

This section provides an overview of selected current practices of learning outcomes assessment. It summarises the characteristics of 18 assessment instruments designed to measure higher education learning outcomes across HEIs on a regional or national level. Examples are drawn from Australia, Brazil, Canada, Mexico, the UK and the USA. Based on this illustrative evidence, this section provides comparative information on some of the conceptual, methodological and organisational aspects in assessing learning outcomes. Four questions are being considered in this section:

1. *What* is being assessed? Section 3.1 gives information on the type of outcomes assessed by each of the instruments considered.
2. *How* are these outcomes being assessed? Section 3.2 presents the processes of designing, developing, administering and reporting on the test, while section 3.3 gives information on the technical features of the assessment instruments such as format, number of items, and duration of assessment.
3. *Who* is being assessed by each instrument? Section 3.4 summarizes the characteristics of assessment participants, providing information on the target population, the selection of participants, the coverage of test application and the incentives for students and institutions to participate in the assessment.
4. *Why* is the assessment being applied? Section 3.5 indicates whether the purpose of the assessment is to evaluate individual students, programs, institutions or education systems. It gives information on the type of results that are being sought and on the ways in which these results are used by different stakeholders.

Throughout this section, it will become apparent that the characteristics presented in the different sub-sections are closely interlinked. For example, the choice of methodological design depends very much on the types of outcomes one tries to capture and on the purpose the results are expected to be used for.

3.1 Outcomes assessed

Building on the typology provided in the previous section, Table 1 gives an overview of the types of outcomes that each assessment instrument seeks to capture.

Most *direct assessments* of learning outcomes focus on cognitive learning outcomes. The Brazilian ENC-Provao and its successor exam ENADE focus on domain-specific knowledge and skills that are considered essential and common to all HEI curricula in the specific domain. The ENADE proposes specific tests for 13 different subject areas, but it also assesses general content knowledge and skills.

The tests used in Mexico also take this approach: The general exit exam (EGEL) focuses on essential subject-specific knowledge and skills and is available for 33 different subject areas. The exit exam of the technical university track (EGETSU) is available for all 19 subject areas offered, but in addition it also includes assessment of generic skills necessary for most domains such as English and IT.

Similarly, many of the assessment instruments used in the USA assess both generic and domain-specific cognitive learning outcomes. The Major Field Test for example is available for 15 undergraduate disciplines and for MBAs, and the Collegiate Learning Assessment focuses on performance-based tasks set in the context of broad disciplines (natural sciences, social sciences, humanities, arts).

Several assessments and surveys do not include domain-specific items and focus mostly on generic knowledge and skills. The Australian Graduate Skills Assessment (GSA) and Course Experience Questionnaire (CEQ) are interested in skills such as critical thinking, problem solving and written communication, and ACER is currently considering to add basic management skills, IT skills and research skills to the GSA. Several postgraduate entry tests, such as the EXANI-III in Mexico and the GRE in the USA also examine generic skills such as verbal and mathematical reasoning and capacities in information use.

Student *surveys and questionnaires* often focus on non-cognitive outcomes or on competencies. The Australian CEQ asks graduates to rate their teamwork skills, ability to plan work and confidence in tackling unfamiliar situations. The National Survey of Student Engagement (NSSE) used in the USA and Canada focuses on how engaged students are with their HEI experience. Issues of interest include how undergraduates spend their time, what they have gained from courses and how they use HEI services.

Surveys delivered to former students such as the Australian Graduate Destination Survey (GDS), the Canadian Youth in Transition Survey (YITS), and the UK Destinations of Leavers from Higher Education (DLHE) investigate the degree to which recent graduates have developed occupational competencies. Most of them provide information on the types of jobs graduates have gone into and whether they have pursued further studies. As discussed in section two, labour market outcomes are not always accurate reflections of the actual competencies gained, as many factors beyond occupational competencies of graduates influence them.

Table 1: Outcomes assessed by the instruments considered in this study

Country	Test Name, Introduction Date	Type of outcomes assessed
Australia	Graduate Skills Assessment (GSA), 2000	Generic skills: Critical Thinking, Problem Solving, Written Communication (ACER is currently considering modifications such as the addition of basic skills, management skills, IT skills, research skills). Domain-specific knowledge and skills: (Not yet included but ACER is currently considering the possibility of testing elements within various broad Field of Study groups) Non-cognitive outcomes: Interpersonal understanding.
Australia	Course Experience Questionnaire (CEQ), part of the Graduate destination survey since 1993	Generic skills: Problem Solving, Analytic Skills, Written Communication Skills. Non-cognitive outcomes: Teamwork skills, Student satisfaction with the following: Teaching, Goals and Standards, Workload, Assessment. General competencies: Confidence in tackling unfamiliar situations, Ability to plan work.
Australia	Graduate Destination Survey (GDS), 1972	Occupational competencies: Employment outcomes approximately 4 months after graduation: availability for employment, sectors of employment, average annual salaries, graduates' job search activities. Further study activities, such as mode of study (full/part-time), levels of study, fields of education, and institution.
Brazil	Exame Nacional de Cursos (ENC or "Provão"), 1995-2003	Domain-specific knowledge and skills that are considered essential and common to all HEI curricula in the specific domain. Available for 26 subject areas.
Brazil	Exame Nacional de Desempenho dos Estudantes (ENADE), 2004	Domain-specific knowledge and skills that are considered essential and common to all HEI curricula in the specific domain. Available for 13 subject areas. General content knowledge: Among the assessed themes are biological and social diversity, public policies, social networks, citizenship, and current events and problems. Generic skills: Ability to infer, interpret poetic texts, establish common points, identify associations, reflect, deduct, and understand graphics.

Canada	National Graduate Survey (NGS), 1978. Follow-up Survey of Graduates, 1987	Occupational competencies: Employment outcomes two years and five years after graduations: Information on the number, characteristics, and duration of all jobs held since graduation, on the length of job search, the match between education and occupation. Graduate satisfaction with their HEI experience.
Canada	Youth in Transition Survey (YITS)	Occupational competencies: Questions for the 3 rd cycle of assessment (target population then aged 22-24) include question on postsecondary education and engagement and employment outcomes.
Mexico	Exámen Nacional de Ingreso al Posgrado (EXANI-III), 1997	Generic skills: Verbal and mathematic reasoning, Capacities to infer, analyse and synthesize. Competencies in information use: organise, obtain and understand information.
Mexico	Exámen General Para el Egreso de la Licenciatura (EGEL), 1994	Domain-specific knowledge and skills that are considered essential and common to all HEI curricula in the specific domain. Available for 33 different subject areas.
Mexico	Exámenes Generales para el Egreso del Técnico Superior Universitario (EGETSU), 2000	Domain-specific knowledge and skills: Comprehension levels and problem-solving skills needed in the student's major field. Tests are available for all 19 areas of the Technical University Track. General content knowledge and generic skills: knowledge and ability necessary for all careers, namely social and economic knowledge, IT, and English.
UK	Destinations of Leavers from Higher Education (DLHE), 2002 (replaced the "First Destination Supplement")	Occupational competencies: Employment and further study outcomes six months after graduation: how many graduates are in employment, the types of jobs they go into, and how many go onto further study.
USA	Collegiate Assessment of Academic Proficiency (CAAP), 1988	Generic skills: Writing (objective and essay), reading, mathematics, science reasoning, critical thinking, curricular content drawn from all fields.
USA	Measure of Academic Proficiency and Progress (MAPP), 2006 (replaced the ETS "Academic Profile" test, 1992-2006)	Generic and domain-specific skills: Reading and critical thinking are measured in the context of humanities, social sciences, or natural sciences. Writing, mathematics.
USA	Tasks in Critical Thinking, 1992 (discontinued)	Generic and domain-specific skills: inquiry, analysis, communication skills. Performance based tasks set in the context of broad disciplines (natural sciences, social sciences, humanities, arts).
USA	Major Field Tests, 1990 (based on the GRE Subject Tests)	Domain-specific knowledge and skills that are considered most important within each major field of study: factual knowledge, ability to analyse and solve problems, ability to understand relationships, ability to interpret material including graphs, diagrams, and charts based on material related to the field. Available for 15 undergraduate disciplines and for MBAs.
USA	Collegiate Learning Assessment (CLA), 2002	Generic and domain-specific skills: critical thinking, analytic reasoning, written communication, ability to use information. Competencies: Real-life tasks such as preparing a memo or policy recommendation by using different types of documents and data that must be reviewed and evaluated. Performance based tasks set in the context of broad disciplines (natural sciences, social sciences, humanities, arts).
USA and Canada	National Survey of Student Engagement (NSSE), 2000 (in Canada since 2004)	Non-cognitive outcomes: information on student engagement: how undergraduates spend their time and what they gain from courses, extracurricular activities, and HEI services.
Test centres in the USA, Canada and other countries	Graduate Record Examination (GRE) General Test, 1966	Generic skills: verbal reasoning, quantitative reasoning, and analytical writing.

3.2 Assessment process

Based on Table 2 below, this section provides basic information for the 18 assessment instruments included in this study, i.e. how the different assessment instruments are financed, designed, implemented and administered. This section compares the different types of organisations involved in the assessment process, as well as the frequency of assessment and the ways in which assessment results are reported.

3.2.1 Organizations involved and funding of assessment

Table 2 shows that all considered assessment instruments are designed by organisations external to the HEIs under assessment. Typically, the external supplier processes the completed tests and prepares a report for the institution. Organizations involved in assessments include government bodies, national statistics agencies, HEI federations, private testing companies, and non-for-profit organizations. For some national exams, such as the EGEL and the EGETSU (Mexico), representatives of academia are involved in the test development.

Tests developed by private assessment agencies are most widely used in the USA, where HEIs order the assessment materials directly from the test supplier and administer them on campus. The most important American assessment agencies are ETS and ACT. The Brazilian national exam (ENADE) is also developed and analysed by private assessment agencies, but it is publicly funded and involves no cost for HEIs. In Australia, some places in the GSA are funded by the Department of Education, Sciences, and Training, but HEIs must pay for additional places. In Mexico, the students themselves have to pay for their participation in the national graduation exam (EGEL).

Even if externally-designed assessments are relatively cost-effective for HEIs, some countries report that students and institutions may be opposed to the very idea of implementing such external instruments on their campus. The test questions may be seen as not relevant for some programmes and the results may not be directly applicable to improving teaching and learning at individual HEIs. Moreover, results may not be available at disaggregate levels, which is where they would be most important for the purposes of institutional learning and improvement (Borden and Owens, 2001). On the other hand, some assessment agencies now offer customized options to HEIs, such as adding locally-developed questions or disaggregating data according to local needs.

3.2.2 Frequency of assessment

Table 2 further indicates that some assessment instruments are designed to give a snapshot of learning outcomes at a specific moment, whereas others are designed to measure student progress over time.

Single testing means that assessments are applied only once, at the end of a period of learning. If such single testing focuses on domain-specific knowledge outcomes, such as the EGEL (Mexico), and the MFT (USA), it is likely that these outcomes are at least to some degree linked to what has been taught throughout the HEI programme. But when single tests focus on more general knowledge and skills, it may be difficult to prove that such outcomes actually stem from the HEI experience. Results of one single test will mostly reflect students' cumulative learning outcomes rather than institutional impact. When program/institution-level performance is to be evaluated, it may be difficult to provide evidence that such learning can actually be attributed to the higher education experience.

Many studies have shown that a considerable number of other factors beyond the HEI experience may influence student learning outcomes (Pascarella and Terenzini, 2005). Therefore a number of assessment

instruments now focus on measuring the “value added” to student learning by HEIs. This can be done through cross-sectional or longitudinal assessment.

Cross-sectional assessment: Several of the presented assessment instruments, such as the GSA (Australia) and the ENADE (Brazil), apply tests to samples of freshmen and senior students. Such simultaneous testing of students at entry level and students at graduation level is referred to as *cross-sectional assessment* (Terenzini, 1989). In this approach, entering students are used as a control group with supposedly similar entry qualifications as current graduates. Observed differences are taken as an indication of the degree to which the HEI experience contributes to student learning processes.

It should be noted that cross-sectional assessment does not focus on a panel in which the same students are tested over time. One limitation of this approach is that it is based on the assumption that current freshmen share important characteristics with current seniors at the time when they entered the HEI. Also, the approach does not account for self-selection of students through continuation and drop-out. Changes in admission standards may also produce important differences in freshmen and senior characteristics (Terenzini, 1989). According to Pascarella (1987), such problems of non-equivalent groups may be reduced by controlling for age and entering academic aptitude.

Longitudinal assessment: Some assessments measure learning outcomes of the same group of individuals over time. For direct assessments, longitudinal design is less frequently used than for surveys. However, some of the tests available in the USA encourage application in longitudinal design (e.g. the CAAP, the MAPP and the CLA). These tests exist in multiple forms in order to allow for pre- and post-testing. Ideally, the performance of a group of students should be measured at the time of matriculation, and then the same group should be tested again after a certain period of time. The major advantage of longitudinal assessment designed in this way is that it allows controlling for input factors, as students’ entering characteristics are taken into account. Another way of comparing student performance at entry level and at graduation level is to compare graduates’ results with their performance in admission tests.

However, longitudinal assessment does not usually control for environmental factors such as off-campus experiences that may also influence outcomes. Theoretically, this problem could be overcome if an assessment followed over the same period of time a control group of high-school graduates who do not attend an HEI but who have equivalent personal and academic characteristics as entering students (Terenzini, 1989). But it is very difficult to obtain a sample of students with similar characteristics who do not attend HEIs. The YITS (Canada) measures development of individuals in the 18-20 age cohort every two years, independently of whether they attend an HEI or not. Such data could be used to compare the development of individuals who did not pursue post-secondary studies with the development of those who did.

Some drawbacks of longitudinal assessment are that a certain proportion of the observed students drop out of studies over time, and that such assessment takes many years to complete and thus involves high direct and indirect costs (Terenzini, 1989). Moreover, Terenzini (1989) describes “ceiling effects” as a limitation of measuring difference scores between college entry and graduation: students with high entry scores may have much less room for improvement and consequently may show smaller gains than students with low initial scores. According to Verhine and Dantas (2005), another problem of value-added assessment is the risk that an institution might encourage freshmen students to perform badly on the tests in order to boost its value-added score. Institutions might also hold back weaker students in higher years whose performance may jeopardize the institution’s value-added results.

Table 3: Assessment process

Country	Test Name, Introduction Date	Initiator/Sponsor of Test Development	Responsibility for test design and analysis	Administration of assessment	Frequency of assessment	Single, cross-sectional or longitudinal testing	Reporting of assessment results
Australia	Graduate Skills Assessment (GSA), 2000	Federal Government: Department of Education, Science and Training (DEST), formerly Department of Education, Training and Youth Affairs (DETYA)	Australian Council for Educational Research (ACER)	Administered on HEI campus, supervised by faculty	Twice a year, once for entry and once for exit examinations.	Cross-sectional assessment of students at entry level and students at graduation level.	Personalised reports for students, containing total score on each component and relative achievement compared to all other participants of the same year. HEIs receive data on individual student performance and aggregated institutional reports.
Australia	Course Experience Questionnaire (CEQ), part of the Graduate destination survey since 1993	Federal Government: Department of Education, Science and Training (DEST), formerly Department of Education, Training and Youth Affairs (DETYA)	Graduate Careers Council of Australia (GCCA); Australian Council for Educational Research (ACER)	Questionnaires are supplied by GCCA and sent out to students by HEIs	Annually	Single testing	Data for each HEI is reported to the DEST. HEIs receive institutional summary reports. Results are made public in a variety of aggregations and levels of detail by universities, GCA, ACER and the government. The press and a commercial publication ("The Good Universities Guide") draw on results to establish rankings for public consumption
Australia	Graduate Destination Survey (GDS), 1972	Federal Government: Department of Education, Science and Training (DEST), formerly Department of Education, Training and Youth Affairs (DETYA)	Graduate Careers Council of Australia (GCCA), Australian Council of Educational Research (ACER), University of Melbourne's Information Technology Service (ITS)	Questionnaires are supplied by GCCA and sent out to students by HEIs	Annually	Single testing	Data for each HEI is reported to the DEST. HEIs receive institutional summary reports. Results are made public in a variety of aggregations and levels of detail by universities, GCCA, ACER and the government. A national file, national tables, a media release and GradStats (a four-page summary of results) are publicly available.

Brazil	Exame Nacional de Cursos (ENC or "Provão"), 1995-2003	Federal government	CESGRANRIO (specialized assessment agency)	Organised by Instituto Nacional de Estudos e Pesquisas Educacionais 'Anísio Teixeira' (INEP) on a nation-wide testing date, administered by municipalities	Annually	Single testing	Results were disclosed through technical reports, institutional bulletins (per area of study and course), and student bulletins. Students' individual bulletins were only available to the students themselves. Annual course classifications were made public and rankings were taken up by the media.
Brazil	Exame Nacional de Desempenho dos Estudantes (ENADE), 2004	Federal government	CESGRANRIO and CESPE (specialized assessment agencies)	Organised by Instituto Nacional de Estudos e Pesquisas Educacionais 'Anísio Teixeira' (INEP) on a nation-wide testing date, administered by municipalities	Once every three years.	Cross-sectional assessment of freshmen and senior students	Results are disclosed through technical reports, institutional bulletins (per area of study and course), and student bulletins. Students' individual bulletins are only available to the students themselves. Results are reported in a discrete manner that draws little attention from the media.
Canada	National Graduate Survey (NGS), 1978. Follow-up Survey of Graduates, 1987	Federal Government: Human Resources Development Canada (HRDC)	Statistics Canada (StatCan) for HRDC	Answers are collected by regional StatCan offices through Computer Assisted Telephone Interviews (CATI)	Periodically. NGS in 1978, 1984, 1988, 1992, 1997, 2002; Follow-up surveys three years later on the original respondents (since 1987)	Longitudinal survey of the same students two years after graduation and five years after graduation	Individual student or HEI information is not disclosed. HEIs may obtain institutional summary reports. Public-access data files are only identified by the type of record (HEI type and region).
Canada	Youth in Transition Survey (YITS), 2000	Federal Government: Human Resources Development Canada (HRDC)	Statistics Canada (StatCan) for HRDC	Answers are collected by StatCan offices through Computer Assisted Telephone Interviews (CATI)	Every two years	Longitudinal survey of two age cohorts, ages 15 and 18-20	Aggregated data is publicly available. No identifiable data about individuals is disclosed.
Mexico	Exámen Nacional de Ingreso al Posgrado (EXANI-III), 1997	Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES),	Centro Nacional de Evaluación para la Educación Superior (CENEVAL)	Administered by CENEVAL on HEI campus.	Individual students or institutions sign up for pre-set national testing dates.	Single testing	Individual student results are disseminated electronically and can be consulted by the students themselves and by the institution they applied for.
Mexico	Exámen General Para el Egreso de la Licenciatura (EGEL), 1994	Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES)	Centro Nacional de Evaluación para la Educación Superior (CENEVAL)	Administered by CENEVAL on HEI campus.	Individual students or institutions sign up for pre-set national testing dates.	Single testing	Results are confidential. HEIs receive institutional summary reports. Individual students receive bulletins indicating their absolute test score. Certificates of achievement are provided for students scoring at or above national average

EDU/WKP(2008)2

Mexico	Exámenes Generales para el Egreso del Técnico Superior Universitario (EGETSU), 2000	Coordinación General de Universidades Tecnológicas (CGUT)	Centro Nacional de Evaluación para la Educación Superior (CENEVAL)	Administered by CENEVAL on HEI campus.	Institutions sign up for pre-set national testing dates.	Single testing	Individual student results are confidential. HEIs receive institutional summary reports. Individual students receive bulletins indicating their absolute test score. Certificates of achievement are provided for students scoring at or above national average
UK	Destinations of Leavers from Higher Education (DLHE), 2002 (replaced the "First Destination Supplement")	Federal Government: Department for Education and Skills (DfES) and other government bodies	Higher Education Statistics Agency (HESA) commissioned expert group	Questionnaires are supplied by HESA and administered by the Careers Offices of each HEI.	Annually	Single testing	Aggregate results are publicly available on the internet. HEIs receive institutional summary reports.
USA	Collegiate Assessment of Academic Proficiency (CAAP), 1988	Not applicable	ACT	HEIs order materials and administer them to students	Flexible	Multiple forms of each module allow for pre- and post-testing (cross-sectional or longitudinal assessment)	Institutional summary report, student roster reports, student score reports, certificates of achievement for students scoring at or above national average, up to three previously specified subgroup reports (e.g. by gender, ethnicity or major). Sub-scores are given for each test component.
USA	Measure of Academic Proficiency and Progress (MAPP), 2006 (replaced the ETS "Academic Profile" test, 1992-2006)	Not applicable	Educational Testing Service (ETS), The College Board	HEIs order materials and administer them to students	Flexible	Multiple forms allow for pre- and post-testing (cross-sectional or longitudinal assessment)	Student roster reports, institutional summary score reports. Sub-scores are given for each component of the test. National data available by class level and by Carnegie classification (institution type). MAPP scores are fully comparable to Academic Profile scores.
USA	Tasks in Critical Thinking, 1992 (discontinued)	New Jersey Department of Higher Education	Educational Testing Service (ETS)	HEIs order materials and administer them to students	Flexible	Single testing	Scores were reported as the percentage of students demonstrating proficiency in each of the three skill areas. No national data available.

USA	Major Field Tests, 1990 (based on the GRE Subject Tests)	Not applicable	Educational Testing Service (ETS)	HEIs order materials and administer them to students	Flexible	Single testing	Institutional reports include individual proficiency scores, departmental summary with department mean-scaled scores, and demographic information. Percentile tables for all seniors taking the current form of each test are published each year. Departments may purchase sub-scores or group assessment indicators that the tests may support.
USA	Collegiate Learning Assessment (CLA), 2002	Major Funders: Carnegie Corporation of New York, William and Flora Hewlett Foundation, Ford Foundation, Lumina Foundation, Christian A. Johnson Endeavor Foundation, Teagle Foundation	Council for Aid to Education (CAE)	Test is administered over the Internet to students who take the measures online at an Internet-enabled site on their campus.	Annually	Most HEIs chose cross-sectional testing. Longitudinal testing is possible and recommended	Students receive their score reports online. HEIs may receive a copy of individual reports with the student's permission. Institutional reports for participating HEIs. Aggregated results are available in comparison with "similarly situated" students (in terms of SAT scores), with similar sectors (Carnegie classification) and across the national higher education system as a whole. Online database with comparative national data (HEIs remain anonymous).
USA and Canada	National Survey of Student Engagement (NSSE), 2000 (in Canada since 2004)	The Pew Charitable Trust	National expert team chaired by Peter Ewell (NCHEMS)	Questionnaires are sent out by NSSE (a joint venture between the Indiana University and NCHEMS)	Annually	Cross-sectional survey of freshmen and senior students.	Institutional reports for participating HEIs. National comparisons by academic level and by Carnegie classification (institution type) publicly available.
Test centres in the USA, Canada and other countries	Graduate Record Examination (GRE) General Test, 1966	Graduate Record Examinations (GRE) Board (independent board, affiliated with the Association of Graduate Schools (AGS) and the Council of Graduate Schools (CGS)).	Educational Testing Service (ETS)	Computer-based or paper-based test centres in the USA, Canada, and other countries	Flexible	Single testing	Score reports are sent to the student and to up to four designated institutions

3.3 Assessment Design

This section provides an overview of the technical features of the assessments, such as type of instrument, format, number of items, and duration of assessment.

3.3.1 Type of instrument

Table four distinguishes between direct assessments of knowledge and skills and indirect assessments of learning, such as self-reporting by students. These two types of instruments differ in a number of characteristics, especially in their "relative distance from the construct of learning" (Ewell, 2005, p.21). The choice of the instrument largely depends on the type of results the assessment seeks to obtain.

Direct assessments of knowledge and skills are the most obvious instrument to measure learning outcomes on a large scale: They can yield easily comparable information and the conditions of assessment can be carefully controlled to make results and comparisons between groups externally credible (Ewell, 1991). Direct assessments are traditionally focussed on *cognitive* learning outcomes, such as the EGEL (Mexico), the ENADE (Brazil), and the various tests available in the United States. Interestingly, however, the GSA examination in Australia includes a non-cognitive component on "interpersonal understanding". This component assesses students' ability to make subtle inferences about roles, relationships, behaviours, feelings, attitudes and motives, as well as the capacity to recognize potentially appropriate responses to complex interpersonal problems.

The standardized direct assessments considered in this study are normally applied outside the normal instructional context. But a variant of using standardized assessments is to embed them into the normal curriculum. In the USA, standardized tests, such as the MFT are sometimes integrated into capstone courses. Such courses require students to demonstrate cumulative and integrative learning of what has been taught in different courses during the years leading up to that course (Maeroff, 2006). The advantage of embedding standardized assessment within capstone courses is that testing occurs automatically as a requirement of the curriculum so that students do not have to be specifically motivated. Also, if the test is embedded into the normal curriculum, it is more likely to be representative of the content that is actually taught (Ewell, 2005). On the other hand, some authors have voiced concerns that organizing courses around a standardized test may reduce the diversity of learning environments, as programmes may concentrate excessively on teaching to the test (Maeroff, 2006).

Self-assessments, on the other hand, do not examine the learning itself, but focus on *indirect* evidence of learning. Surveys may ask students or graduates to rate their own learning outcomes or to report on their medium and long-term behavioural outcomes. The NSSE (USA) for example measures the ways in which students become involved in various learning activities. The CEQ (Australia) asks students to rate in how far courses helped them to develop knowledge, skills and interpersonal understanding.

Graduate student surveys can also be a means to gather information on secondary proxies of learning, such as employment rates and further study. Information provided by students through surveys such as the NGS (Canada) and the DLHE (UK) can be linked to existing statistical records held by institutions or government agencies such as Statistics Canada or the Higher Education Statistics Agency (UK). This combined data can provide useful information about the relationships between education and employment.

3.3.2 Format

Among the direct assessments presented in Table four, one can distinguish between metric-based and open-ended formats. Most direct assessment instruments combine both formats.

Metric-based assessments automatically produce quantitative scores because the different items composing them have clearly established answer schemes (Ewell, 2005). The best example is the standardized multiple-choice test used in most assessments including the GSA (Australia), EXANI-III (Mexico), EGEL (Mexico), EGETSU (Mexico), CAAP (USA), MAPP (USA) and MFT (USA). For metric-based tests, assessment designers only have to construct the test items and scoring scheme, but they do not become involved with rating student answers. Such assessments provide highly comparable results. Their downside is that they can only test a narrow range of learning outcomes.

Open-ended assessments demand students to construct written answers or to provide individual results that cannot always be automatically graded. Such assessments then rely on experts who make judgements about the quality of individual answers. Although scoring guides and examples of performance levels generally accompany such assessments, they are subject to normal variability in human judgement (Ewell, 2005). On the other hand, such assessments have the advantage of capturing a wider range of learning outcomes, such as written expression, synthesis and organisation skills. Examples include writing tasks and essay questions used by GSA (Australia), ENADE (Brazil), CAAP (USA) and MAPP (USA).

3.3.3 Standards of quality used in assessments

Criterion-referenced assessment. Assessments that make judgements about absolute levels of performance are referred to as *criterion-referenced*. Only very few assessments apply absolute criteria and benchmarks for what constitutes “quality” or “mastery”. The two Mexican graduation exams (EGEL and EGETSU) are some of the few examples that measure attainment against pre-established criteria of desired performance, corresponding to specific test scores. Various experts emphasize the difficulty of establishing such specifications for higher education tests given the complexity of the competencies required by careers that call for a university degree (Verhine and Dantas, 2005).

In practice, criterion-referenced assessments such as EGEL and EGETSU often set a “minimum standard” of performance, rather than setting a standard of excellence. Minimum standards are frequently used for professional certification and licensing examinations. They establish what constitutes the “least amount” of knowledge and abilities necessary to prove competence in a specific domain. For example, the EGEL (Mexico) for medicine students tests the knowledge and abilities that *every young doctor* who just graduated needs to have (Hughet, 2000). The idea of such assessment is to avoid that there are institutions whose graduates perform below the essential minimum (Hughet, 2000).

Norm-referenced assessment. Most standardized assessments in higher education classify those being assessed based on a comparison among them (Verhine and Dantas, 2005). Such assessment is referred to as *norm-referenced*. Norm-referenced results have meaning only in comparison with the results of other students or the same students at different points of time (ETS, 2006). They do not reflect proficiency of HEI graduates on absolute standards, but only their relative standing (Pascarella and Terenzini, 2005). According to McMillan (1988), it is difficult to value or evaluate such relative proficiencies: there may be relative gains in knowledge and abilities, but the absolute level of achievement may still be less than adequate.

Table 4: Assessment design

Country	Test Name, Introduction Date	Instrument	Format	Number of items	Duration of Assessment
Australia	Graduate Skills Assessment (GSA), 2000	Direct assessment	Multiple-choice test and writing tasks	One multiple-choice test, two writing tasks (one reporting task and one argument task)	Three hours. (2 hours for the multiple-choice test and one hour for writing tests)
Australia	Course Experience Questionnaire (CEQ), part of the Graduate destination survey since 1993	Survey	Questionnaire	25 items asking students to rate their course satisfaction and generic skills development	Five to ten minutes
Australia	Graduate Destination Survey (GDS), 1972	Survey	Questionnaire	19 items on employment, 6 items on further study	Five to ten minutes
Brazil	Exame Nacional de Cursos (ENC or "Provão"), 1995-2003	Direct assessment	Information not available	Information not available.	Four hours
Brazil	Exame Nacional de Desempenho dos Estudantes (ENADE), 2004	Direct assessment	Objective questions and essay questions	30 field-specific questions and ten general study questions	Four hours
Canada	National Graduate Survey (NGS), 1978. Follow-up Survey of Graduates, 1987	Survey	Computer-assisted telephone interview (CATI)	The CATI questionnaire is 113 pages long and contains 18 sections. Most respondents answer only a portion of the questions within each section. Some respondents skip entire sections that are not applicable to them.	Information not available
Canada	Youth in Transition Survey (YITS), 2000	Survey	Computer-assisted telephone interview (CATI)	Information not available	Information not available
Mexico	Exámen Nacional de Ingreso al Posgrado (EXANI-III), 1997	Direct assessment	Multiple-choice test	120 test items. 54% (66 items) test general intellectual abilities, 46% (54 items) test competencies in information use	Four hours
Mexico	Exámen General Para el Egreso de la Licenciatura (EGEL), 1994	Direct assessment	Multiple-choice test	Depending on the subject area, 100 to 300 items.	Several sessions on a weekend (in total, between eight and ten hours)

Mexico	Exámenes Generales para el Egreso del Técnico Superior Universitario (EGETSU), 2000	Direct assessment	Multiple-choice test	250 test items. 35% (87 items) constitute the general area component. 65% (163 items) constitute the specific area component.	Six hours (two sessions of three hours each)
UK	Destinations of Leavers from Higher Education (DLHE), 2002 (replaced the "First Destination Supplement")	Survey	Questionnaire	56 items	Approximately 15 minutes
USA	Collegiate Assessment of Academic Proficiency (CAAP), 1988	Direct assessment	Multiple-choice test and essay questions	Users can choose among six different skill modules. Each module has up to 72 questions. Users may add up to nine additional items.	40 minutes for each module
USA	Measure of Academic Proficiency and Progress (MAPP), 2006 (replaced the ETS "Academic Profile" test, 1992-2006)	Direct assessment	Multiple-choice test, optional essay question	Long form contains 108 multiple-choice questions and takes 100 minutes. Short form contains 36 questions. Optional essay is available.	Two standard forms (two-hour tests) and six abbreviated forms (40-minute tests)
USA	Tasks in Critical Thinking, 1992 (discontinued)	Direct assessment	Open-ended questions and performance-based problem-solving tasks	Information not available	90 minutes for each task
USA	Major Field Tests (MFT), 1990 (based on the GRE Subject Tests)	Direct assessment	Multiple-choice test	Information not available	Two hours (three hours for MBA)
USA	Collegiate Learning Assessment (CLA), 2002	Direct assessment	Performance-based tasks	Information not available	90 minutes
USA and Canada	National Survey of Student Engagement (NSSE), 2000 (in Canada since 2004)	Survey	Questionnaire	About 90 questions in five broad areas of engagement	Approximately 15 minutes
Test centres in the USA, Canada and other countries	Graduate Record Examination (GRE) General Test, 1966	Direct assessment	Multiple-choice test and writing tasks	Information not available.	Up to three hours (for the computer-based test)

3.4 Participants

This section summarizes the characteristics of assessment participants, providing information on the target population, the selection of participants, the coverage of test application and the incentives for students and institutions to participate in the assessment.

3.4.1 Target populations

Table five shows that most assessment instruments considered in this paper are geared towards assessing student outcomes at the end of a period of learning. Senior students near completion of their undergraduate degree are the most frequent target group for direct program/institution-level assessments. Examples are the MAPP (USA), the GSA (Australia), and the ENADE (Brazil). An advantage of assessing learning outcomes at undergraduate level is that at this level many programmes share common objectives, especially in terms of developing generic academic skills. Assessment of post-graduate learning outcomes is more difficult to design as it requires highly specified and differentiated assessment instruments. Moreover, assessments of undergraduate learning concern a larger portion of society as enrolment is higher than in more advanced levels of study.

There are few examples of program/institution-level assessments of individuals after graduation. This is probably due to the fact that it is difficult to organise large-scale exams outside the university context and to motivate those who have already graduated for test participation. Where direct assessments exist for students after graduation, they are usually focussed on individual results and involve high stakes for the participant, such as graduate school admission tests or certification exams. The EGEL exam in Mexico is open for students both before and after graduation, as long as they have covered 100% of their credits. The possibility of receiving an additional nationally recognized qualification may motivate individuals to participate even if they have already received their graduation diploma.

Surveys concerned with course and campus experiences may target either students or graduates. The NSSE (Canada and USA) is distributed to both freshmen and senior students, while the CEQ in Australia is sent out to graduates four months after graduation. Surveys concerned with secondary indicators of student learning such as employment and further study obviously target individuals some time after graduation. Examples include the GDS (Australia), the NGS (Canada), and the DLHE (UK).

3.4.2 Selection of participants

General assessment of the entire target population through direct examinations is possible only if the assessment is made mandatory or if it involves high stakes for each individual. The only example of direct general assessment of the entire student population was the ENC-Provão in Brazil, which was discontinued in 2003. The application of tests to such a large population requires very considerable human and financial resources. General assessment may be more appropriate for indirect measurements such as surveys because they are less expensive to administer. Various graduate surveys (NGS in Canada, DLHE in the UK) send out questionnaires to the entire target population.

Sampling is an assessment method that allows obtaining accurate estimates of student performance at much lower cost, by applying tests to a representative sample of the overall target population. In order to generate accurate evidence, samples must be of sufficient size and proportionally represent different groups of the student population. Sampling is obviously less costly than general testing. But it may lack credibility for some stakeholders (Terenzini, 1989). In Brazil, the sampling approach was criticised by some stakeholders because it may lead to distortions when institutions list only those candidates who are best prepared for the test (Verhine and Dantas, 2005).

3.4.3 Incentives

A common problem of assessments aiming to measure programme quality or institutional performance is the lack of incentives for students to perform at their best. In most of these assessments students do not have a direct stake in the outcomes. Although students may ultimately be negatively affected if their HEI scores badly on such evaluations, the test scores may not *directly* matter to them.

Different attempts have been made to motivate student to participate in the exams. The Provão (Brazil) made exam participation a requirement for graduation. It is interesting to note, however, that showing up for the exam was mandatory, but answering the questions was not (Verhine and Dantas, 2005). Consequently, many students left their exam sheet blank or did not try very hard (Verhine and Dantas, 2005). In the UK, it has been tried to motivate students to participate by promising a £1 donation for charity for each participant in the new longitudinal DLHE study, which is currently being piloted (HESA website, 2006). Several HEIs using the CLA report offering students small honoraria or gifts to motivate them for participation.

One way of incentivising student performance in such assessments is to involve some positive stakes for students, such as a possible award or certification. The ENADE (Brazil) offers awards and scholarships to the students who perform best in the exam. Government bodies in Australia and Mexico have designed standardised examinations (GSA in Australia, EGEL / EGETSU in Mexico) as nationally-recognized additional qualifications which students can add to their curriculum vitae when they apply for a job or for graduate school entry. The Australian and Mexican governments encourage employers and graduate schools to require the test results as evidence of student performance.

However, involving high stakes for students in program/institution-level assessments may also introduce a motivation bias. If the exam is important for the students' career, they will specifically prepare for the test, so that test results may become more a reflection of the degree to which students prepare individually for a specific examination rather than of the degree to which the HEI experience has enhanced their learning.

Table 5: Assessment participants

Country	Test Name, Introduction Date	Target Population	Selection of participants	Coverage of test application	Incentives for Participation	
					For test takers	For institutions
Australia	Graduate Skills Assessment (GSA), 2000	Students at entry level and students at graduation level	Limited, self-selected sample (Voluntary assessment)	About 2000 students from about 20 universities from a variety of fields of study participate each year.	Students may add their GSA score to their curriculum vitae when they apply for a job.	At entry level: results help identify poorly performing students to follow up and offer assistance. At graduation level: results may be used as an additional criterion for graduation or for entry into graduate courses. Results provide information on "value-added" by the institution and on general education learning quality across courses and HEIs
Australia	Course Experience Questionnaire (CEQ), part of the Graduate destination survey since 1993	Graduates who completed requirements for any higher education qualification in the previous calendar year	General assessment of the target population	Surveys are sent to all recent graduates of Australian HEIs (including international students). Response rates are around 60-65%.		HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes.
Australia	Graduate Destination Survey (GDS), 1972	All graduates who completed requirements for any higher education qualification in the previous calendar year	General assessment of the target population	Surveys are sent to all recent graduates of Australian HEIs (including international students). Response rates are around 60-65%.		HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes.

Brazil	Exame Nacional de Cursos (ENC or "Provaão"), 1995-2003	Students at graduation level	General assessment of the target population	All graduating students from courses within previously defined areas of study. In 2003, testing covered more than 460,000 students enrolled in 6,500 courses totalling over 70% of all graduating students	Participation was mandatory: Reception of graduation diploma was conditional to participation. Awards were given to students with the best performances.	Participation was mandatory. Courses and institutions receive feedback regarding the performance of their students compared to students in the same field enrolled in other courses/HEIs.
Brazil	Exame Nacional de Desempenho dos Estudantes (ENADE), 2004	Students at entry level (having covered between 7% and 22% of the curriculum) and students at graduation level (having covered at least 80% of the curriculum) within previously defined areas of study	A random sample is selected from a list including all eligible students. If courses have less than 20 students, all students are tested.	In 2004, the randomly selected sample included 140,340 students enrolled in 2,184 courses, totaling to 51% of freshmen students and 69% of graduating students. 2,830 students that were not randomly chosen signed up for voluntary participation	Participation is mandatory for the randomly sampled students: Reception of graduation diploma is conditional to participation. Awards are given to students with the best performances.	Participation is mandatory. Courses and institutions receive feedback regarding the performance of their students in absolute and relative terms. As opposed to the "Provaão", ENADE does not divulge results via the media. Results are communicated more discretely through bulletins on the internet.
Canada	National Graduate Survey (NGS), 1978. Follow-up Survey of Graduates, 1987	Graduates from all public HEIs. NGS: two years after graduation. Follow-up survey: five years after graduation (same respondents)	A sample is selected using a stratified design to provide accurate estimates by province, program, and field of study	Initial sample sizes range from 35,000 to over 40,000 graduates. Participation is voluntary. The average response rate is around 65%		Individual HEIs are not the focus of assessment, thus no possible negative impact. HEIs receive feedback on their graduates' labour market outcomes. HEIs may analyse and report on their results for internal purposes
Canada	Youth in Transition Survey (YITS), 2000	Age cohorts 15 and 18-22 of the general population (not restricted to students and graduates)	Stratified multi-stage sample design based on the use of the Canadian Labour Force Survey sample	Participation is voluntary. The initial sample size was 29,000 persons. The response rate for the 18-20 cohort is approximately 80%		Not applicable

EDU/WKP(2008)2

Mexico	Exámen Nacional de Ingreso al Posgrado (EXANI-III), 1997	Graduates applying for entry into graduate study programmes	General assessment of all students applying for entry into graduate schools that use the test	All graduates applying for admission or scholarships in graduate schools that use the test. 13,604 applicants took the test in 2006.	Participation is mandatory for entry into some graduate programmes. Allows students to provide evidence of their proficiency to apply for scholarships.	Individual HEIs are not the focus of assessment, thus no possible negative impact. Graduate HEIs receive comparable results on student performance.
Mexico	Exámen General Para el Egreso de la Licenciatura (EGEL), 1994	Students from non-technical HEIs having covered 100% of the curriculum (graduated or not)	HEIs or individual students can sign up for voluntary participation. If HEIs sign up for assessment, they decide on size and characteristics of the student sample.	Information not available.	Gives a complementary and nationally-comparable qualification to students. Certificates of achievement are provided for students scoring at or above national average	HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes. Results may be used as an additional criterion for graduation or for higher-level course entry
Mexico	Exámenes Generales para el Egreso del Técnico Superior Universitario (EGETSU), 2000	Students from technical HEIs having covered 100% of the curriculum (graduated or not)	General assessment of all eligible candidates	All graduate level students of all 48 technical HEIs throughout the country	Gives a complementary and nationally-comparable qualification to students. Certificates of achievement are provided for students scoring at or above national average	HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes. Results may be used as an additional criterion for graduation or for higher-level course entry
UK	Destinations of Leavers from Higher Education (DLHE), 2002 (replaced the "First Destination Supplement")	Recent graduates from publicly-funded HEIs who obtained a relevant qualification and who studied full-time or part-time (approximately six months after graduation)	General assessment of all eligible candidates	Questionnaires are sent to all eligible students. In 2002/03, 77% of the full-time qualifiers (251,300 students) and 70% of the part-time qualifiers (60,900 students) responded.		Individual HEIs are not the focus of assessment, thus no possible negative impact. HEIs receive feedback on their graduates' labour market outcomes. HEIs may analyse and report on their results for internal purposes

USA	Collegiate Assessment of Academic Proficiency (CAAP), 1988	All types of students	States or HEIs decide whether to assess students and determine the size and characteristics of the student sample.	Between 1988 and 2001, the test has been used by more than 600 HEIs and more than 450,000 students	Incentives for test takers may be provided by some HEIs. Motivation questions on the objective tests help determine how seriously students took the test.	HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes. Results may provide information on "value-added" by the institution and on general education learning quality across courses and HEIs.
USA	Measure of Academic Proficiency and Progress (MAPP), 2006 (replaced the ETS "Academic Profile" test, 1992-2006)	All types of students	States or HEIs decide whether to assess students and determine the size and characteristics of the student sample.	The Academic Profile (1992-2006) on which the MAPP is based has been used by 375 HEIs and 1 million students	Incentives for test takers may be provided by some HEIs.	HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes. Results may provide information on "value-added" by the institution. Results show student performance in general education compared to students from other courses/HEIs.
USA	Tasks in Critical Thinking, 1992 (discontinued)	All types of students	States or HEIs decide whether to assess students and determine the size and characteristics of the student sample.	Between 1992 and 2001, the test has been administered at 35 institutions to 200-500 students at each institution	Incentives for test takers may be provided by some HEIs.	HEIs receive feedback on their students' results. HEIs may analyse and report on their results for internal purposes. Results may provide information on "value-added" by the institution. Results show student performance in critical thinking compared to students from other courses/HEIs.
USA	Major Field Tests, 1990 (based on the GRE Subject Tests)	Senior students (4-year colleges)	States or HEIs decide whether to assess students and determine the size and characteristics of the student sample.	More than 500 colleges and universities employ the test per year. In the 1999-2000 academic year, more than 1,000 departments from 606 HEIs administered the test to nearly 70,000 students.	Test is often given as a capstone course or in the last semester of study as part of a graduation requirement.	HEIs may incorporate the assessment into their course syllabi and make the exam a graduation requirement. HEIs may analyse and report on results for internal purposes. Results show student performance in the specific area of study compared to students from other HEIs.

EDU/WKP(2008)2

USA	Collegiate Learning Assessment (CLA), 2002	Freshmen and senior students	A sample of 100 freshmen and 100 seniors per year and HEI. The sampling design involves administering separate components of the full set to different randomly selected sub-samples of students ("matrix-sampling"), thereby minimizing testing time yet allowing complete coverage of different instruments and content areas.	134 colleges have participated between 2002 and 2005	Some HEIs have provided participants with gifts, prizes or a small honorarium.	HEIs receive feedback on their results. Results can be aggregated by type of instruction, gender, race/ethnicity, social economic status or other criteria, thus allowing to improve services to raise minority success. Results provide information on "value-added" by the institution. HEIs remain anonymous on the CLA Database.
USA and Canada	National Survey of Student Engagement (NSSE), 2000 (in Canada since 2004)	First-year and senior students (4-year colleges)	A random sample is selected from a list including all first-year and senior students.	Since 2000, almost 1,000 different North American universities and colleges have administered NSSE to more than 1,160,000 students. Minimum sample sizes are determined by undergraduate enrolment (sample sizes vary between 450 and 1,000 students per HEI)		HEIs receive feedback on their students' results. HEIs may analyse and report their results for internal purposes. Results allow to identify aspects of the undergraduate experience that can be improved through changes in policies.
Test centres in the USA, Canada and other countries	Graduate Record Examination (GRE) General Test	Graduates applying for entry into graduate or professional schools	General assessment of all students applying for entry into graduate or professional schools that require the test results	Information not available	Participation is mandatory for entry into many graduate programmes in the USA and other English-speaking countries	Individual HEIs are not the focus of assessment, thus no possible negative impact. Graduate HEIs receive comparable information on student performance.

3.5 Assessment purpose and results

This section looks at the purpose of assessments to analyse whether they assess individual students, programs, institutions or education systems. It describes the type of results that are being sought and the ways in which these results are used by different stakeholders.

3.5.1 Focus of assessment

All learning outcomes assessments use the results of individual students as primary data. However, depending on the purpose of assessment, the student may not be the primary focus of assessment (Terenzini, 1989). A distinction can be made between those tests that are geared towards individual-level assessment and those that are geared towards group-level assessment. Many of the instruments combine several levels of assessment.

Individual-level assessment focuses primarily on the results achieved by each individual student, and is less interested in group performance. Postgraduate entry tests such as the GRE (USA) and the EXANI-III (Mexico), for example, consider the performance of each individual student in order to facilitate decisions on student selection. Individual-level assessments may also be geared towards identifying poorly performing students so that HEIs can follow up and offer special assistance. The GSA in Australia for example has such an individual-level component, although it also serves to yield aggregate results.

It has been proposed that individual-level assessments such as postgraduate entry tests could be used to draw conclusions about the quality of the undergraduate studies students have pursued. However, individual-level assessments are generally concerned with the cumulative outcomes of student learning obtained through the entire educational pathway. Such exams are less concerned with the degree to which a particular program or institution contributed to these learning outcomes.

Group-level assessment focuses on aggregated results. It does not consider the results of individual students as ends in themselves but uses them to establish average group-level performance indicators. The term “group” may refer to different ways of aggregating student outcomes, such as by program, institution or educational system. Group-level assessments may be used to demonstrate in how far the program or institution has made a difference in student learning outcomes.

Program-level assessment looks at performance variations among students who pursue the same program at different institutions. Program-level assessment often focuses on achievements in domain-specific knowledge and skills. Such assessment allows for comparisons between programmes that use different instructional approaches to achieve the same learning outcomes. Examples include field-specific exams such as the ENC-Provão in Brazil (discontinued in 2003) or the MFT (USA), which establish performance comparisons between students in the same area of study, but enrolled in different programmes or institutions.

Institution-level assessment looks at aggregate learning outcomes of students enrolled at different institutions, independently of the particular program they are enrolled in. Assessments may test generic academic abilities across programmes, or they may include domain-specific questions for each program. Examples include the GSA in Australia, the ENADE in Brazil, and various tests in the USA

System-level assessment looks at country-level variations between student performances. It focuses on the impact of system-level factors including societal factors, structural characteristics and educational policy frameworks on student performance. Only one assessment instrument included in this paper is used at HEIs in more than one national system: the National Survey of Student Engagement (NSSE) is administered at HEIs in both the USA and in Canada. System-level differences are reported on the web and in the national media, comparing between student engagement in Canada and the USA. System-level

comparisons could potentially be drawn from graduate admission tests, in which students from a very large number of educational systems participate, such as the GRE (USA).

3.5.2 Use of assessment findings

Formative assessment aims to yield results that are immediately useful in guiding subsequent action (Ewell, 1991). Volkwein (2003, p.7) describes formative assessment as a “feedback loop in which teaching influences learning, learning influences outcomes, and assessment of outcomes is used to influence/improve teaching and, ultimately, learning.”

At the individual level, formative assessment involves giving regular feedback to students about their progress and development in terms of knowledge and abilities (Terenzini, 1989). Teachers also use assessment to detect individual development needs of students and to better respond to these needs through adequate teaching design. The GSA (Australia) is a good example of such formative assessment use: entry level examinations are used to identify students who perform poorly, so that special assistance and follow-up can be offered.

At the program or institutional level, formative assessment may facilitate program modifications and increased teaching effectiveness. Formative assessment on the HEI level is based on the assumption that systematic self-assessment can lead to continuous institutional learning and improvement (Astin, 1977; Bowen, 1977; Ewell, 1984). HEIs can use commercially available assessment tools for the purpose of formative assessment. This a current practice in the United States, with tests such as the CAAP and the MAAP. Formative assessment design promotes institutional self-reflection and evidence-based program design. It allows HEIs to benchmark performance, analyse trends, evaluate individual programmes, and develop and improve the curriculum.

Summative assessment aims at responding to accountability demands of some organisationally higher authority. It involves judging performance for a decision or record (Ewell, 2005).

At the individual level, summative assessment can have a sorting function, separating qualified students from unqualified students (Terenzini, 1989). Standardized assessments may be used at various points of a program for such sorting purposes. At the individual level, summative assessment can also refer to having to prove to an employer that one is qualified for a certain position. Students use results from the GSA (Australia) and from the EGEL and EGETSU (Mexico) for such purposes.

At the program or institutional level, aggregated assessment results may be used as part of a larger national or state accountability framework. The stakes involved for programmes and institutions vary from country to country. In most countries, accreditation agencies do not take student learning outcomes into consideration. In some countries, however, HEIs are obliged to engage in some sort of local outcomes assessment as a prerequisite for accreditation and public fund reception. The administration of nation-wide or state-wide standardized assessment purely to create institutional performance indicators is still very rare (Ewell, 2005).

Only Brazil and some federal USA states make accreditation and resource allocation directly contingent on satisfactory results in standardized outcomes assessments. In other countries, such as Australia, the government may use results to benchmark national student performance, to inform policy decisions and to plan for the need of the higher education sector.

Aggregated assessment results may also be used to provide information to prospective students, their parents, and employers. Some HEIs and governments are very concerned with shifting the understanding of quality more towards a focus on effective learning and “fitness for purpose” rather than prestige and status (Baldwin, 2000). Findings from learning outcomes assessment may provide valuable evidence of HEI quality in terms of teaching and learning.

Table 6: Purpose and use of assessment results

Country	Test Name, Introduction Date	Focus of assessment	Type of results yielded	Use of assessment results
Australia	Graduate Skills Assessment (GSA), 2000	Institutions	General education results of entry-level students and graduation-level students	<p>HEIs: At entry level: identify poorly performing students to follow up and offer assistance. At graduation level: use results as an additional criterion for entry into graduate-level courses. Benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms, develop and improve curricula, determine student eligibility for upper-division studies.</p> <p>Government: Collect information on the quality of learning outcomes across HEIs for national and potentially international benchmarking of graduate skills.</p> <p>Employers: The Government promotes the test to employers and supports its use as a standard recruitment tool.</p>
Australia	Course Experience Questionnaire (CEQ), part of the Graduate destination survey since 1993	Programmes, Institutions	Graduate satisfaction with teaching and learning. Self-reported gains in academic skills related to the HEI experience.	<p>HEIs: Benchmarking, trend analysis, evaluation of programmes, curriculum development and improvement. Provide national accountability data.</p> <p>Government: Ensure quality and performance management within HEIs. Inform student choice. Assess and plan for the needs of the HE sector. Since 2005, results from the CEQ are used for performance-based incentive funding through the national "Learning and Teaching Performance Fund (LTPF)".</p>
Australia	Graduate Destination Survey (GDS), 1972	Programmes, institutions	Information on employment and further study	<p>HEIs: Benchmarking, trend analysis, evaluation of programmes, curriculum development and improvement to optimize labour market and further study outcomes. Provide national accountability data.</p> <p>Government: Ensure quality and performance management within HEIs. Inform student choice. Assess and plan for the needs of the HE sector. Since 2005, results from the GDS are used for performance-based incentive funding through the "Learning and Teaching Performance Fund (LTPF)".</p>
Brazil	Exame Nacional de Cursos (ENC or "Provaão"), 1995-2003	Programmes	Inter-institutional performance comparisons between students from the same field of study	<p>HEIs: Good scores were widely used for commercial purposes (advertisements & publicity). Results often served to mobilize students and professors to make a joint effort to maintain good scores/improve bad ones.</p> <p>Government: Since 2001, the test served as a guidance for accreditation and re-accreditation, but punitive measures were only taken in extreme cases.</p> <p>General Public: Provaão results were widely divulged by the media to inform prospective students and society at large about the quality of learning across HEIs.</p>

EDU/WKP(2008)2

Brazil	Exame Nacional de Desempenho dos Estudantes (ENADE), 2004	Programmes, institutions	Differences in cognitive results between entry-level students and graduation-level students	Students: Prove their performance according to national standards to potential employers. HEIs: Benchmarking, trend analysis, evaluation of programmes, curriculum development and improvement. Provide national accountability data. Government: ENADE is one aspect of a combination of performance indicators used for HEI evaluations. Employers: may ask for a candidate's test results as objective evidence of proficiency in the professional area.
Canada	National Graduate Survey (NGS), 1978. Follow-up Survey of Graduates, 1987	Programmes and institutions	Information on the integration of graduates into the labour market. Links between education and labour market outcomes.	HEIs: Benchmarking, trend analysis, evaluation of programmes, curriculum development and improvement to optimize labour market and further study outcomes. Provide national accountability data. Government: Assess and plan for the needs of the HE sector.
Canada	Youth in Transition Survey (YITS)	Education system	Information on how youth manage the transitions from school/HEI to the labour market	Human Resources Skills Development Canada: Collect information on the patterns of, and influences on, major transitions in young people's lives, particularly with respect to education, training and work. Aid policy and program development. Provide information to educators, social and policy analysts and advocacy groups.
Mexico	Exámen Nacional de Ingreso al Posgrado (EXANI-III), 1997	Institution types and regions	Cumulative academic ability. Mastery of the generic academic skills required to undertake graduate study	HEIs: Use individual student results to compare performance of applicants and to facilitate decision-making on student admission and/or scholarship attribution.
Mexico	Exámen General Para el Egreso de la Licenciatura (EGEL), 1994	Students	Cumulative results (specific to the area of study). Mastery of the essential knowledge and abilities required to start professional practice	Students: Prove their performance according to national standards to potential employers. HEIs: Benchmarking, trend analysis, evaluation of programmes, curriculum development and improvement. Some HEIs use results as an additional criterion for certification or graduation. Employers: may ask a candidate's test results as objective evidence of proficiency in the professional area.
Mexico	Exámenes Generales para el Egreso del Técnico Superior Universitario (EGETSU), 2000	Students, Programmes, or Institutions	Cumulative results (specific to the area of study). Mastery of the essential knowledge and abilities required to start professional practice	Students: Prove their performance according to national standards to potential employers. HEIs: Benchmarking, trend analysis, evaluation of program analysis, curriculum development and improvement. Results may be used by individual HEIs as a criterion for certification or graduation. Employers: may ask a candidate's test results as objective evidence of proficiency in the professional area.

UK	Destinations of Leavers from Higher Education (DLHE), 2002 (replaced the "First Destination Supplement")	Programmes, Institutions	Information on employment and further study	HEIs: Benchmarking, trend analysis, evaluation of programmes, curriculum development and improvement to optimize labour market and further study outcomes. Provide national accountability data. Government: Assess and plan for the needs of the HE sector.
USA	Collegiate Assessment of Academic Proficiency (CAAP), 1988	Students, Programmes, institutions	Depending on simple or value-added administration: cumulative outcomes in general academic skills or growth in general academic outcomes while at college	HEIs: Satisfy accreditation and accountability reporting requirements, benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms, develop and improve curricula, determine student eligibility for upper-division studies. States: the CAAP has been used to track entire systems over a period of time, using test results as a benchmark for progress within a public accountability framework. State-wide results may be published in a "public report card" and compared with the results of other states.
USA	Measure of Academic Proficiency and Progress (MAPP), 2006 (replaced the ETS "Academic Profile" test, 1992-2006)	Students, Programmes, Institutions	Growth in generic academic skills while at college	HEIs: Satisfy accreditation and accountability reporting requirements, benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms, develop and improve curricula, determine student eligibility for upper-division studies, counsel individual students for academic achievement.
USA	Tasks in Critical Thinking, 1992 (discontinued)	Students, Programmes, Institutions	Performance-based measure of college-level higher order thinking skills	HEIs: Satisfy accreditation and accountability reporting requirements, benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms, develop and improve curricula, determine student eligibility for upper-division studies.

EDU/WKP(2008)2

USA	Major Field Tests, 1990 (based on the GRE Subject Tests)	Students, Programmes	Mastery of concepts, principles, and knowledge expected of students at the conclusion of an academic major in specific subject areas.	Students: Test is often given as a capstone course or in the last semester of study as part of a graduation requirement. HEIs: Scores may be used for medium to high-stakes decisions. Document proficiency in the specific area in the last semester of study to measure effectiveness of departmental curricula. Satisfy accreditation and accountability reporting requirements, benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms, develop and improve curricula, determine student eligibility for upper-division studies.
USA	Collegiate Learning Assessment (CLA), 2002	Programmes, Institutions	Performance-based measure of college-level higher order thinking skills	HEIs: Provide accountability data, create institution-wide criteria for course adoption, benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms, develop and improve curricula.
USA and Canada	National Survey of Student Engagement (NSSE), 2000 (in Canada since 2004)	Programmes, Institutions	Secondary indicators of learning: Information on student participation in learning opportunities during the college experience. Self-reported gains in academic skills related to the college experience.	HEIs: Develop and improve curricula and services to enhance student engagement. Satisfy accreditation and accountability reporting requirements (NSSE provides an "Accreditation Toolkit" facilitating the use of NSSE results for regional accreditation). Benchmark and analyse trends, document/demonstrate program effectiveness and improvement over time, compare students' achievement levels with national user norms. General public: Aggregated results are publicly available and provide information about what students gain from their HEI experiences. Government: Data can be used as an indicator of institutional effectiveness in accrediting processes. Data supports national and sector benchmarking processes.
Test centres in the USA, Canada and other countries	Graduate Record Examination (GRE) General Test, 1966	Students	Cumulative academic ability. Mastery of the generic academic skills required to undertake graduate study	HEIs: Compare performance of applicants and to facilitate decision-making on student admission

REFERENCES

- AAC&U (American Association of Colleges and Universities) (2002): "Greater Expectations: A New Vision for Learning as a Nation Goes to College", Washington DC, <http://www.greaterexpectations.org/>
- AAC&U (2004): "Our Students' Best Work: A Framework for Accountability Worthy of Our Mission", A Statement from the Board of Directors of AAC&U, Washington, DC
- ABET (Accreditation Board for Engineering and Technology) (1997), Engineering Criteria 2000 Third Edition: Criteria for Accrediting Programmes in Engineering in the United States, Baltimore, Maryland, <http://www.abet.org/EAC/eac2000.html>
- ACER (Australian Council for Educational Research) (2001), "Graduate Skills Assessment. A Summary Report", Occasional Paper Series, Higher Education Division, Department of Education, Training and Youth Affairs
- Adam, S. (2004), "Using Learning Outcomes. A consideration of the nature, role, application and implications for European education of employing learning outcomes at the local, national and international levels", Paper prepared for the United Kingdom Bologna Seminar, 1-2 July 2004, Heriot-Watt University, Edinburgh. Scotland.
- Allan, J. (1996), "Learning Outcomes in Higher Education", Studies in Higher Education, Vol. 21, No.1
- American Psychological Association (1985), Standards for Educational and Psychological Testing, Washington DC
- Anaya, G. (1999), "College Impact on Student Learning: Comparing the Use of Self-Reported Gains, Standardized Test Scores, and College Grades", Research in Higher Education, Vol.40, No.5
- Anderson, L.W. and Sosniak, L.A. (1994), Bloom's taxonomy: A forty-year retrospective. Ninety-third yearbook of the National Society for the Study of Education, University of Chicago Press, Chicago
- Andrich, D. (2002). "A Framework Relating Outcomes Based Education and the Taxonomy of Educational Objectives", Studies in Educational Evaluation, Vol. 28, p.35-59
- Astin, A. (1970), "The Methodology of Research on College Impact I" and "The Methodology of Research on College Impact II", Sociology of Education, No. 43
- Astin, A.W. (1984), "Student Involvement: A Developmental Theory for Higher Education", Journal of College Student Personnel, Vol. 25, No. 4
- Benjamin, R. and Chun, M. (2003), "A New Field of Dreams: The Collegiate Learning Assessment Project", Peer Review, Vol. 5, No. 4, <http://www.aacu.org/peerreview/pr-su03/pr-su03feature2.cfm/>.

- Baldwin, G. and James, R. (2000), "The Market in Australian Higher Education and the Concept of Student as Informed Consumer", *Journal of Higher Education Policy and Management*, Vol. 22, No. 2
- Bloom, B., Krathwohl, D. (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals*, by a committee of college and university examiners. Handbook I: Cognitive Domain.
- Borden, V.M. and Owens, J.L. (2001), "Measuring Quality: Choosing Among Surveys and Other Assessments of College Quality", Centre for Policy Analysis, American Council on Education, Washington, D.C.
- Bowen, H.R. (1977), *Investment in Learning: The Individual and Social Value of American Higher Education*, Jossey-Bass Publishers, San Francisco, CA
- Bratti, M., McKnight, A., Naylor, R., Smith, J. (2004), "Higher education outcomes, graduate employment and university performance indicators", *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Vol.167, No.3
- Byrne, J.V. (2006), "Public Higher Education Reform Five Years After. The Kellogg Commission on the Future of State and Land-Grant Universities", National Association of State Universities and Land-Grant Colleges And the W.K. Kellogg Foundation
<http://www.nasulgc.org/Kellogg/KCFiveYearReport.pdf>
- Chun, M. (2002), "Looking Where the Light is Better: A Review of the Literature on Assessing Higher Education Quality", peerReview, Winter/Spring
- Clewett, R.M. (1996), "Flourishing Virtues – Intellectual and Other: A Consideration of the Proper Focus of Higher Education", Conference on Values in Higher Education,
<http://web.utk.edu/~unistudy/ethics96/rmc.html>
- Dearing, R. (1997), *Higher Education in The Learning Society*, The report of the National Committee of Inquiry into Higher Education.
- Eisner, E. W. (1979), *The Educational Imagination: On the Design and Evaluation of School Programs*, Macmillan, New York, NY
- Eraut, M. (1990), "Identifying the Knowledge which Underpins Performance" in Black, H. and Wolf, A., *Knowledge and Competence: Current Issues in Education and Training*, Careers and Occupational Information Centre, Employment Department, London, UK
- Ewell, P. T. (1984), *The Self-Regarding Institution: Information for Excellence*, National Centre for Higher Education Management (NCHEMS), Boulder, CO
- Ewell, P.T. (1991), "To Capture the Ineffable: New Forms of Assessment in Higher Education", *Review of Research in Education*, Vol.17, pp.75-125
- Ewell, P. (2005), "Applying Learning Outcomes Concepts to Higher Education: An Overview, prepared for the University Grants Committee",
http://www.hku.hk/caut/seminar/download/OBA_1st_report.pdf
- Gagné, R. M. (1977), *The Conditions of Learning*, Holt, Rinehart & Winston, Inc., New York, NY

- Gormley, W.T., Weimer, D.L. (1999), *Organizational Report Cards*. Harvard University Press, Cambridge, MA.
- IHEP (Institute for Higher Education Policy) (ed.) (2007), *College and University Ranking Systems. Global Perspectives and American Challenges*, Washington, DC
- Inkeles, A. (1966), "Social structure and the socialization of competence", *Havard Educational Review*, 36
- Huguet, A. (2000), "El CENEVAL y la Evaluación Externa de la Educación en México", *Revista Electrónica de Investigación Educativa*, Vol.2, No.2
- IHEP (Institute for Higher Education Policy) (ed.) (2007), *College and University Ranking Systems. Global Perspectives and American Challenges*, Washington, DC
- Klein, P., Kuh, G. D., Chun, M., Hamilton, L., Shavelson, R. (2005), "An Approach to Measuring Cognitive Outcomes across Higher Education Institutions", *Research in Higher Education*, Vol.46, No.3
- Longman, Green New York, NY, USA
- Klein, S., Benjamin, R., Shavelson, R., Bolus, R. (forthcoming), "The Collegiate Learning Assessment: Facts and Fantasies", White Paper, Council for Aid to Education.
- Klemp, G.O. (1977), "Three Factors of Success" in Vermilye, D.W. (ed), *Relating Work and Education* San Francisco, Jossey-Bass, San Francisco, CA
- Kolb, D. A. (1981), "Learning styles and disciplinary differences", in A. W. Chickering (ed.) *The Modern American College*, Jossey-Bass, San Francisco, CA
- Kolb, D. A. (1981) "Experiential Learning Theory and the Learning Style Inventory: A Reply to Freedman and Stumpf", *Academy of Management Review*, Vol.6, No.2
- Kuh, G.D. (2001), "Assessing What Really Matters to Student Learning", *Change*, Vol.33, No.3
- Maeroff, G. (2006), "Grading Higher Education", *Beyond the Rankings: Measuring Learning in Higher Education*, Hechinger Institute on Education and the Media, Columbia University,
- Marzano, R.J. (2001), *Designing a new taxonomy of educational objectives*, Corwin Press
- McGuinness S. (2003), "University Quality and Labour Market Outcomes", *Applied Economics*, Vol.35, No.18
- McMillan, J.H. (1988), "Beyond Value-Added Education: Improvement Is Not Enough", *The Journal of Higher Education*, Vol. 59, No.5
- Melton, R. (1996), "Learning Outcomes for Higher Education", *British Journal of Educational Studies*, Vol. 44, No. 4
- Middle States Commission on Higher Education (1996), *Framework for Outcomes Assessment*, Philadelphia, PA,

<http://www.montclair.edu/pages/InstResearch/AssessmentDocs/Middle%20States/Middle%20States%20Framework%20for%20Outcomes%20Assessment%201996.pdf>

Montgomery, J. (1991), "Social Networks and Labor Market Outcomes: Toward an Economic Analysis", *American Economic Review*, Vol.8, No.5

Morgan, A. (2002), "The Utilisation of NVQs in Higher Education", *Education and Training*, Vol.44, No.2

Organisation of Economic Co-operation and Development (OECD) (2000), "RBM in the development co-operation agencies: a review of experience", *Development Assistance Committee (DAC) paper*.

Otter, S. (1992), "Learning Outcomes in Higher Education", *A Development Project Report*, Unit for the Development of Adult Continuing Education (UDACE)

Pascarella, E.T. (1987), "Are Value-Added Analyses Valuable?", in *Assessing the Outcomes of Higher Education*, Proceedings of the 1986 ETS Invitational Conference, Educational Testing Service, Princeton, NJ

Pascarella, E., Smart, J. and Smylie, M. (1992), "College tuition costs and early career socio-economic achievement: do you get what you pay for?", *Higher Education*, 24

Pascarella, E.T., Terenzini, P.T (2005): *How College Affects Students. A Third Decade of Research*, Vol. 2, Jossey-Bass, San Francisco, CA

Pike, G.R. (1999), "The Constant Error of the Halo in Educational Outcomes Research", *Research in Higher Education*, Vol.40, No.1

Posner, G.J. (1992), *Analyzing the Curriculum*, McGraw-Hill, USA

Rychen, D.S. (2004), "Key Competencies for All: An Overarching Conceptual Frame of Reference", *DeSeCo Strategy Paper on Key Competencies*, <http://www.unal.edu.co/diracad/competencias/KeyCompetences.pdf>

Sadlak, J. and Liu, N.C. (eds.) (2007), *The World-Class University and Ranking: Aiming Beyond Status*, UNESCO-CEPES, Cluj University Press, Bucharest, Romania

Salmi, J. and Saroyan, A. (2006), "League Tables as Policy Instruments: Uses and Misuses",

Schalock, R. L. (2001), *Outcome-Based Evaluation*, 2nd ed., Kluwer Academic / Plenum Publishers, New York

Seagraves, L., Kemp, I.J., Osborne, M.J. (1996), "Are Academic Outcomes of Higher Education Provision Relevant to and Deliverable in the Workplace Setting?", *Higher Education*, 32, pp.157-176

Shavelson, R.J., and Huang, L. (2003), "Responding Responsibly to the Frenzy to Assess Learning in Higher Education", *Change*, Vol.35, No.1

Spady, W. G. (1988), "Organizing for Results: The Basis of Authentic Restructuring and Reform", *Educational Leadership*, Vol. 46, No.2

Spence, A.M. (1973), "Job Market Signaling", *Quarterly Journal of Economics*, Vol.87, No.3

- Terenzini, P. (1997), "Student Outcomes Information for Policy-Making, NCES 97-991", paper prepared for the Council of the National Postsecondary Education Cooperative Working Group on Student Outcomes from a Policy Perspective, USA Department of Education / National Centre for Education Statistics, Washington, DC 1997
- Terenzini, P.T. (1989), "Assessment with Open Eyes: Pitfalls in Studying Student Outcomes", *The Journal of Higher Education*, Vol. 60, No.6
- Verhine, R.E. and Dantas, L.M. (2005), "Assessment of Higher Education in Brazil: From the Provão to ENADE", Document prepared for the World Bank, Responsible party: Alberto Rodriguez
- Volkwein, J.F. (2003), "Implementing Outcomes Assessment on your Campus", *Research and Planning E-Journal*, Vol. 1, No.1, http://www.rpgroup.org/publications/eJournal/volume_1/Volkwein_article.pdf

Electronic Resources

NC State University, University Planning & Analysis, Internet Resources for Higher Education Outcomes Assessment: <http://www2.acs.ncsu.edu/UPA/assmt/resource.htm>

ANNEX: DATA SOURCES (BY COUNTRY)

Australia

GSA (Graduate Skills Assessment)

ACER (Australian Council for Educational Research)
<http://www.acer.edu.au/tests/university/gsa/intro.html>

DEST (Department of Education, Science and Training, Australia)
http://www.dest.gov.au/sectors/higher_education/publications_resources/other_publications/graduate_skills_assessment.htm

ACER (Australian Council for Educational Research) (2001), “Graduate Skills Assessment. A Summary Report”, *Occasional Paper Series*, Higher Education Division, DETYA (Department of Education, Training and Youth Affairs)

Hambur, S. et al. (2002), “Graduate Skills Assessment. Stage One Validity Study”, ACER, Department for Education, Science and Training (DEST), Evaluations and Investigations Programme,
http://www.dest.gov.au/NR/rdonlyres/FB832CDF-A333-4026-9C55-56F3A55ACE0B/808/03_2.pdf

Murdoch University, Teaching and Learning Center, Australia,
<http://www.tlc.murdoch.edu.au/eddev/evaluation/gsa/gsareport.html#costs>

GDS (Graduate Destination Survey) and CEQ (Course Experience Questionnaire)

Graduate Careers Council of Australia, <http://www.graduatecareers.com.au/content/view/full/867>

Wilson, K.L. et al. (1997), “The Development, Validation and Application of the Course Experience Questionnaire”, *Studies in Higher Education*, Vol. 22, No. 1

Access Economics PTY Limited (2005), “Review of Higher Education Outcome Performance Indicators”, Report for the Department for Education, Science and Training (DEST),
http://www.dest.gov.au/NR/rdonlyres/0E41FC0F-F8E3-4465-9EDE-CE0FDEFCC185/6314/review_higher_ed_outcome_performance_indicators_re.pdf

Baldwin, G. and James, R. (2000), “The Market in Australian Higher Education and the Concept of Student as Informed Consumer”, *Journal of Higher Education Policy and Management*, Vol. 22, No. 2

AVCC (Australian Vice-Chancellors' Committee) and GCCA (Graduate Careers Council of Australia), “Standard Recommended Methodology for the Graduate Destination Survey, Course Experience Questionnaire and Postgraduate Research Experience Questionnaire”,
<http://www.graduatecareers.com.au/content/view/full/1456>

Brazil**ENC-“Provão” (Exame Nacional de Cursos) and ENADE (Exame Nacional de Desempenho dos Estudantes)**

INEP (Instituto Nacional de Estudos e Pesquisas Educacionais 'Anísio Teixeira'),

On ENC-Provao: <http://www.inep.gov.br/superior/provao/default.asp>

On ENADE: <http://www.inep.gov.br/superior/enade/default.asp>

Verhine, R.E. and Dantas, L.M. (2005), “Assessment of Higher Education in Brazil: From the Provão to ENADE”, Document prepared for the World Bank, Responsible party: Alberto Rodriguez

Sobrinho, D.S. (2006), “Changes in Brazilian Higher Education and Their Effects on the Enhancement of Teaching (1995-2005)”, *New Directions for Higher Education*, Vol. 2006, No.133

Wolff, L. (2004), “Educational Assessments in Latin America: The State of the Art”, *Applied Psychology: An International Review*, Vol. 53, No.2

Barreyro, G.B. (2006), “Evaluación de la Educación Superior Brasileira: El SINAES”, *Revista de la Educación Superior*, Vol. 35, No. 137, Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES)

Canada**NGS (National Graduate Survey) and Follow-up Survey of Graduates**

Statistics Canada, <http://stcwww.statcan.ca/english/sdds/5012.htm>

Statistics Canada (2003), “Selected Statistics Canada Data Sources for Research on Transitions and Learning Outcomes”, CESC-SSHRC Education Research Initiative, http://www.statcan.ca/english/rdc/pdf/data_sources.pdf

Allen, M. and Butlin, G. (2003) “Finding their way: a profile of young Canadian graduates”, *Education, skills and learning – Research papers*, Statistics Canada, <http://www.statcan.ca/english/research/81-595-MIE/81-595-MIE2003003.pdf>

YITS (Youth in Transition Survey)

StatCan (Statistics Canada): www.statcan.ca/english/sdds/4435.htm

HRDC (Human Resources Development Canada), (2000) “Youth in Transition Survey – Project Overview”, <http://dsp-psd.pwgsc.gc.ca/Collection/Statcan/81-588-X/81-588-XIE2000001.pdf>

Mexico**EXANI-III (Exámen Nacional de Ingreso al Posgrado)**

EDU/WKP(2008)2

CENEVAL (Centro Nacional de Evaluación para la Educación Superior):
<http://www.ceneval.edu.mx/portalceneval/index.php?q=info.fichas.ficha7>

EGETSU (Exámenes Generales para el Egreso del Técnico Superior Universitario)

CENEVAL: <http://www.ceneval.edu.mx/portalceneval/index.php?q=info.fichas.ficha50>

EGEL (Exámen General Para el Egreso de la Licenciatura)

CENEVAL, <http://www.ceneval.edu.mx/portalceneval/index.php?q=info.examenes>

Alvarado, R., “Sobre la Evaluación de la Educación Superior y Algunos de Sus Problemas”, CENEVAL,

Hughet, A. (2000), “El CENEVAL y la Evaluación Externa de la Educación en México”, *Revista Electrónica de Investigación Educativa*, Vol.2, No.2

UK

DLHE (Destinations of Leavers from Higher Education)

HESA (Higher Education Statistics Agency): <http://www.hesa.ac.uk/manuals/06018/dlhe0607.htm>

Bratti, M. et al. (2003), “Higher Education Outcomes, Graduate Employment and University Performance Indicators”, *Warwick Economic Research Papers*, No.692

USA

CAAP (Collegiate Assessment of Academic Proficiency)

ACT: <http://www.act.org/caap/>

MAAP (Measure of Academic Proficiency and Progress)

ETS (Educational Testing Service): <http://www.ets.org/maap>

Tasks in Critical Thinking

Borden, V.M. and Owens, J.L. (2001), “Measuring Quality: Choosing Among Surveys and Other Assessments of College Quality”, Center for Policy Analysis, American Council on Education, Washington, D.C.

Major Field Tests

ETS: <http://www.ets.org/mft>

GRE (Graduate Record Examination)

ETS: <http://www.ets.org/gre>

ETS (2005), “2005-2006 Guide to the Use of Scores”, Published for the Graduate Record Examinations Board, <ftp://ftp.ets.org/pub/gre/994994.pdf>

NSSE (National Survey of Student Engagement)

NSSE: <http://nsse.iub.edu/index.cfm>

CLA (Collegiate Learning Assessment)

Council for Aid to Education (CAE): http://www.cae.org/content/pro_collegiate.htm

CAE: “CLA in Context”, <http://www.cae.org/content/pdf/CLA.in.Context.pdf>

Klein, S., Benjamin, R., Shavelson, R., Bolus, R. (forthcoming), “The Collegiate Learning Assessment: Facts and Fantasies”, White Paper, Council for Aid to Education

USA – Reviews of Assessment Instruments:

Borden, V.M. and Owens, J.L. (2001), “Measuring Quality: Choosing Among Surveys and Other Assessments of College Quality”, Center for Policy Analysis, American Council on Education, Washington, D.C.

Ewell, P.T. (1991), “To Capture the Ineffable: New Forms of Assessment in Higher Education”, *Review of Research in Education*, Vol.17, pp.75-125

Maeroff, G. (2006), “Grading Higher Education”, *Beyond the Rankings: Measuring Learning in Higher Education*, Hechinger Institute on Education and the Media, Columbia University,

Pascarella, E.T., Terenzini, P.T (2005): *How College Affects Students. A Third Decade of Research*, Vol. 2, Jossey-Bass, San Francisco, CA

Shavelson, R.J., and Huang, L. (2003), “Responding Responsibly to the Frenzy to Assess Learning in Higher Education”, *Change*, Vol.35, No.1

Volkwein, J.F. (2003), “Implementing Outcomes Assessment on your Campus”, *Research and Planning E-Journal*, Vol. 1, No.1, http://www.rpgroup.org/publications/eJournal/volume_1/Volkwein_article.pdf

EXISTING OECD EDUCATION WORKING PAPERS

- No.1 *Teacher Demand and Supply: Improving Teaching Quality and Addressing Teacher Shortages* (2002), Paulo Santiago.
- No.2 *Teacher Education and the Teaching Career in an Era of Lifelong Learning* (2002), John Coolahan.
- No.3 *Towards an Understanding of the Mechanisms That Link Qualifications and Lifelong Learning* (2003), Friederike Behringer, Mike Coles.
- No.4 *Measuring Educational Productivity in Standards-Based Accountability Systems: Introducing the SES Return on Spending Index* (2005), Martin Hampel.
- No. 5 *PISA 2000: Sample Weight Problems in Austria* (2006), Erich Neuwirth.
- No. 6 *Funding Systems and their Effects on Higher Education Systems – International Report* (2007), Franz Strehl, Sabine Reisinger and Michael Kalatschan.
- No. 7 *On the Edge: Securing a Sustainable Future for Higher Education* (2007), OECD/IMHE-HEFCE.
- No. 8 *Globalisation and Higher Education* (2007), Simon Margison and Marijk van der Wende.
- No. 9 *Understanding the Regional Contribution of Higher Education Institutions: A Literature Review* (2007), Peter Arbo and Paul Benneworth.
- No. 10 *Effects of Tertiary Expansion – Crowding-out Effects and Labour Market Matches for the Higher Educated* (2007), Bo Hansson.
- No. 11 *Skilled Voices? Reflections on Political Participation and Education in Austria* (2007), Florian Walter and Sieglinde K. Rosenberger.
- No. 12 *Education and Civic Engagement: Review of Research and a Study on Norwegian Youths* (2007), Jon Lauglo and Tormod Oia.
- No. 13 *School Accountability, Autonomy, Choice, and the Level of Student Achievement: International Evidence from PISA 2003* (2007), Ludger Wössmann, Elke Lüdemann, Gabriela Schütz and Martin R. West.
- No. 14 *School Accountability, Autonomy, Choice, and the Equity of Student Achievement: International Evidence from PISA 2003* (2007), Gabriela Schütz, Martin R. West, Ludger Wössmann.

THE OECD EDUCATION WORKING PAPERS SERIES ON LINE

The OECD Education Working Papers Series may be found at:

The OECD Directorate for Education website: www.oecd.org/edu/workingpapers

The OECD's online library, SourceOECD: www.sourceoecd.org

The Research Papers in Economics (RePEc) website: www.repec.org

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

Go to www.oecd.org

Click on "My OECD"

Sign up and create an account with "My OECD"

Select "Education" as one of your favourite themes

Choose "OECD Education Working Papers" as one of the newsletters you would like to receive

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