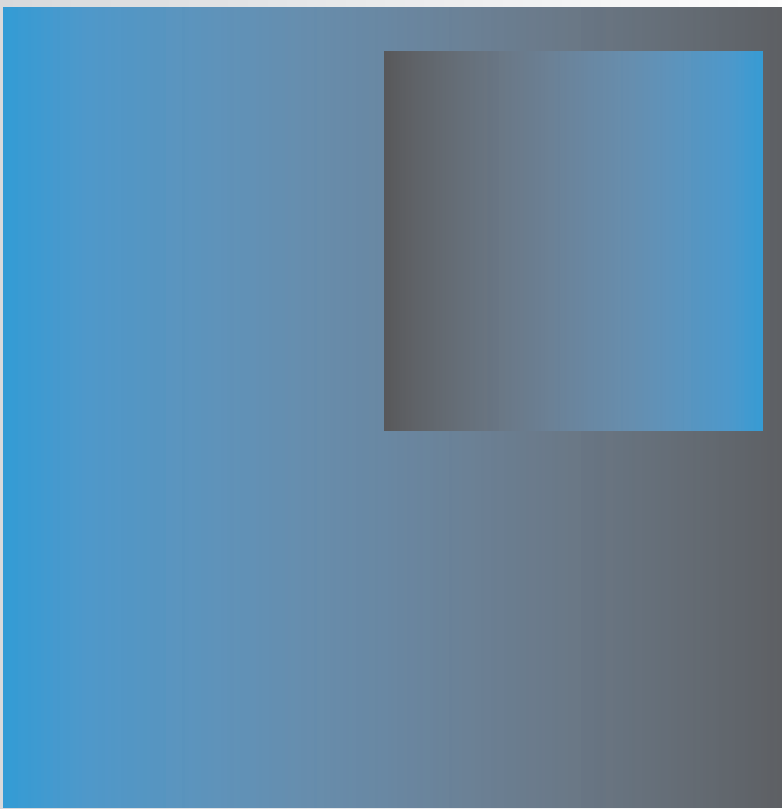
Higher Education Management and Policy

JOURNAL OF THE PROGRAMME
ON INSTITUTIONAL MANAGEMENT
IN HIGHER EDUCATION



OECD

Volume 22/1
2010

Higher Education Management and Policy

Volume 22/1

Journal of the Programme on Institutional Management in Higher Education



ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

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

ISSN 1682-3451 (print)

Corrigenda to OECD publications may be found on line at: *www.oecd.org/publishing/corrigenda*.

© OECD 2010

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

Higher Education Management and Policy

- A journal addressed to leaders, managers, researchers and policy makers in the field of higher education institutional management and policy.
- Covering practice and policy in the field of system and institutional management through articles and reports on research projects of wide international scope.
- First published in 1977 under the title *International Journal of Institutional Management in Higher Education*, then *Higher Education Management* from 1989 to 2001, it appears three times a year.

Information for authors wishing to submit articles for publication appears at the end of this issue. Articles and related correspondence should be sent to the editor:

Professor Vin Massaro
Professorial Fellow
LH Martin Institute for Higher Education Leadership and Management
Melbourne Graduate School of Education, The University of Melbourne
153 Barry Street
Carlton, Victoria 3010, Australia
E-mail contact: imhe@oecd.org

To subscribe send your order to:

OECD Publications Service
2, rue André-Pascal, 75775 Paris Cedex 16, France

2010 subscription (3 issues):
EUR 129 USD 169 GBP 90 JPN 18 300

Online bookshop: *www.oecdbookshop.org*

The Programme on Institutional Management in Higher Education (IMHE) is the OECD membership forum serving managers in higher education institutions, researchers and policy makers in national and regional authorities. IMHE provides strategic analysis and advice on institutional leadership, management, research and innovation in a global knowledge economy and reform and governance in higher education.

IMHE is the only OECD forum open to higher education institutions. Established in 1969, IMHE now has 240 members from 50 different countries. IMHE's current activities include the study of the impact of ranking on higher education, a feasibility study of assessment of higher education outcomes, reviews of higher education in regional development and a project on quality teaching in higher education.

For more information on the concrete benefits of IMHE membership, visit *www.oecd.org/edu/imhe/join*.

Editorial Advisory Group

Elaine EL-KHAWAS
George Washington University, United States (Chair)

Philip G. ALTBACH
Boston College, United States

Chris DUKE
RMIT University, Australia

John GODDARD
Newcastle University, United Kingdom

Leo GOEDEGEBUURE
University of Melbourne, Australia

Ellen HAZELKORN
Dublin Institute of Technology, Ireland

Salvador MALO
Instituto Mexico de la Competitividad, Mexico

Vin MASSARO
University of Melbourne, Australia

V. Lynn MEEK
University of Melbourne, Australia

Robin MIDDLEHURST
University of Surrey, United Kingdom

José-Ginés MORA
University of London, United Kingdom

Jan SADLAK
School of Social Psychology and Humanities/SWPS, Poland

Jamil SALMI
The World Bank, United States

Michael SHATTOCK
University of London, United Kingdom

Sheila SLAUGHTER
University of Georgia, United States

Andrée SURSOCK
European University Association, Belgium

Ulrich TEICHLER
INCHER-Kassel, Germany

Luc WEBER
Université de Genève, Switzerland

Akiyoshi YONEZAWA
Tohoku University, Japan

Frank ZIEGELE
Centre for Higher Education Development, Germany

Table of Contents

The state of learning outcomes assessment in the United States
George D. Kuh and Peter T. Ewell 9

Defining and monitoring academic standards in Australian higher education
Hamish Coates 29

Women in science: the persistence of gender in Australia
Sharon Bell 47

Quality assurance in higher education as a political process
Michael L. Skolnik 67

System accreditation: an innovative approach to assure and develop the quality of study programmes in Germany
Tanja Grendel and Christoph Rosenbusch 87

The intended and unintended effects of the Bologna reforms
Sybille Reichert 99

Knowledge as a common good: the societal relevance of scientific research
Lex M. Bouter 119

Higher Education Management and Policy
Volume 22/1
© OECD 2010

The state of learning outcomes assessment in the United States

by
George D. Kuh and Peter T. Ewell
Indiana University Center for Postsecondary Research, United States,
and National Centre for Higher Education Management Systems, United States

Worldwide, economic and other factors are pressing institutions of higher education to assess student learning to insure that graduates acquire the skills and competencies demanded in the 21st century. This paper summarises the status of undergraduate student learning outcomes assessment at accredited colleges and universities in the United States. Three-quarters of institutions have established learning outcomes for all their students, a necessary first step in the assessment cycle. Most schools are using a combination of institution-level and programme-level assessments. Quality assurance requirements in the form of regional and specialised accreditation, along with an institutional commitment to improve, are the primary drivers of assessment. While there is considerable assessment activity going on, it does not appear that many institutions are using the results effectively to inform curricular modifications or otherwise to enhance teaching and learning. The paper closes with recommendations for various groups that can advance the assessment and institutional improvement agenda.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

État de l'évaluation des résultats de l'enseignement aux États-Unis

par
George D. Kuh et Peter T. Ewell
Indiana University Center for Postsecondary Research, États-Unis
et National Centre for Higher Education Management Systems, États-Unis

À travers le monde, des facteurs, notamment économiques, poussent les institutions de l'enseignement supérieur à évaluer la formation des étudiants afin de garantir que les jeunes diplômés acquièrent les compétences et les savoirs exigés au XXIe siècle. Cet article résume l'état de l'évaluation des résultats de l'enseignement des étudiants de premier cycle dans les universités et les collèges accrédités aux États-Unis. Trois-quarts des institutions ont établi des résultats sur la formation de tous leurs étudiants, ce qui constitue une première étape nécessaire en vue du cycle d'évaluation. La plupart des écoles associent des méthodes d'évaluation au niveau de l'institution et au niveau des programmes. Les exigences d'assurance de la qualité sous la forme d'accréditations régionales et spécialisées, ainsi qu'un engagement institutionnel visant à l'amélioration de l'enseignement, sont les éléments clés de l'évaluation. Tandis qu'une activité d'évaluation significative est en cours, il semble que seules de rares institutions utilisent les résultats de manière efficace en vue de la modification des cursus ou d'une autre forme d'amélioration des processus d'enseignement et d'apprentissage. Cet article se conclut par des recommandations à destination des différents groupes sur la question de l'évaluation et de l'amélioration des institutions.

Introduction

The recent world wide economic downturn threatens both affordability and quality in higher education everywhere. In the United States, rising enrolments, coupled with reductions in public and private support, make meeting these challenges even more difficult. This means, among other things, that colleges and universities must become smarter and better at assessing student learning outcomes, at using the resulting data to inform resource allocation and other decisions and at communicating to their constituents how well they are performing.

Courses, credits, certificates and degrees are important proxies for student accomplishment, but they are only proxies. It is the broad range of intended outcomes that students attain during college that yields the personal, economic, and societal benefits promised by higher education. How well are colleges and universities in the United States delivering on this promise? In this paper, we provide a partial answer to this question by reporting the results of a recent survey of academic leaders undertaken by our organisation, the National Institute for Learning Outcomes Assessment (NILOA).

Learning outcomes assessment in the United States in a worldwide context

Systematic assessment of student learning outcomes began in the United States about three decades ago and has progressed steadily in extent and sophistication. Before situating the topic in a worldwide context, we need to note that the term "assessment" in the United States generally refers to the systematic process of gathering evidence of the extent to which *groups* of students – for example, those enrolled in a particular institution or programme of study, or those sharing a similar characteristic such as gender, age, or socio-economic class – perform *in the aggregate* in attaining particular levels of knowledge or skill, in order to judge effectiveness or improve provision. This contrasts with common usage elsewhere in the world, in which the term "assessment" is frequently used to refer to the process of examining individual students in order to award degrees, marks, or grades.

As in other countries, assessment of learning outcomes in this sense of the term in the United States is undertaken for two occasionally divergent purposes. The first is accountability. As part of our national approach to quality assurance in higher education, institutional accrediting organisations

all require institutions to undertake assessment, and institutional assessment approaches are examined as part of the process of external review. This, to some extent, mirrors the growing trend toward accreditation as a quality assurance approach in much of the world, and the increasing emphasis on competency attainment in many of these processes (Santiago *et al*., 2008). As one of many examples, successive rounds of quality review in Hong Kong have examined learning outcomes more deeply in each round (University Grants Committee of Hong Kong, 2005). Similarly, the Quality Assurance Agency (QAA) in the United Kingdom recently documented a steady increase in references to student learning outcomes and outcomes assessment in its ongoing programme of academic audit (QAA, 2007). Finally, a 2004 review of the Bologna Process in Europe revealed a growing emphasis on learning outcomes references and assessment initiatives at the local, national, and international levels (Adam, 2004).

State governments in the United States have also periodically established examination programmes to directly evaluate the quality of instruction at public institutions. South Dakota, for instance, requires all students to complete a standardised examination of generic outcomes such as reading, writing, mathematics and critical thinking in order to graduate. Until recently, moreover, Florida mandated that all students achieve certain scores on a standardised examination as a condition for advancing from their second to third year of study. This parallels Brazil's use of standardised examinations from 1996-2002 to evaluate student progress (Schwartzman, 2010). Finally, government bodies throughout the world, including many states in the United States, routinely use the results of professional licensure examinations as part of their assessment of programme quality.

The second application of assessment of learning outcomes is to provide guidance in improving teaching and learning. In order to provide effective guidance, though, tools for gathering appropriate evidence need to be far more fine-tuned than those appropriate in the context of accountability, and must lend themselves to considerable disaggregation to reveal patterns of strength and weakness across different kinds of students and different dimensions of ability. This frequently demands moving beyond standardised examinations toward popular alternatives like "curriculum-embedded" assessments and student portfolios. Under the curriculum-embedded approach, faculties collectively identify specific assignments located at key points in a curricular sequence that can be used to examine particular learning outcomes at particular levels of performance. Doing this requires the institution to create its own version of a "qualifications framework"; many countries have put one in place, and it is a centrepiece of the evolving Bologna Process in Europe. It also demands "secondary reading" of student responses by faculty who did not teach the class in which the response was generated – a process

reminiscent of the external examiner approach used in the United Kingdom and other Commonwealth countries (Lewis, 2010). Under the "portfolio" process, students post examples of their work in an accessible electronic medium, grouped under learning outcomes specified by the institution or programme, as evidence that they have mastered each area. This evidence is then evaluated by teaching staff using specially developed scoring guides or rubrics. Use of portfolios as both a pedagogical and a programmatic assessment mechanism is also becoming increasingly popular in other parts of the world. For example, factors affecting the successful introduction of portfolios in the Netherlands were recently discussed in an international quality journal (Tartwijk *et al.*, 2007).

Using assessment to improve teaching and learning can be considerably enhanced if assessment results can be benchmarked against established standards or across institutions. Such benchmarking not only enables institutions and programmes to know where they stand, but also allows them to identify potential "best practices" that they can learn from. For example, a group of more than 50 independent colleges in the United States is engaged in an ongoing benchmarking collaboration, centred on a common performance assessment (CIC, 2008). Growing international interest in benchmarking the results of common learning outcomes assessments is evidenced by the Assessing Higher Education Learning Outcomes (AHELO) initiative recently launched by the Organisation for Economic Co-operation and Development in economics, engineering and generic skills (OECD, 2009).

In addition to such direct measures of student learning outcomes, the results of commonly administered surveys of currently enrolled students and recent graduates are generally included within the realm of assessment. Results drawn from such surveys, if they address specific student experiences and teaching-learning practices, can be used in combination with results of direct assessments of learning outcomes to target what needs to be improved for which student populations. Such surveys have been a part of the United States assessment landscape for more almost 50 years for purposes of benchmarking and institutional improvement. Now, similar applications are emerging in other countries. For example, a recently-published volume on the use of alumni surveys in Europe cited examples from Germany, the United Kingdom, Spain, and the Netherlands (Weerts and Vidal, 2005). Results of student and alumni surveys are also prominent in institutional accreditation reports in the United States and have been used by states to provide common accountability benchmarks for public institutions. The National Survey of Student Engagement (NSSE), for instance, is regularly administered to students in states like Kentucky, North Carolina and South Dakota (Ewell, 2010). This application is similar to the Australian Student Course Experience Questionnaire (SCEQ)

first administered in 1999 to stratified random samples of undergraduate students and repeated annually ever since (Harris and James, 2010).

In sum, although the systematic assessment of student learning outcomes began a bit earlier in the United States than its applications elsewhere, practice in all countries is rapidly converging. Because of its importance in both accountability and improvement contexts, assembling and interpreting evidence of what students know and can do as a result of their tertiary educational experience is becoming much more common and may eventually be mandatory. However, as documented in the NILOA survey described below, the biggest challenge to be faced in the United States is to use the resulting knowledge to improve teaching and learning. This is probably the case elsewhere as well.

What academic leaders in the United States say about learning outcomes assessment

The NILOA survey on United States assessment practices was conducted electronically in the second quarter of 2009. In it, we asked provosts or chief academic officers at all regionally accredited, undergraduate-degree-granting, two and four-year public, private, and for-profit institutions in the United States (n = 2 809) about the assessment activities underway at their institutions and how assessment results are being used. The NILOA questionnaire (NILOA, 2009) was organised around four broad questions:

1. What learning outcomes are you measuring at your institution?
2. How are you assessing these outcomes and using the results?
3. What are the major factors prompting assessment at your institution?
4. What do you need to further learning outcomes assessment at your institution?

All told, of the 1 518 institutions contacted, 53% responded. The characteristics of these participating institutions reflect the national profile in their institutional sectors, size (based on enrolments) and geographic regions.

Major findings

Eight observations summarise the current state of outcomes assessment in the United States and suggest that more assessment activity may be underway in American higher education than some government officials and others might assume.

1. Most institutions have identified a common set of learning outcomes that apply to all students

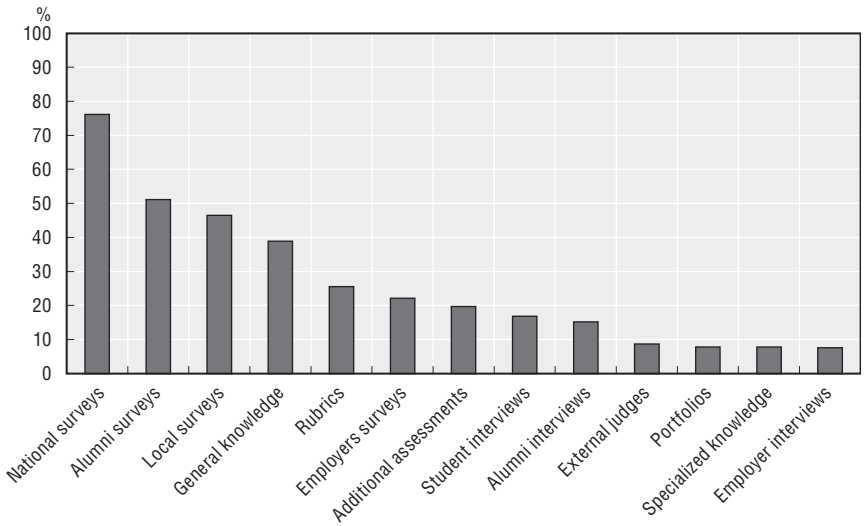
About three-quarters of all institutions reported having common learning outcomes for all their undergraduate students. This is consistent with another recent study of Association of American Colleges and Universities member schools (Hart Research Associates, 2009). Unsurprisingly, given their complexity and wide range of programmes, larger research-intensive institutions were less likely to have common learning outcomes for all undergraduate students than were colleges that award primarily baccalaureate or associate degrees.

2. Most institutions use a combination of institution-level and programme-level assessment approaches

We asked provosts what approaches were used by their institutions to assess learning outcomes, such as nationally normed measures of general knowledge and skills, portfolios, national or locally developed surveys, and alumni and employer surveys and interviews. We also asked if the tools or approaches were used with institutionally valid samples so that claims could be made about overall institutional performance or if the assessment approach focused at the programme level. Assessment tools and approaches understandably vary depending on what the data are intended to represent.

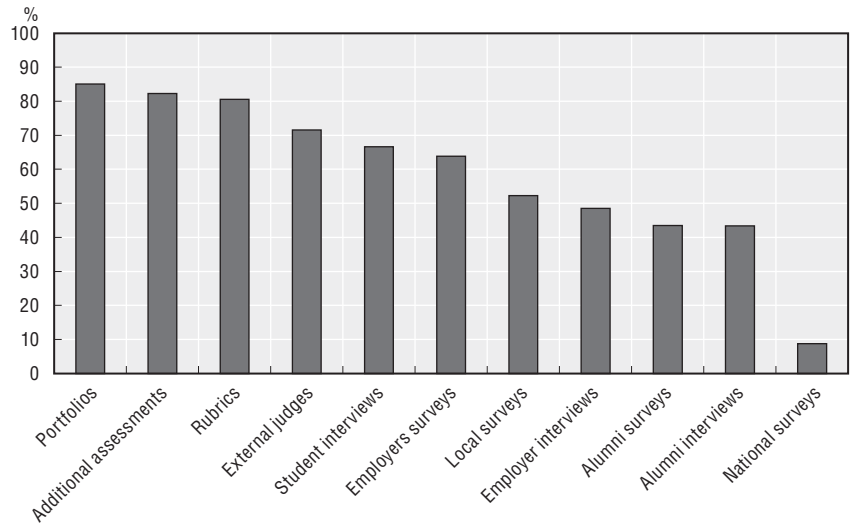
- The vast majority (92%) of all colleges and universities use at least one assessment approach or tool with institutionally valid samples; two-thirds of all schools use three or more.
- Nine out of ten schools use at least one institutional-level and one programme-level assessment approach; 77% use two or more of each type and 58% use three or more of each.
- The most frequent approach used with an institutionally valid sample was a national survey. More than three-quarters (76%) of all schools reported using surveys at the institution-wide level.
- Two-fifths (39%) of all campuses reported using a standardised measure of general knowledge and skills.
- External expert judgments of student work (9%), tests of specialised knowledge (8%), student portfolios (8%) and employer interviews (8%) were used much less often with institutionally valid samples (Figure 1).
- At the programme level the most popular approaches to assessing learning outcomes were student portfolios, measures of specialised knowledge and other performance assessments, and rubrics (Figure 2), as more than 80% of institutions indicated at least one of their academic programmes was using one of these approaches.

Figure 1. **Institution-level assessments of learning outcomes for all institutions**



Source: Kuh, G.D. and S.O. Ikenberry (2009), *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education*, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment.

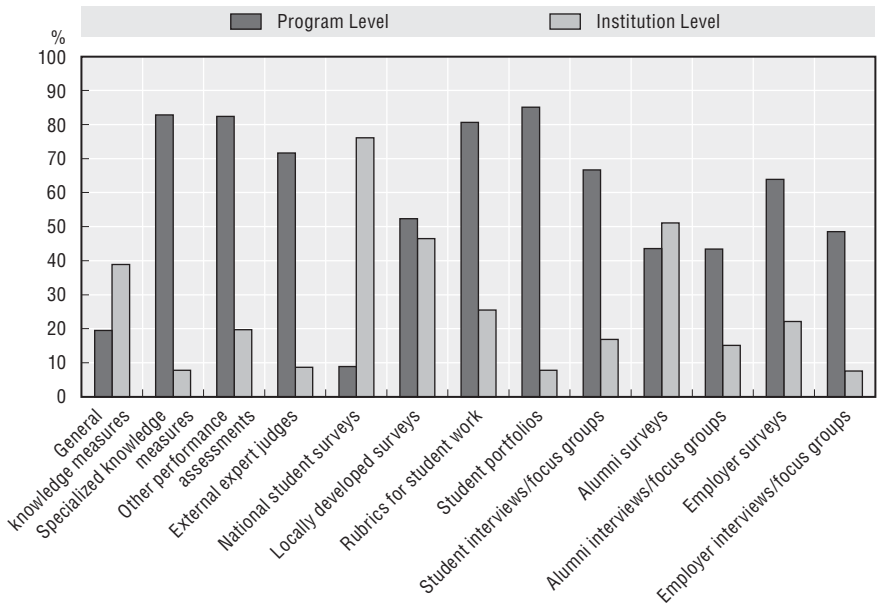
Figure 2. **Programme-level assessments of learning outcomes for all institutions**



Source: Kuh, G.D. and S.O. Ikenberry (2009), *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education*, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment.

- Community colleges and other associate degree-granting institutions were more likely to use general knowledge assessments at the programme level.
- More for-profit schools employed multiple approaches using institutionally valid samples, but fewer collected information at the programme level, reflecting the more discipline-focused nature of their curricula (Figure 3).
- At more than seven out of ten institutions, at least one department was using:
  - ❖ specialised knowledge measures;
  - ❖ performance assessments other than grades;
  - ❖ external judgments of student performance;
  - ❖ rubrics;
  - ❖ portfolios;
  - ❖ student interviews;
  - ❖ surveys of employers.

Figure 3. **Types of assessment information**

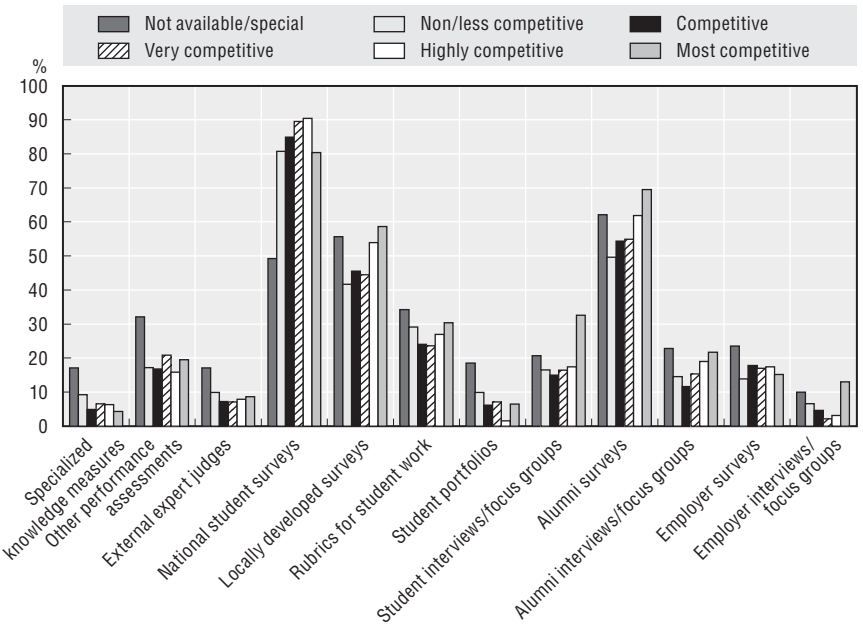


Source: Kuh, G.D. and S.O. Ikenberry (2009), *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education*, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment.

3. Assessment approaches and uses of assessment results vary systematically by institutional selectivity

In general, institutions that have less competitive admissions standards were more likely to administer standardised measures of general knowledge with institutionally valid samples. Colleges and universities with the most competitive admissions standards more frequently used locally developed instruments to collect information from students and alumni (Figure 4).

Figure 4. **Institutional-level assessment by selectivity**



Source: Kuh, G.D. and S.O. Ikenberry (2009), *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education*, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment.

- About half of the least competitive schools employ general knowledge tests compared with only about one-fifth of the most competitive institutions.
- At least four-fifths of all schools use nationally normed student surveys, except for institutions that do not have selectivity data available, where only half do so.

The **uses** of assessment data at institutions of varying selectivity tell a different story. While the most competitive colleges and universities collect information at rates generally comparable to their less-selective counterparts, they do not report using it nearly as often – with one exception: reporting to

the governing board. To illustrate, the most competitive institutions are least likely to use assessment data for:

- revising learning goals;
- responding to calls for accountability;
- informing strategic planning;
- improving instructional performance;
- evaluating units and programmes;
- allocating resources;
- reporting to the public.

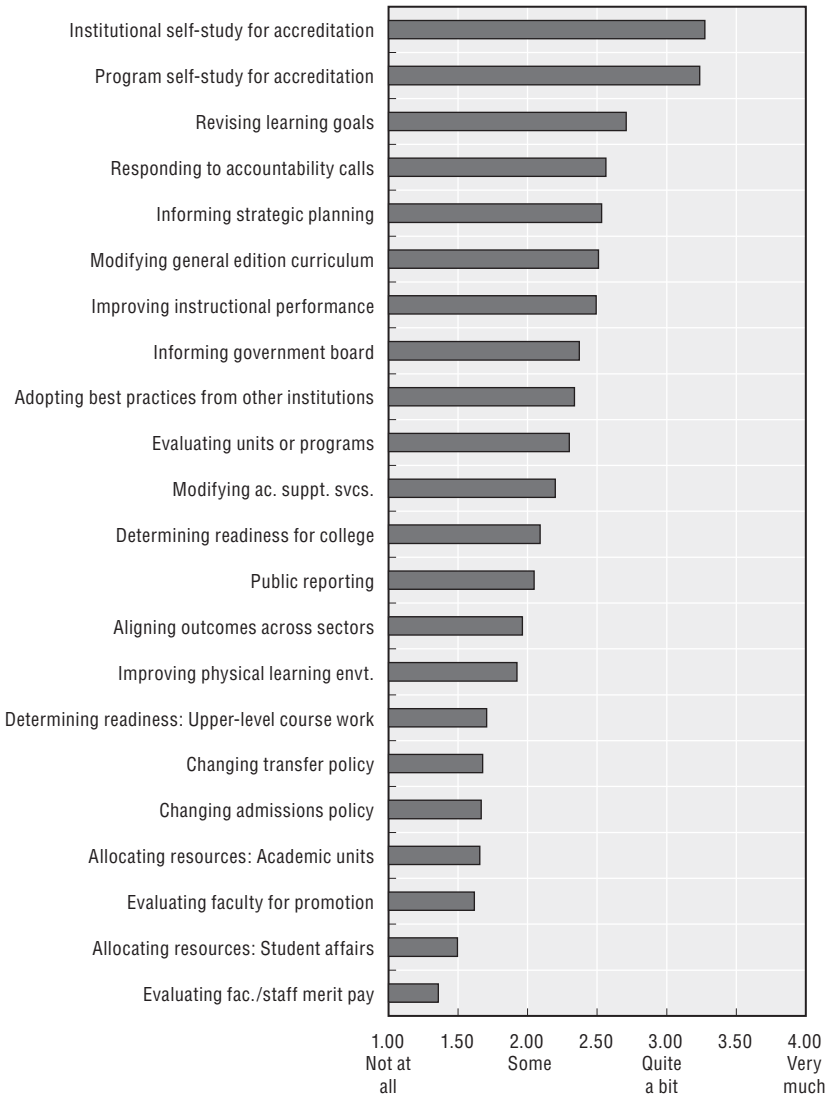
4. The most common use of assessment data is related to accreditation

The most common uses for student learning data reported by provosts were for preparing for institution and programme accreditation and, to a lesser degree, for revising undergraduate learning goals. Assessment results were used to a lesser extent for making day-to-day decisions about resources, admissions or transfer policies, faculty and staff performance, etc. (Figure 5).

The patterns of assessment data use varied somewhat by institution type:

- Fewer doctoral-granting universities were using outcomes data for determining student readiness for upper-level course work, improving instructional performance, evaluating departments, allocating resources to academic departments, and informing strategic planning.
- At the same time, more doctoral institutions were using results to respond to calls for accountability such as the Voluntary System of Accountability (VSA) and to meet specialised academic programme accreditation requirements.
- Baccalaureate schools were more likely to incorporate assessment results for making faculty promotion and tenure decisions, consistent with their focus on undergraduate education.
- Community colleges and other associate degree-granting institutions reported using outcomes data for aligning curricula across sectors, determining student readiness for college course work, improving instructional performance, and allocating resources to academic units – all encouraging findings.
- For-profit schools reported the most frequent use of assessment data in every category of use. While only 34 for-profit schools are represented in these data, they represent more than half of the accredited for-profit institutions that award degrees, which were the two criteria for inclusion in the sample. So, the results for this group of institutions probably are as reliable as for the schools in other categories.

Figure 5. **Uses of assessment data for all schools**



Source: Kuh, G.D. and S.O. Ikenberry (2009), *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education*, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment.

5. *Assessment is driven more by accreditation and a commitment to improve than external pressures from government or employers*

The three most influential forces driving outcomes assessment reported by provosts are the requirements of regional accreditors, the requirements of

specialised accreditors, and an institutional commitment to improvement. Somewhat less influential in this regard were national calls for accountability or mandates from trustees or state co-ordinating boards.

The relative importance of different factors prompting outcomes assessment varied somewhat in predictable ways by institution type:

- Community colleges and other associate degree-granting institutions were more responsive to co-ordinating and governing board mandates. In the United States, these are all public institutions, with oversight and funding provided by state governments.
- Baccalaureate institutions accorded relatively greater importance to a campus commitment to improvement as a reason for assessing learning outcomes. Master's institutions gave regional and specialised accreditation relatively greater weight.
- Initiatives such as the "Voluntary System of Accountability" recently put in place as an accountability vehicle by national associations of public universities seemed to be more influential at doctoral-degree-granting institutions relatively less influential at those campuses was faculty and staff interest in improving student learning.
- For-profit schools indicated that every one of the eight observations was influential in driving assessment activity, again suggesting a sharper focus on learning outcomes assessment at those schools.

6. Learning outcomes assessment is undercapitalised

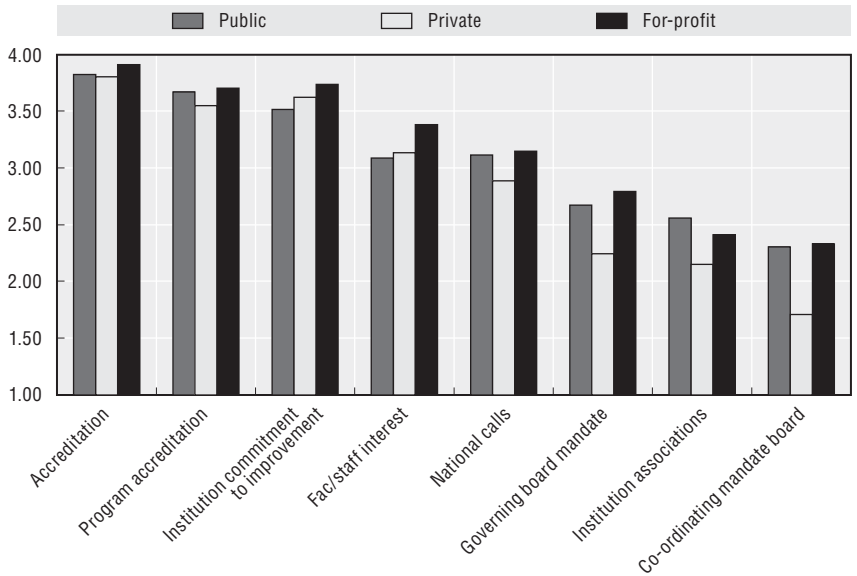
Given the importance of higher education to the future of society and the resources devoted to the enterprise, investment in assessment staff is modest.

- Although four-fifths of all institutions had a person or unit charged with co-ordinating or implementing assessment campus wide, only a quarter assigned more than one full-time equivalent (FTE) person to assessment.
- Almost half (47%) of doctoral institutions reported having one or more staff, while only one-fifth (19%) of community colleges and other associate degree-granting schools had at least one person focused on outcomes assessment.

7. The two greatest needs to advance student learning outcomes assessment are greater involvement of faculty and more assessment expertise, resources and tools

Two-thirds of all schools noted that more faculty involvement would be helpful. Three-fifths wanted more assessment expertise. Rated least important in terms of need was information about assessment policies and practices at other schools (18%) and presidential support (9%).

Figure 6. **Assessment drivers by control type**



Source: Kuh, G.D. and S.O. Ikenberry (2009), *More than you Think, Less than we Need: Learning Outcomes Assessment in American Higher Education*, Urbana, IL: University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment.

8. Most institutions plan to continue outcomes assessment work despite budgetary challenges

Although more than half of all institutions predicted that the recession would not affect their assessment activities, a not insignificant number (one-fifth) indicated that a decrease in institutional support was possible. Understandably, about 15% of all schools (and more at public institutions) were not certain about what might happen at the time the survey was conducted.

Outcomes assessment in the United States is work in progress

Our survey revealed that a fair amount of assessment work is going on in colleges and universities in the United States. Challenges to additional progress remain, however. Student performance evaluation is so embedded in the everyday work of teaching, testing and grading that many faculty members interpret calls for documenting outcomes at the programme or institution level – if not as an outright threat – as a redundant exercise or worse: a waste of time and resources that could be more profitably invested elsewhere. Thus, it was not surprising that gaining faculty co-operation and engagement was at the top of provosts' wish lists.

Campus culture also plays a role. As noted earlier, the most selective institutions are the least likely to use assessment data for improvement or accountability. Some faculty and staff at prestigious, highly selective campuses wonder why documenting something already understood to be superior is warranted. They have little to gain and perhaps a lot to lose. On the other hand, many colleagues at lower-status campuses often feel pressed to demonstrate their worth; some worry that they may not fare well in comparison with their better resourced, more selective counterparts. Here, too, anxiety may morph into a perceived threat if the results disappoint.

Accreditation is a catalyst for improvement and accountability

Accreditation is the primary vehicle for quality assurance in American higher education and the major driver of learning outcomes assessment. A fair amount of assessment work is with institutionally valid samples, especially using student and alumni surveys as well as standardised measures of general knowledge and skills. Equally important, various assessment approaches are being used at the program level – in engineering, business and teacher education, for example. Such work often animates improvement. The curricular changes in engineering and engineering technology education stimulated by ABET, Inc., are especially instructive, because much of the impetus originated outside the academy by practitioners via the accreditors. Second, the changes featured discipline-specific assessment strategies to evaluate the efficacy of the changes in a formative and summative manner.

That same convergence of improvement and accountability forces is influencing institution-wide regional accreditation. While the focus of regional accreditation is improvement, external accountability forces are shaping and sharpening its expectations to press for more extensive assessment of student learning and using the results for improvement and making institutional performance more transparent.

While some observers see these two purposes – **improvement and accountability** – if not at odds, at least in tension with each other (Ewell, 2009), campuses seem to suggest that their assessment efforts are substantively influenced by both factors.

Sustaining assessment work

Allocating resources to assessment is an expression of institutional priorities, culture and values. Some institutions have more resources than others to devote to student learning outcomes assessment; colleges and universities that offer a substantial variety of programmes should spend more on it. While in the past campuses were left to determine the quality of effort they would direct to assessing student learning, the time has come for a

systematic analysis of what institutions of varying levels of organisational and programmatic complexity should invest to get assessment right and to ensure effective use of the results.

The degree to which an institution or programme is likely to expend resources on improving student learning is a function of its knowledge about how well its students are learning what is important and its knowledge of what to do to improve learning outcomes. How well are individual courses coming together as a cohesive whole? Are the essential learning goals and expectations for students being met? Do engineering graduates have the crucial knowledge and skills? Is the nurse sufficiently well prepared to care for the patient? Is the newly minted graduate capable of critical thinking? Does (s)he have the analytical and communication skills the campus promises and employers expect?

Focusing on these and related questions about outcomes can be the common ground that brings together those who demand greater accountability by documenting accomplishment and those whose primary interest in assessment is enhancing accomplishment. States and higher education associations can play an important role in bridging this divide.

Recommendations and potential actions

The looming challenge for higher education in the United States is to convince naysayers among the faculty that assessment is not a threat and to find ways to thoughtfully **use** assessment data to inform decisions, improve programmes, and communicate effectively with the public. Indeed, it is this last point – productively **using** learning outcomes results to inform decision making and to improve teaching and learning – that remains the most important unaddressed challenge related to student learning outcomes assessment in our country. Simply posting a number on an institution's website or checking a box on an accreditation report are of little value to students, parents, or policy makers. Equally important, such actions do nothing to improve access, affordability, or accomplishment.

Who needs to do what to advance the assessment of student learning outcomes in ways that improve the current state and future prospects of higher education in the United States?

- **Presidents, provosts, and other academic leaders** must make quality assurance an institutional priority. They need to tell their assessment professionals what the institution needs to know and why, and assign a group to evaluate and periodically report on the quality and utility of the learning outcomes assessment efforts underway. They should determine whether the resources allocated to assessment are sufficient for the scope and growing importance of the task, and know how the results are being used, if at all, by

whom and for what purposes. They must champion effective use of the results to make decisions. Finally, they need to keep the governing board informed about the degree to which a culture of evidence is taking root.

- **Governing board members** must ensure their institution has a system of academic quality control supported by the assessment of student learning and the use of those results for continuous improvement. They should receive an annual update on these efforts. Institutional leaders need to ensure that their board chairs keep the issue on the agenda for at least one meeting each year.
- **Faculty members** must systematically collect data about student learning, carefully examine and discuss these results with colleagues, and use this information to improve student outcomes. This challenging process may well reveal shortcomings on the part of students, instructors, the curriculum, and institutions. But by making sure these data are used to improve and not penalise, the exercise need not, and should not, be threatening. If assessment results are to be meaningfully interpreted and if changes are to be made to improve outcomes, faculty leadership and involvement are crucial.
- **Assessment and institutional research personnel** should revisit the rationale for using various tools and approaches to be sure they yield the kind of information that the institution needs to respond to improvement and accountability mandates. They should present results in ways that will speak to faculty and policy makers and will answer their important questions. They need to point to areas that assessment data indicate require attention and design subsequent data collection activities that will determine whether changes in teaching and learning approaches have had the desired effects.
- **Student affairs staff** must share their perspectives on the student experience by participating on the campus assessment committee and self-study committees. They should partner with academic affairs to promote a deeper, more widespread awareness and understanding of common undergraduate learning outcomes among faculty, staff and students. They need to use outcomes assessment results to orient and inform student affairs practice.
- **Faculty developers** must become familiar with the campus assessment activities and results and use this information to design professional development opportunities for faculty, student affairs professionals, librarians and others who work with students.
- **Prospective students and parents** should look for learning outcomes information about students who attend the institutions they are considering. If it is not publicly accessible on an institution's website, they

should ask someone in the institutions' admissions offices for data about how their students perform on different kinds of measures.

- **Higher education associations** must keep learning outcomes assessment on their agenda. Much of the campus assessment activity reported by provosts would not be underway without the initiatives of these organisations. *Statewide planning and co-ordinating boards* must confirm that all institutions under their scope of influence have effective internal systems of academic quality control supported by assessment data that conform to the expectations of both regional and specialised accreditation bodies. They should use language that removes the spectre of threat from assessment work, and offer incentives for campuses to develop and share sound practices of outcomes assessment.
- **Accrediting groups** must not slacken their efforts to promote assessment and the use of student learning outcomes. They need to sharpen accreditation standards as they are applied to: *i)* collecting institution- and programme-level data about student performance, *ii)* using assessment results to improve student performance and institutional quality; and *iii)* making assessment results available internally and externally. In all of these areas, they must hold institutions accountable.
- **Foundations** should keep learning outcomes assessment on their funding agendas. They should devote more attention to programmes and incentives that encourage institutions to use outcomes data productively. Accrediting groups, both regional and specialised, should be encouraged to be vehicles for campus change that is constructive and attainable.

These suggested action steps are necessary but not sufficient to strengthen American higher education through more effective knowledge of student learning outcomes and to use that knowledge to improve. While more assessment work is underway than many think in the United States, it is considerably less than what is needed to ensure US college graduates are prepared to manage the challenges of the 21st century and to secure the future to which we aspire.

The authors:

George D. Kuh
Chancellor's Professor and Director
Center for Postsecondary Research, Indiana University
1900 East 10th Street, Eigenmann Hall #419
Bloomington IN 47406-7512
United States
E-mail: kuh@indiana.edu

Peter T. Ewell
Vice President
National Center for Higher Education Management Systems
3035 Center Green Drive
Suite 150
Boulder, CO 80301-2251
United States
E-mail: peter@nchems.org

The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the organisations that funded this research: Carnegie Corporation of New York, Lumina Foundation for Education and The Teagle Foundation.

References

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", Scottish Executive, Edinburgh, Scotland.

Council of Independent Colleges (CIC) (2008), *Evidence of Learning: Applying the Collegiate Learning Assessment to Improve Teaching and Learning in the Liberal Arts Experience*, Council of Independent Colleges, Washington DC.

Ewell, P.T. (2009), *Assessment, Accountability, and Improvement: Revisiting the Tension*, NILOA Occasional Paper No. 1, University of Illinois and Indiana University, National Institute of Learning Outcomes Assessment, Urbana, IL.

Ewell, P. (forthcoming), "The US National Survey of Student Engagement (NSSE)", D. Dill and M. Beerkens (eds.), *Public Policy for Academic Quality: Analyses of Innovative Policy Instruments*, Springer, New York, pp. 87-102.

Harris, K-L. and R. James (2010), "The Course Experience Questionnaire, Graduate Destination Survey, and Learning and Teaching Performance Fund in Australian Higher Education", D. Dill and M. Beerkens (eds.), *Public Policy for Academic Quality: Analyses of Innovative Policy Instruments*, Springer, New York, pp. 103-124.

Hart Research Associates (2009), *Learning and Assessment: Trends in Undergraduate Education* (a survey among members of the Association of American Colleges and Universities), Washington DC, *www.aacu.org/membership/documents/2009MemberSurvey_Part1.pdf*.

Lewis, R. (forthcoming), "External Examiner System in the United Kingdom", D. Dill and M. Beerkens (eds.), *Public Policy for Academic Quality: Analyses of Innovative Policy Instruments*, Springer, New York, pp. 21-38.

National Institute for Learning Outcomes Assessment (NILOA) (2009), National Survey of Learning Outcomes Assessment, *www.learningoutcomeassessment.org/documents/NILOA_ForWeb_5-22.pdf*.

Organisation for Economic Co-operation and Development (2009), *OECD Feasibility Study for the International Assessment of Higher Education Learning Outcomes*, OECD publishing, *www.oecd.org/document/22/0,3343,en_2649_35961291_40624662_1_1_1_1,00.html*.

Quality Assurance Agency (QAA) (2007), *Outcomes from Institutional Audit, The Adoption and Use of Learning Outcomes*, Gloucester, United Kingdom.

Santiago, P. *et al*. (2008), *Tertiary Education for the Knowledge Society*, Vol. 1, OECD Publishing.

Schwartzmann, S. (forthcoming), "The National Assessment of Courses in Brazil", *in* D. Dill and M. Beerkens (eds.), *Public Policy for Academic Quality: Analyses of Innovative Policy Instruments*, Springer, New York, pp. 309-332.

Tartwijk (van), *J. et al*. (2007), "Factors Influencing the Successful Introduction of Portfolios", *Quality in Higher Education*, Vol. 13, 1 April 2007, pp. 69-80.

University Grants Committee of Hong Kong (2005), *Education Quality Work: The Hong Kong Experience*, University Grants Committee of Hong Kong, Hong Kong.

Weerts, D. and J. Vidal (2005), "Enhancing Alumni Research: European and American Perspectives", *New Directions for Institutional Research*, No. 126, Jossey-Bass, San Francisco.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Defining and monitoring academic standards in Australian higher education

by
Hamish Coates
Australian Council for Educational Research (ACER), Australia

This paper outlines the need for adopting a more scientific approach to specifying and assessing academic standards in higher education. Drawing together insights from large-scale studies in Australia, it advances a definition of academic standards, explores potential indicators of academic quality and looks at approaches for setting standards. As learner outcomes need to be placed at the forefront of work on academic standards, this paper concludes by exploring the implications of this position for student assessment and institutional change.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Définition et suivi des standards académiques de l'enseignement supérieur en Australie

par
Hamish Coates
Australian Council for Educational Research, Australie

Le présent essai souligne la nécessité de l'adoption d'une approche plus scientifique de la spécification et de l'évaluation des standards académiques de l'enseignement supérieur. Partant des aperçus tirés d'études à grande échelle réalisées en Australie, il articule une définition de standards académiques, explore des indicateurs potentiels de qualité académique et examine des approches en vue d'établir des standards. À l'heure où les résultats des étudiants doivent constituer la priorité du travail sur la question des standards académiques, le présent essai aboutit à l'examen des implications de ce positionnement pour l'évaluation des étudiants et pour l'évolution des institutions.

Counting standards in higher education

Universities are responsible for maintaining academic standards. Indeed, having autonomy over standards is an intrinsic factor which contributes to their status. This privilege partly derives from the capacity of universities to conduct scholarly research and generate knowledge through teaching – to explore the foundations and contours of knowledge itself. Hence it could come as a surprise that while universities have responsibility for maintaining standards, relatively little scholarly research has been done on what this actually means, particularly in relation to student learning and outcomes. To impel progress in this area, this paper outlines a case for giving greater attention to the specification and assessment of academic standards in higher education. It argues, in particular, that significantly more emphasis needs to be placed on assessing learning outcomes.

Today, it is possible for students at different or even the same universities who have the same underlying competence to perform at comparable levels on the same assessment task, yet receive very different grades. This unfortunate state of affairs may arise from bias or lack of precision in marking processes. Alternatively, it may reflect variability in the tasks or in grading or reporting systems. Even when common tasks, marking processes and reporting metrics are used, these can still mask variability that goes unchecked by the absence of cross-validation. Such practice, where it occurs, is unsatisfactory. Assessment results are high stakes for students, play a significant and formative role in shaping individuals' outcomes, and impact graduates' broader social and professional experiences.

Given the various forms of validation that are used to ensure the quality of academic research, it is surprising that universities can lack procedures for verifying the standards of education. Of course, many excellent accountability mechanisms have been implemented widely by departments, institutions and systems. Continuous improvement, benchmarking, monitoring, risk assessment and auditing all refer to mechanisms for defining, measuring, assuring and enhancing academic standards. Yet while regulatory activity has grown along with the importance of higher education to the knowledge economy, there remains little formal conceptualisation of what is meant by "academic standards" and how these can be assessed. This is disquieting, for there is an increasing realisation (see, for instance, Salmi, 2009) that collegial approaches are no longer singularly sufficient to support the accountability

requirements of contemporary higher education. The drive to define and enhance standards is not new, but expansion of higher education institutions and systems creates new and increased pressure for change.

Close to the heart of much contemporary discourse in this area is a growing interest in developing new ways of documenting what students learn, know and can do. Of course, universities have always had students' learning and development as a core part of their mission. Discussions about higher education quality, however, have tended to focus on institution-level inputs or, in the last few decades, on teaching processes. Recent interest, however, flags a shift towards focusing more explicitly on students – on what they are doing and achieving. Developing, monitoring and reporting achievement and broader graduate outcomes is increasingly seen as an institutional and even system-wide responsibility.

This interest is international in scope. In Brazil, for instance, the pioneering national courses examination, the *Provão,* has been administered since 1996, and now provides external data across more than a dozen disciplines. In the United States, the Collegiate Learning Assessment (CAE, 2009) has been used by over 400 institutions for around a decade to collect objective data on learning outcomes. *Measuring Up* (NCPPHE, 2008) and related reports (see, for instance: Ewell, 2009; Ikenberry and Kuh, 2009; Spellings, 2006) have for some years flagged the absence of information on student learning metrics, and the Voluntary System of Accountability (VSA, 2009) has been developed to drive improvement in this area. In the United Kingdom, the Quality Assurance Agency (QAA) announced early in 2009 that it would look at how to build on the work of external examiners to make more explicit and comparable statements about achievement standards. Cross-nationally, the OECD's International Assessment of Higher Education Learning Outcomes (AHELO) feasibility study (OECD, 2009a) is examining the feasibility of assessing and comparing graduate learning across nations, systems, languages and cultures.

In line with these ideas, this paper brings together insights from large-scale research activities in Australia (AUQA, 2009; Coates, 2007a, 2007b, 2007c, 2007d, 2008, 2009a) to explore the foundations of a generalisable approach to specifying and assessing academic standards in higher education. The following section offers a working definition of academic standards, and is followed by a discussion of the methods used for selecting indicators of quality and establishing performance thresholds. The final sections affirm the importance of focusing on learner outcomes, and look at the implications of this for student assessment and institutional change.

Many of the ideas in this paper are provocative and would not be quick or easy to implement. It is contended, however, that they would result in a

significant improvement to academic work. With a more contextualised sense of how they have performed, for instance, students will have a better idea of how their work will be measured and reported, and a clearer picture of what their final reported performance signifies. Employers will have greater confidence in the meaning of the results that graduates present to them, regardless of the institution from which they come. Academic staff will be able to face and rebut with greater confidence the frequent accusations of falling standards and "dumbing down".

What "academic standards" means

The concept of academic standards lies at the heart of higher education quality. Perhaps because of this it has proved difficult and even elusive to define. James (2003, p. 189) has suggested that this may be due to the "highly abstract nature of standards… [or because the] higher education sector has had little sustained discussion and analysis of standards in recent years". Whatever the reason, it is useful to offer a brief deconstruction and working definition of the concept as a foundation for this paper and subsequent practice.

In education, as in many other industries, the term "standards" is used both in substantive and descriptive ways. Standards can refer to shared measures against which comparisons can be made. This is the "what" of standards. Examples include teaching quality, learning outcomes or information resources. In line with the OECD (2009b) and other common use, such phenomena are referred to as "indicators" in this paper.

"Standards" is also used to refer to varying levels of quality or performance, as in phrases such as low performance, high quality or teaching excellence. This is the "how much" of standards. This use is reinforced by technical methods for setting standards (Angoff, 1971). Similarly, the Australian Department of Education, Science and Training (DEST, 2002) states that "Academic standards usually refer to student performance and levels of achievement on a particular piece of assessment, in a subject, or at the end of a degree". The Australian Universities Quality Agency (AUQA, 2007) defines standards as "A specification or other precise criteria designed to be used consistently as a rule, guideline or definition of a level of performance or achievement". Anderson, Johnson and Milligan (2000) write that "It is assumed that standards can be represented on a scale, at least to the extent of distinguishing pass and fail and various classes of honours attainment". Following such use, and consonant with the definition proposed by AUQA, the term "standards" is used in this paper to denote agreed levels of quality or performance on an indicator of student achievement.

As with standards, the term "academic" is used in many different ways. In broad terms, academic might be used to qualify a wide range of activities

that pertain to the production, accumulation and enhancement of knowledge. This work is conventionally distinguished into the three overlapping dimensions of education, research and service. This paper focuses on education and, as the subsequent discussion brings out, places particular emphasis on student learning and outcomes.

It is proposed, therefore, that a useful working definition of academic standards could be "agreed levels of academic quality" or, more fully, "agreed levels of performance on indicators of academic quality". This implies the need to define indicators of academic quality.

Academic quality indicators

Developing indicators to measure complex phenomena is a complex task. Valued phenomena are not necessarily easy to define, measure and report. Even simple indicators must be accompanied by considerable qualifications and caveats. Indicators can carry different meanings in different contexts, be underpinned by different or differently collected data, or carry different implications for policy and practice. Along with the many educational and practical considerations, the definition of outcomes is a necessarily value-laden process and, as such, it can be difficult to reach consensus about which domains are relevant for a given situation.

Luckily, a considerable amount of work has been undertaken over the last 50 years to develop processes for the specification of higher education quality indicators (Bauer, 1966; Cuenin, 1988; Kells, 1993; Davis, 1996; Cave *et al*., 1997; Linke, 1991; National Commission on Excellence in Education, 1983; Coates, 2009b). Any selection of indicators needs to balance technical, practical and substantive considerations. In summary, it is vital that indicators are valid, relevant to key phenomena, stable across contexts, transparent, non-trivial, responsive to change, auditable, efficient to collect, preferably readily available, as simple as possible, quantifiable and generalisable. Of course, higher education is a deeply contextualised activity, and indicators for universities should link with those used by schools, industry, the professions, vocational education and training and the broader community.

While data collections proliferated in the 1990s in many countries, in step with the global expansion of higher education and the quality movement, much of this proceeded without reference to an underlying set of indicators. In Australia, for instance, the national indicator system currently in use was developed in the late 1980s (Linke, 1991) for use in a very different kind of higher education system. Of course, institutions and institutional networks developed indicators to drive their own continuous improvement, but on the whole they are not based on the kind of research and consultation that facilitates cross-institutional generalisation.

Indicators are most powerful when they are located within a multidimensional framework. The educational framework that consists of input-process-output stages is one of the most general. This framework usually takes a hierarchical form, and further distinguishes each of the stages into individual, instructional and institutional levels. This basic but robust framework has been used across a large number of diverse education systems and contexts (Bottani and Tuijnman, 1994; Astin, 1985; Ewell and Jones, 1996; Jaeger, 1978). Its most general expression is found in the OECD's Indicators of Education Systems (INES) programme (OECD, 2009b).

In 2007, research was conducted with five Australian universities to develop an academic standards model (Coates, 2007c). An iterative process of review and consultation was undertaken and led to indicators given in Table 1.

Table 1. **Indicators of education quality**

	Outcomes	Processes	Inputs
Higher education learners	• Graduation rates • Graduate destinations • Learning outcomes • Graduate capabilities • Work readiness • Satisfaction	• Student engagement • Retention and progress	• Entry levels • Entry pathways • Student diversity • Student characteristics • Student aspirations
Higher education teachers	• Teaching experience • Teaching resources	• Teaching processes • Course management • Support systems	• Staff characteristics • University enculturation • Educational resources • Curriculum
Higher education institutions	• Institutional growth • Institutional reputation • Community engagement	• Academic governance • Academic management • Academic culture • Staff development • Quality systems	• Institutional characteristics • Institutional resources • Industry engagement

Source: Coates (2007c).

This framework is not advanced as a definitive articulation of academic quality indicators, but as one means of illustrating the broad nature of phenomena that need to be considered to build a rich picture of the complex business of university education.

Indicators of academic standards should reflect high-level concepts. However, it is necessary to underpin indicators with tangible data elements, hereafter referred to as "measures". These measures provide evidence of what is being achieved, and thereby play a very important role in the assessment of academic standards. It is important that the measures are carefully

considered, for without careful consideration of data requirements, there is a risk that data availability will dictate overarching quality objectives.

Each indicator might be underpinned by one or more measures. Data elements can be differentiated as follows: those based on fact, on subjective feedback, or on objective assessment. The distinguishing characteristic of the measures is that they are, or could be, operationalised in terms of specific data. Certain measures may be international or national in scope, while others may be relative to specific disciplines or institutions. It is essential that measures are formed in technically and educationally appropriate ways.

Inevitably, given the scale and complexity of higher education, an intricate information architecture is required to map measures onto indicators. It is useful to consider a few examples to illustrate the relationship between indicators and potential measures:

- Graduation rates may be assessed by calculating the percentage of students beginning a programme who graduate within a specified period of time.
- Student engagement is most commonly measured using surveys, or sometimes via time diaries or data collected through administrative systems.
- Teaching experience could be measured through reference to factors such as qualifications, contact hours, professional recognition, or excellence awards.
- University enculturation may be operationalised by measuring the effort institutions make to include teachers in the institution, particularly teachers who are new to the profession or have a contingent employment status.
- Community engagement may be defined variously in terms of alumni programme size, employer satisfaction feedback, participation in service learning, or sponsorship of public events.

Setting performance standards

The enumeration of operational measures facilitates the setting of academic standards. A range of methods are used in education to define quality standards. Standards can be set through reference to the performance of other individuals or groups, an approach which is often used in assessing and reporting student learning. Standards can also be set with reference to specific distributions, often specified in terms of percentiles, as is common in institutional benchmarking processes (McKinnon, Walker and Davis, 2000) or the calibration of student mark distributions. The Angoff standards-setting procedure (Angoff, 1971) has been developed as a robust method for defining the criteria which distinguish varying levels of performance. Standards can be set through empirical or normative analysis, through accreditation or moderation, or through trend analysis. A recent review of extant processes (AUQA, 2009) suggests that application of the word “appropriate” may currently

be among the most common means of setting standards, however this approach fails to provide an objective performance expectation.

A hybrid approach is commonly used that suits the indicators and measures under study and the contexts of development. Standards for high-stakes student assessments, for instance, might well be best set using psychometric modelling along with review by discipline experts. Standards for course curriculum and relevance might be set using rating criteria and expert moderation, either during course development, or on an ongoing and iterative basis. Professional accreditation processes play an important standards-setting function in many disciplines.

From a technical perspective, the standards-setting process should result in the specification of a series of thresholds that map out a continuum of increasing performance. Such thresholds operationalise the indicator in terms of its underpinning measures. The measures therefore play a role in defining the indicator in terms of current and potential planning and practice.

By far the most important standards from a public policy perspective are those that provide assurance that universities are exceeding minimum levels of performance. From a continuous improvement perspective, however, there would also be value in moving beyond minimal standards to set thresholds that measure gradations of increasing performance and potential. Specifying different levels of performance may assist institutions measure improvement against targets, and recognise particularly accomplished forms of performance excellence.

As with the indicators of academic quality, the standards would often be national or international in scope, although a few may be specific to individual institutions, courses or even teachers. Such standards provide a basis for review, audit, monitoring and enhancement processes, and they would derive much of their value from having external points of reference. It is important to stress that setting *a priori* standards does not imply or reinforce any move towards standardisation of institutional practice. Rather, as in the quality control of scholarly research, it provides scientific foundations that provide a springboard for diversity and excellence.

Developing approaches focused on learner outcomes

It is argued in this paper that learner outcomes are the most important of all the quality indicators categorised in Table 1. These, in many respects, can be considered an "educational bottom line". Of course, many variables influence student performance, including student background, initial knowledge and skill, programme design, student effort, teaching resources and teaching quality. Gathering data about these types of input and process variables and evaluating them are very important exercises, particularly for

each institution's own continuous improvement. Limiting the scope of quality assessment and review strictly to these variables, however, cannot substitute for a direct focus on achievement itself. This is primarily because the various inputs and processes interact in complex ways, and are not deterministic. In large-scale analyses, for instance, it is common for context and demographic variables to explain a relatively small portion of student outcomes (Coates, 2008; Coates and Edwards, 2009).

An explicit focus on learner outcomes examines the net learning effect of all the variables operating together. It allows the attained level of achievement to be assessed and recorded (as grades on student transcripts, for instance), and it allows evaluation of how well the teaching and learning system is working. As Salmi (2009, p. 7) writes, "accountability should not focus on the way institutions operate, but on the results that they actually achieve". Learner outcomes are a basic focus of learners, teachers, governments and graduate employers. Despite considerable investment and development, if students fail to learn and achieve, it is unlikely that teaching and support resources could be considered to be of high quality. Institutional policies may appear coherent, but would be hollow unless related to student and graduate outcomes. Focusing attention on all facets of the educational process is important for quality management and improvement, but this is not by itself sufficient.

Strangely, given the basic importance of learner outcomes, they are demonstrably absent from many quality assurance systems around the world. This is no doubt partly because the assessment of learner outcomes has traditionally been seen as the private preserve of teaching staff. Academic or institutional autonomy are frequently cited reasons for upholding the privacy of student assessment, even though these principles are sustained in research contexts that involve extensive forms of peer validation and review. Another reason may be that unlike research, teaching staff have not typically been rewarded for developing high-quality assessments, which would have an understandable impact on the motivation of staff to invest effort in this area. Of course, very few academics are specialists in psychometrics or student assessment and without training could hardly be expected to have a grasp of key principles and practices. Such training could inform teachers of options for developing good practice. The balance of this paper explores what this may entail.

It is important to set the preconditions for such work by creating a culture that values learner outcomes. Engaging staff in conversations about learner outcomes has been developed as a mechanism for doing this. In the United Kingdom, this culture has come through the development of benchmark statements (QAA, 2009). It is reflected in the learning outcomes frameworks developed as part of the European Tuning Process (EC, 2009) and the "metarubrics" developed in the United States by the Association of American

Colleges and Universities (AACU, 2009). Developing a focus on learner outcomes can occur through professional accreditation, or via the use of objective assessments of learner engagement and capability. In Australia, for instance, many institutions have developed lists of "graduate attributes" (see, for instance, Barrie, 2009). While such work does not go so far as to yield evidence of performance, it can provide a foundation for institution-wide thinking about the broad outcomes expected of learners and graduates.

There have been enormous advances in educational assessment over the last hundred years, much of which is captured by Linn (1989), Keeves (1988), AERA (1999), NRC (2001) and OECD (2003). Important aspects of this assessment methodology, however, have yet to be applied to higher education. Universities and academics are responsible for monitoring and assuring academic standards, and it is critical that continuous efforts are made to enhance the standard of assessment itself. Documentation of a validated suite of measurement approaches lies at the heart of any work on academic standards, for it provides the means that can be used to assure academic quality and enhance teaching and learning practice.

Currently, as contended at the beginning of this paper, many tasks are developed by individual teaching staff for use in specific subjects whose content may change in various ways from year to year. Teaching staff often develop such resources over relatively short periods of time, for localised purposes and with limited resources or knowledge of assessment methodology. As a result, student knowledge and skill is often measured using uncalibrated tasks with unknown reliabilities and validities, scored normatively by different raters using unstandardised rubrics and then, often with little moderation, adjusted to fit percentile distributions which are often specified *a priori* by departments, faculties or institutions.

It is possible to develop validated assessment tasks for a large number of higher education subjects. Achieving consistency across tasks can be vital, because variations in task severity will register as variations in student achievement, regardless of actual competence. Broad subjects based on a single textbook, which take a "shrink-wrapped" approach, can be accompanied by assessment materials. These materials can incorporate formative assignments for continuous assessment as well as validated examinations or items. The tasks themselves could be supported by notes for managing the assessment, analysing data, interpreting results and reporting achievement. A degree of flexibility would presumably need to be designed into the tasks to both encourage and support local adaptations. These assessments could be designed to fit different levels and fields of study, and may include performance tasks, portfolios, open ended questions, constructed response items and multiple choice questions. The validated tasks for these mass subjects could take many different forms, their defining characteristic being

that they are designed to optimise the measurement, diagnosis and enhancement of learning.

Many higher education subjects are specialised in nature or small in scale, however, and it may not be feasible to develop fully validated assessments. It is important, nonetheless, that the resource-consuming nature of assessment task design does not inhibit high-quality practice. In such instances, the most appropriate approach may be to train academic staff. An awareness of basic principles of assessment design and advanced practice would develop the capacity of teaching staff to enhance their own assessment tasks and activities. It would also have more general pedagogical benefits, by requiring academics to think not just about what and how they teach, but about what students are expected to learn and how they should be assessed.

Training teaching staff in assessment could be coupled with a process of assessment task review, in which technical experts or academic colleagues offer feedback on assessment tasks and approaches, and ensure that tasks are of appropriate quality. This feedback may reference quality criteria for student assessment. Of course, this currently happens for many courses and assessments (see, for instance: QAA, 2008), but the process is by no means universal. The largely individualised development of assessment tasks can make it difficult to develop informed and generalisable criteria which map out thresholds of increasing performance. It can be difficult, as a result, for institutions to assure the quality of the tasks which are themselves used to set academic standards.

Moderation processes might be used to ensure the generalisability of assessment standards and outcomes. In general, moderation requires teaching staff to review samples of student work to assure the comparability of standards across contexts. Such moderation may be conducted on an *ad hoc* basis, as often already occurs. It is preferable to design robust and scalable management systems, however, to ensure that outcomes can be quality assured. Moderation could be managed by a cross-institutional agency, as in many senior secondary contexts, or perhaps by a cluster of discipline-specific agencies. The UK External Examiner system illustrates one implementation of moderation in higher education (QAA, 2008). It might involve statistical calibration processes to help equate standards, highlight unusual scores and to manage moderation processes.

Along with the development of formative assessment practice, objective tests can be used to measure critical thinking, problem solving and numeracy skills. Such tests have become popular over the last decade for monitoring the standards of institutional provision. The most widespread, the Collegiate Learning Assessment (CAE, 2009), "presents realistic problems that require students to analyse complex materials and determine the relevance to the

task and credibility". Responses are analysed to assess critical thinking, analytical reasoning, problem solving and communication. The Graduate Skills Assessment test (ACER, 2009) uses multiple choice questions and writing tasks to assess critical thinking, problem solving, interpersonal understandings and written communication. The ETS Proficiency Profile (ETS, 2009) measures critical thinking, reading, writing and mathematics. These tests have the advantage of providing objective estimates of each participant's performance. Data provide external points of reference which can help validate assessment processes and inform moderation and final grading. Similar triangulation may be obtained by drawing, where appropriate, on licensing examinations, consistent feedback from graduate employers or professional bodies, or other information about the performance of graduates.

Translating advances into practice

The notion of "academic standards" touches on most areas of a university's operations, yet is complex and difficult to define. There is, accordingly, no single or simple means of defining, assessing, monitoring and enhancing academic standards. Any single approach is likely to be overly simplistic and prescriptive and, as a result, is likely to underestimate the complexity and significance of the matters at hand, running the risk of promoting standardisation rather than diversification of practice.

A series of new policies and practices would likely be required to support any extensive movement towards implementing the ideas in this paper. In specific instances, new practices might be developed from existing pockets of excellence, or transferred from sectors and systems in which they have been developed and tested over many decades. However, certain forms of more general cultural change would likely be required, along with substantial investments in training and systems. These changes and investments may be costly, but their value lies in the significant information that they would provide on individual learning and growth. Of course, the deeply embedded nature of work surrounding academic standards means that universities have an intrinsic warrant for continuous improvement in this area.

It is almost essential that academics and institutions themselves take the lead in developing this growing aspect of higher education. As Salmi (2009, p. 7) writes, "accountability works better when it is experienced in a constructive way"; "… the most effective accountability mechanisms are those that are mutually agreed or are voluntarily embraced by tertiary education institutions". This is not just because institutions have the authority to accredit their own programmes, ensure academic standards and underpin quality assurance processes. To the extent that student outcomes are stressed, it is vital that progress in this area builds on, rather than breaks, the

authority of teachers and institutions over the development, dissemination and assessment of knowledge. It is important that any measurement of student learning and development is itself collaborative in nature, given the broader individual, social and economic roles such measures will play. Further, it is vital that performance information is reported in multilevel ways that inform and support practice.

This multilevel emphasis is important because, while there are pockets of excellence, developing assessments of performance that simultaneously provide sound information to students, institutions and systems remains a major challenge for higher education. It is common for classroom practice to play out in completely separate ways to the organisation-level assessments managed by institutional researchers, and for both of these to operate in relative isolation from external quality assurance activities. There are of course distinct contexts and foci that shape practice at each of these levels of analysis, yet they share the common basic goal of assessing what learning has been achieved. There would appear to be value, therefore, in identifying approaches that work towards an integration across these levels of analysis.

In this paper it has only been possible to offer a preliminary analysis of the specification and measurement of academic standards. No attempt has been made, for instance, to consider how performance data might be best analysed or how it might be reported to individuals, institutions or regulatory or funding agencies. Further, this paper has not attempted to examine many of the important pedagogical and course management considerations implied by the propositions made about outcomes measurement. Many important questions also surround the value of absolute performance measures compared with measures of individual growth and value added. While graduates and employers often review trends in performance over the duration of a degree, this perspective is rarely captured explicitly in academic transcripts or reports.

In order to encompass the many possibilities for further development in this area, this paper has focused squarely on developing mechanisms for defining and monitoring academic standards in higher education. Facets of the analysis may be considered controversial, but these have been explored with the assumption that much would flow for institutions, systems and – crucially – individuals from making substantive progress in this area. Importantly, it has been contended that standards of education are, or should be, linked in important ways to a capacity to demonstrate that individual learning and development have taken place.

The author:

Dr. Hamish Coates
Australian Council for Educational Research (ACER)
19 Prospect Hill Road (Private Bag 55)
Camberwell, Victoria, 3124
Australia
E-mail: coatesh@acer.edu.au; hamishc@unimelb.edu.au

Ideas expressed in this paper were formed through several projects sponsored by Australasian universities, peak bodies and government agencies.

References

American Educational Research Association (AERA), American Psychological Association (APA) and National Council for Measurement in Education (NCME) (1999), *Standards for Educational and Psychological Testing*, AERA, Washington DC.

Anderson, D., R. Johnson and B. Milligan (2000), *Quality Assurance and Accreditation in Australian Higher Education: An assessment of Australian and International Practice*, Australian Government Department of Education, Training and Youth Affairs, Canberra.

Angoff, W.H. (1971), "Scales, Norms and Equivalent Scores", in R.L. Thorndike (ed.), *Educational Measurement* (2nd ed.), American Council on Education, Washington DC.

Association of American Colleges and Universities (AACU) (2009), *VALUE: Valid Assessment of Learning in Undergraduate Education*, AACU, Washington DC.

Astin, A.W. (1985), *Achieving Educational Excellence: A Critical Analysis of Priorities and Practices in Higher Education*, Jossey Bass, San Francisco.

Australian Council for Educational Research (ACER) (2009), *Graduate Skills Assessment*, ACER, Camberwell.

Australian Universities Quality Agency (AUQA) (2007), *AUQA Audits – 2007 Onwards: Overview of Audit Cycle 2*, AQU, Melbourne.

Australian Universities Quality Agency (AUQA) (2009), *Setting and Monitoring Academic Standards for Australian Higher Education: A discussion paper*, AUQA, Melbourne.

Barrie, S. (2009), *The National Graduate Attributes Project*, Australian Learning and Teaching Council, Sydney.

Bauer, R. (1966), *Social Indicators*, MIT Press, Cambridge.

Bottani, N. and A. Tuijnman (1994), "International Education Indicators: Framework, Development and Interpretation", in Centre for Educational Research and Innovation (CERI), *Making Education Count*, OECD Publishing, Paris.

Cave, M. *et al.* (1997), *The Use of Performance Indicators in Higher Education: The Challenge of the Quality Movement* (3rd ed.), Jessica Kingsley, London.

Coates, H. (2007a), "Developing Generalisable Measures of Knowledge and Skill Outcomes in Higher Education", in *Proceedings of AUQF 2007 – Evolution and Renewal in Quality Assurance*, Australian Universities Quality Agency, Melbourne.

Coates, H. (2007b), "Excellent Measures Precede Measures of Excellence", in *Higher Education Policy and Management,* Vol. 29, No. 1, pp. 87-94.

Coates, H. (2007c), *ATN Academic Standards Model*, Australian Technology Network of Universities, Adelaide.

Coates, H. (2007d), "Universities on the Catwalk: Models for Performance Ranking in Australia", in *Higher Education Management and Policy,* Vol. 19, No. 2, pp. 1-17, OECD Publishing, Paris.

Coates, H. (2008), *Attracting, Engaging and Retaining: New Conversations about Learning. Australasian Student Engagement Report*, Australian Council for Educational Research, Camberwell.

Coates, H. (2009a), "Building Quality Foundations: Indicators and Instruments to Measure the Quality of Vocational Education and Training", *Journal of Vocational Education and Training*, Vol. 61, No. 4, pp. 517-34.

Coates, H. (2009b), *What's the difference? A Model for Measuring the Value Added by Higher Education in Australia*, Higher Education Management and Policy, Vol. 21, No. 1, pp. 69-88, OECD Publishing, Paris.

Coates, H. and D. Edwards (2009), *The 2008 Graduate Pathways Survey: Graduates' Education and Employment Outcomes Five Years after Completion of a Bachelor Degree at an Australian University*, Department of Education, Employment and Workplace Relations, Canberra.

Council for Aid to Education (CAE) (2009), *Collegiate Learning Assessment* (CLA), CAE, New York.

Cuenin, S. (1988), *Performance Indicators in Higher Education: A Study of their Development and Use in 15 OECD Countries*, OECD Publishing, Paris.

Davis, D. (1996), *The Real World of Performance Indicators: A Review of their Use in Selected Countries*, Commonwealth Higher Education Management Service, London.

Department of Education, Science and Training (DEST) (2002), *Higher Education at the Crossroads,* DEST, Canberra, p. 15.

Educational Testing Service (ETS), (2009), *ETS® Proficiency Profile*, ETS, Princeton.

European Commission (EC) (2009), *European Tuning Process*, *http://unideusto.org/tuning*, accessed 1 February 2009.

Ewell, P. (2009), "Stuck on Student Learning", in *Measuring Up 2008: The National Report Card on Higher Education*, The National Centre for Public Policy and Higher Education, San Jose.

Ewell, P.T. and D.P. Jones (1996), *Indicators of "Good Practice",* in *Undergraduate Education: A Handbook for Development and Implementation*, National Centre for Higher Education Management Systems, Colorado.

Ikenberry, S. and G. Kuh (2009), *More Than You Think, Less Than We Need: Learning Outcomes Assessment in American Higher Education*, Indiana University, Bloomington.

Jaeger, R.M. (1978), "About Educational Indicators: Statistics on the Conditions and Trends in Education", in *Review of Research in Education,* Vol. 6, pp. 276-315.

James, R. (2003), "Academic Standards and the Assessment of Student Learning: Some Current Issues in Australian Higher Education", in *Tertiary Education and Management,* Vol. 9, No. 3, pp. 187-198.

Keeves, J.P. (ed.) (1988), *Educational Research, Methodology, and Measurement: An International Handbook*, Pergamon Press, Oxford.

Kells, H. (1993). *The Development of Performance Indicators in Higher Education: A Compendium of Twelve Countries*, OECD Publishing, Paris.

Linke, R. (1991), *Performance Indicators in Higher Education: Report of a Trial Evaluation Study Commissioned by the Commonwealth Department of Employment, Education and Training*, Australian Government Publishing Service, Canberra.

Linn, R.L. (ed.) (1989), *Educational Measurement*, Macmillan Publishing Company, New York.

McKinnon, K., S. Walker and D. Davis (2000), *Benchmarking: A Manual for Australian Universities*, Department of Education, Training and Youth Affairs, Canberra.

National Centre for Public Policy and Higher Education (NCPPHE) (2008), *Measuring Up 2008: The National Report Card on Higher Education*, The National Centre for Public Policy and Higher Education, San Jose.

National Commission on Excellence in Education (1983), *A Nation at Risk: The imperative for Educational Reform,* US Government Printing Office, Washington DC.

National Research Council (NRC) (2001), *Knowing What Students Know: The Science and Design of Educational Assessment*, NRC, Washington DC.

Organisation for Economic Co-operation and Development (OECD) (2003), *PISA 2003 Technical Report*, OECD Publishing, Paris.

OECD (2009a), *International Assessment of Higher Education Learning Outcomes* (AHELO) Feasibility Study, *www.oecd.org/edu/ahelo*, accessed 1 February 2009.

OECD (2009b), *Education at a Glance 2009,* OECD Publishing, Paris.

Quality Assurance Agency (QAA) (2008), *Academic Infrastructure*, *www.qaa.ac.uk*, accessed 1 February 2009.

QAA (2009), *Subject Benchmark Statements*, *www.qaa.ac.uk*, accessed 1 February 2009.

Salmi, J. (2009), "The Growing Accountability Agenda in Tertiary Education: Progress or Mixed Blessing?", in *Education Working Paper Series*, No. 16, the World Bank, Washington DC, p. vii.

Spellings, M. (2006), *A Test of Leadership*, US Department of Education, Washington DC.

Voluntary System of Accountability (VSA) (2009), *Voluntary System of Accountability*, *www.voluntarysystem.org*, accessed 1 February 2009.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Women in science: the persistence of gender in Australia

by
Sharon Bell
LH Martin Institute, The University of Melbourne, Australia

The high levels of participation of women in higher education are currently attracting an unprecedented degree of interest. In this context, a shift is also occurring in equity policy in many countries, and the question of gender equity is now framed by concerns around male participation and male underachievement. However, a number of significant international studies are providing evidence of persistent patterns of horizontal segregation (by discipline) and vertical segregation (by level of seniority and measures of esteem) of women in higher education and research. Research undertaken in Australia by the Federation of Australian Scientific and Technological Societies (FASTS) on women in science provides evidence that this segregation leads to attrition of women from the scientific professions. This paper argues that, not withstanding the need to focus on the most disadvantaged in terms of equity and social inclusion, the question of women in higher education is a half prosecuted agenda, and that premature abandonment of this agenda may impede successful realisation of broader social inclusion and diversity agendas. It may also impact negatively on productivity and innovation.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Les femmes dans la science : la rémanence de la question du genre en Australie

par
Sharon Bell
LH Martin Institute, Université de Melbourne, Australie

Les niveaux élevés de participation des femmes dans l'enseignement supérieur suscitent à l'heure actuelle un intérêt sans précédent. Dans ce contexte, on peut également observer un changement de la politique d'égalité des sexes dans de nombreux pays et la question de l'égalité entre les genres se pose désormais en termes de participation masculine et de contre-performance masculine. Toutefois, un grand nombre d'études internationales de grande envergure fournit la preuve de la rémanence de modèles de discrimination horizontale (par discipline) et de discrimination verticale (par âge et par mesures d'estime) des femmes dans l'enseignement supérieur et dans la recherche. Des recherches menées en Australie par la Fédération des sociétés australiennes pour les sciences et la technologie (FASTS) sur la question de la place des femmes dans la science prouvent que cette discrimination mène à l'exclusion des femmes des professions scientifiques. Cet article démontre que, en dépit de la nécessité de se concentrer sur les plus désavantagés en terme d'intégration sociale, la question de la place des femmes dans l'enseignement supérieur n'est traitée qu'à moitié et que l'abandon prématuré de cette question est susceptible d'entraver le succès de la réalisation d'une intégration sociale plus large et le souci de diversité. Cela pourrait également avoir des conséquences négatives sur la productivité et l'innovation.

Introduction

In October 2009, *University World News* published a special report entitled "Women in Higher Education" arguing that "look to many universities around the world in 2009 and you'll see women outnumbering men". Reports on student enrolments were included from various countries. We learn that in the United States, "from 1998 to 2008, the annual average increase in enrolments was greater for women than men". Overall, nearly 60% of all graduate students in autumn 2008 were women, and they comprised a larger share of total enrolees at the masters and graduate certification level (61%) as well as at doctoral level (51%). In Canada, a "steady climb" in female enrolments has been seen over the past ten years; in Ireland, the latest figures show that females now outnumber males 59% to 41% in Ireland's seven universities. In South Africa, 56% of all students are reported to be female, and in Australia for the past 20 years the "higher education student population has been dominated by women who have increased their numerical superiority over males year by year until now they comprise nearly 58% of the total student body". This national evidence is supported by UNESCO's *Global Education Digest 2009* which reports that "the number of female students in tertiary education rose six-fold between 1970 and 2007 compared with a quadrupling of male enrolments during the same period". In terms of graduation, "women outnumber men in 75 of 98 countries with available data" (*University World News*, 2009, Issue 0098).

Women no longer the second sex?

Such influential headlines, that paint a grim picture for men in higher education, disguise the complex and persistent patterns of gendered distribution of students by discipline, of staff by discipline, by level and by institutional type, and by researchers against measures of excellence and esteem. A closer reading of the above reports hints at the perpetuation of this gendered distribution: according to the UNESCO *Digest* men outnumber women "in engineering, manufacturing and construction in all countries for which data were available", with the opposite proving true "for the cohort of graduates in education, humanities and arts, social sciences, business and law, and health and welfare, where, in almost nine out of ten countries women outnumber men". In the United States "the report confirms patterns of gender-driven field differentiation throughout the decade. Women were more

likely than men to be found in health sciences, education, public administration and services, and similar programmes. Although there were annual increases in engineering, physical sciences and business for women, in 2008 men continued to account for more than half (50.4%) of enrolees." In South Africa, according to the *Higher Education Monitor*, men continue to dominate in science, engineering and technology where they made up 57% of enrolments in 2007. In all other fields of study, more women are enrolled than men. Thus, women comprised 56% of students in business, commerce and management, 59% in the human and social sciences, and 73% in education. The report points out that areas with the greatest gender imbalances are engineering and engineering technology where only 24% of students are women, and health care and health sciences where only 32% of students are men. In Australia, "as is the case across most of the developed world, women have long held top spot in the 'typically female' professions of education and health (where three out of four Australian students are women), food and hospitality, the arts and humanities, and creative arts (where almost two in every three are female)" (*University World News,* 2009, Issue 0098).

Over the past decade, a number of influential national and international reports have drawn attention to these patterns, and concern has particularly been voiced over the persistent patterns of gender inequity in the science and technology disciplines and the implications of this. The international benchmark study in this field was undoubtedly that conducted at Massachusetts Institute of Technology. In 1995, the MIT Dean of Science established a committee to analyse the status of women faculty in the School of Science. The committee discovered that:

> … junior women faculty feel well supported within their departments and most do not believe that gender bias will impact their careers. Junior women faculty believe, however, that family-work conflicts may impact their careers differently from those of their male colleagues. In contrast to junior women, many tenured women faculty feel marginalized and excluded from a significant role in their departments. Marginalization increases as women progress through their careers at MIT. Examination of data revealed that marginalization was often accompanied by differences in salary, space, awards, resources, and response to outside offers between men and women faculty with women receiving less despite professional accomplishments equal to those of their male colleagues. **An important finding was that this pattern repeats itself in successive generations of women faculty. The Committee found that, as of 1994, the per cent of women faculty in the School of Science (8%) had not changed significantly for at least 10 and probably 20 years.** (MIT, 1999, p. 4, emphasis added)

In 1997, the United States National Science Foundation launched the Professional Opportunities for Women in Research and Education (POWRE) programme to support the development of scholarly and institutional leaders in research and education. The POWRE programme was designed to increase the prominence, visibility and influence of women in all fields of academic science and engineering. POWRE sought to enhance women's professional advancement by providing them with funding opportunities that were not ordinarily available through regular research and education grants programmes (National Science Foundation, 1997). In 2001, the National Science Foundation initiated the ADVANCE programme to replace POWRE. The ADVANCE programme was designed to improve the institutional climate, and the recruitment and retention of women faculty in science and engineering (National Science Foundation, 2009).

The Athena Project, funded by a UK-based consortium, was established in the United Kingdom in 1999 to increase the retention and advancement of women scientists in higher education employment (Bebbington, 2002). Under the Athena Project, three ASSET surveys (from 2003 to 2006) were conducted on the career experiences of male and female scientists working in higher education and research in the United Kingdom. The data from these surveys showed that research culture, networking and provision of childcare were crucial to the retention of women researchers (Kingston-Smith, 2008). The Athena Project actively engaged higher education institutions to promote the understanding of employment and cultural practices common to the scientific community. Good practice guidelines were developed covering areas such as establishing personal and professional support (mentoring, networks and career development), having supportive departmental heads, establishing a departmental culture that supports work-life balance, etc. Under Athena, the Athena Scientific Women's Academic Network (SWAN) Charter was developed in 2005 to recognise and celebrate good employment practice for women working in science, engineering and technology (SET) in higher education and research (*www.athenaswan.org.uk*).

But progress has been slow. An OECD sponsored workshop (OECD, 2006) found that women continue to be concentrated in lower-level positions: "just over one-third of United States university faculty are women and this figure is even lower in EU countries. [...] In the EU, women make up less than 20% of senior academic staff in the majority of countries." In Europe, "the percentage of women in the top grades of academia rarely exceeds 20%, and men are three times more likely than women to obtain professorships or their equivalent". Moreover, whilst the number of female researchers has increased, this is at a slow rate and off a low base. Thus "women account for 25-35% of researchers in most OECD countries, with the exception of Japan and Korea (12% each)". Women tend to be concentrated in fields and industries such as

biology, health, agriculture and pharmaceuticals, with low representation in physics, computing and engineering.

The United States National Academies of Sciences, Engineering, and the Institute of Medicine, under the oversight of the Committee on Science, Engineering and Public Policy, created the Committee on Maximizing the Potential of Women in Academic Science and Engineering. Their report (2007) found that "With every step up the academic ladder, from high school on through full professorships, the representation of women in science and engineering drops substantially". There is not a lack of supply – "the problem is not simply the pipeline":

> For over 30 years, women have made up over 30% of the doctorates in social sciences and behavioural sciences and over 20% in the life sciences. Yet, at the top research institutions, only 15.4% of the full professors in the social and behavioural sciences and 14.8% in the life sciences are women – and these are the only fields in science and engineering where the proportion of women reaches into the double digits. […] Women faculty are paid less, are promoted more slowly, receive fewer honors, and hold fewer leadership positions than men (*National Academies*, 2007).

At the same time, the US Athena Project (targeting women in SET who had embarked on careers within their organisations) arose from a 2006 Hidden Brain Drain private sector task force comprising 43 global companies. Between March 2006 and October 2007, the Hidden Brain Drain task force conducted four major surveys of men and women and included 28 focus groups. The task force was concerned that women were concentrated in lower level positions and attrition rates of women are high: "41% of highly qualified scientists, engineers, and technologists on the lower rungs of corporate career ladders are female." At the very time when labour shortages in SET are worsening the Athena Project found that "52% of highly qualified females working for SET companies quit their jobs, driven out by hostile work environments and extreme job pressures". The study concluded that women in SET are "marginalised by lab coat, hard hat, and geek workplace cultures that are often exclusionary and predatory".

Women in science in Australia: still the second sex

The latest student data from the Department of Education, Employment and Workplace Relations (DEEWR) shows that women make up 55% of all undergraduate students and 51.8% of postgraduate students (DEEWR, 2007). However, consistent with the international data cited above, the number of female students is not evenly distributed between the different fields of education. The health and education fields have the highest numbers of female students at 72.9% and 74.0% respectively. This is in marked contrast to

the fields of engineering and information technology where the numbers of female students make up only 15.5% and 18.9% respectively. Other fields such as natural and physical sciences (52%), management and commerce (48.5%) and creative arts (63.2%) hover around 50-60%. This uneven representation of women in the different areas of education (and the workforce) is known as horizontal segregation (Carrington and Pratt, 2003).

In recent decades, there have been many initiatives to encourage more women in science, engineering and information technology. These measures have succeeded in increasing the numbers of female students in some fields. However, more in-depth study of the data shows that women's participation is relatively low in particular disciplines (narrow fields of education) such as mathematical sciences, physics and astronomy and earth and chemical sciences, as well as all fields of information technology and engineering. In fields such as biological sciences as well as in agriculture and environmental studies there is significant participation of women. It should be noted however that where increases in participation at undergraduate level have been achieved this is in part due to the introduction of a new field of education classification in 2001 which introduced the field of information technology, thereby artificially increasing the percentage of women in the field of natural and physical sciences.

This uneven representation of women in the different fields of science is also a manifestation of horizontal segregation, well-documented in the literature (Langford, 2006; Queensland Government Office for Women, 2006; Cervantes, 2006). Moreover, snapshot data clearly indicates that even when relatively high levels of participation at undergraduate and even post-graduate levels have been achieved there are persistently low levels of representation of women at senior levels of the academy – evidence of vertical segregation (Carrington and Pratt, 2003; Bell and Bentley, 2005). The research of Castleman *et al.* (1995) on the payroll data of a sample of universities suggests that women are distributed unevenly amongst high- and low-demand disciplines and that where they form a sizeable minority of academics in high-demand disciplines they remain concentrated in the lower levels of the classification structure).

In 2008, the Federation of Australian Scientific and Technological Societies (FASTS) commissioned a report (Bell *et al.,* 2009) in response to two long-term issues around women's participation in science and technology: first, increasing concern in the Commonwealth Government regarding levels of participation in science subjects in the senior years of high school and the flow-on effects of this; second, growing awareness of the looming personnel shortages facing the academic and research sectors. Australia's science knowledge and skills base is fragile. Many research fields are increasingly dependent on international talent and securing such talent is becoming

extremely competitive. The report argues that identification of strategies to correct these trends through the participation and retention of women is crucial. Some of the major findings of this report are outlined below.

For comparative purposes, the FASTS study was able to draw on a 1995 discussion paper by the Women in Science, Engineering and Technology Advisory Group (WISET). The Advisory Group was tasked to advise the Commonwealth Government on strategies to improve women's participation in SET careers and education. In the 1995 discussion paper, it was argued that:

> Women remain seriously under-represented in some specific disciplines of science, engineering and technology (SET), and furthermore, are not well-represented at the most senior levels in all disciplines. This problem is poorly understood since statistics actually show a significant improvement in women's participation overall in SET-based education, training and employment over the last decade. […] Women make up 51% of the nation's population. Using their talents to the full at all levels of scientific and technological education, training and employment is an economic necessity, and an investment in Australia's future national development. The Advisory Group believes that continued under-representation and under-participation of women in SET-based education, training and employment is not only a cause for social concern on equity grounds, it is also likely to inhibit Australia's capacity to develop internationally competitive research and industries (*Women in Science, Engineering and Technology*, 1995, pp. 5-6).

The minimal rates of change since the production of the 1995 discussion paper are captured in comparative data that directly compare the statistical information presented in the 1995 report with the most current data available. Over the 16-year period from 1992 to 2008 the data indicate a small (2.8%) overall increase in women's participation in the workforce (from 42.3% in 1992 to 45.1% in 2008). The most significant changes registered were in the feminisation of traditional areas of female employment: a 14.2% growth in women's participation in community services (growing from 65.8% to 80%) and an 18.8% growth in government administration and defence (37.1% in 1992 to 55.9% in 2008). In traditionally male-dominated fields there were small to moderate increases in women's participation: 1.8% in agriculture, forestry and fishing (from 29.6% in 1992 to 31.4% in 2009), 3.9 percentage points in transport and storage (19.5% to 23.4%), 5.7% in mining (9.5% to 15.2%) and 12.6% in electricity, gas and water supply (13.6% to 26.2%).

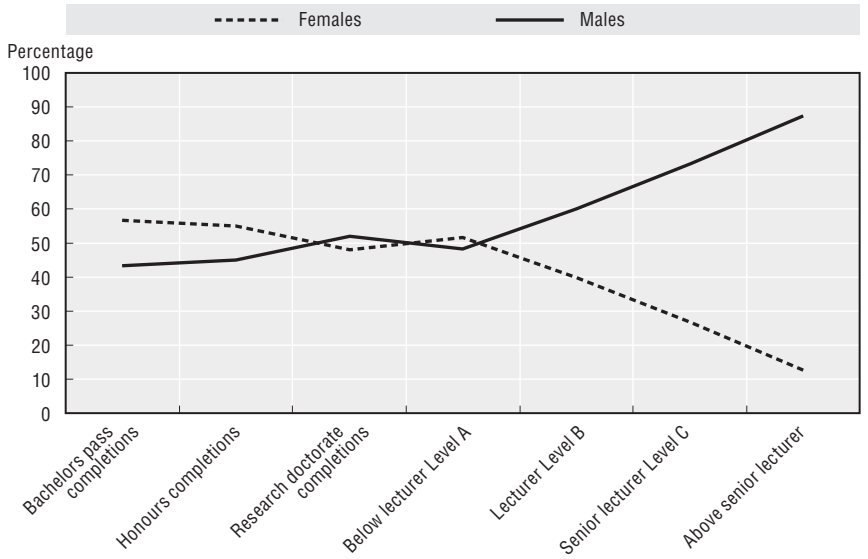
In terms of leadership, the category of female administrators and managers has grown less than 4% from 25.1% in 1992 to 29.0% in 2008. Yet female participation in professional fields has increased over 11% from 42.4% to 53.4%.

More specific occupational data indicates that women constituted only 18.1% of full-time professionals in the field of design, engineering science and transport in 1996 and this only grew by 4.2% to 22.3% in 2009. ICT professionals did not fare as well; while females constituted 19% of all full-time ICT professionals in 1996, this number fell by 3.8% to 15.2% in 2008. The percentage of women in full-time engineering, ICT and science technician roles similarly dropped from 18.9% in 1996 to 17.1% in 2008. Moreover, earnings for women in these fields, and indeed in highly feminised fields, have remained consistently lower than their male counterparts.

Figures 1 to 4 are constructed from the percentages of male and female completions at each level of higher education, and from academic staff profile data in 2007. In Australia, academic positions are structured into five levels, from Level A through to Level E. Level A is the equivalent of an associate lecturer and Level E is a full professor. Seniority may be regarded as a proxy measure of success.

This data is **not** cohort-based longitudinal data.

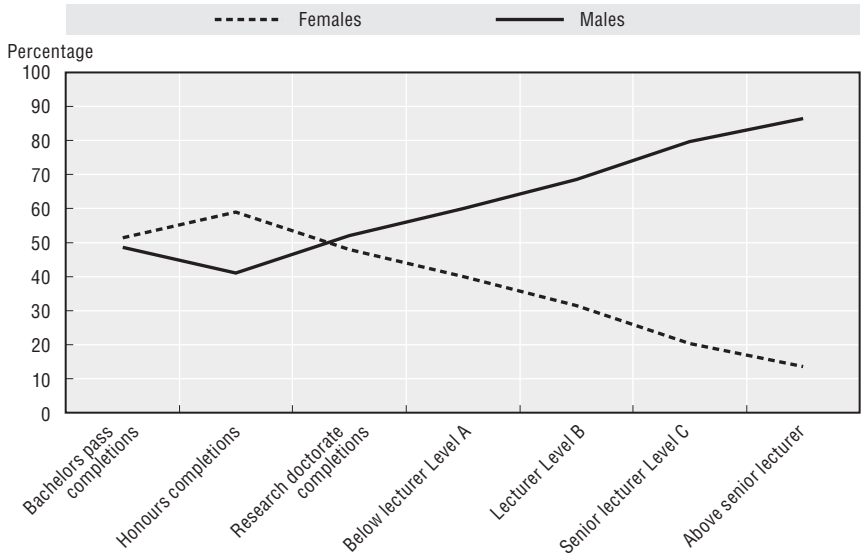
Figure 1. **Academic profiles by gender**
Natural and physical sciences, 2007



Source: DEEWR Selected Higher Education Student Statistics 2007; Department of Education, Science and Training, *Special Report*, FTE Staff in AOU Groups, 2007.

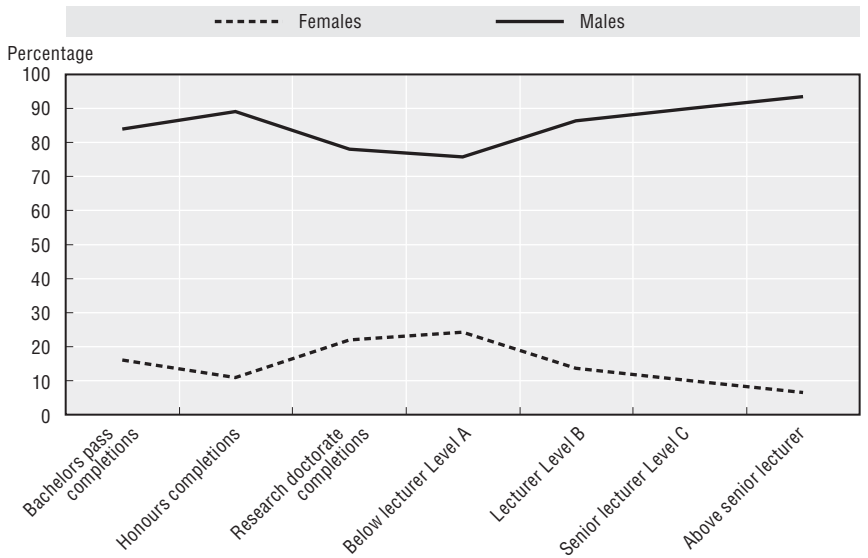
The fact that women are under-represented in senior academic positions does indicate barriers to "success" but does not necessarily equate to attrition. Women may be moving from the academy into productive work in industry,

Figure 2. **Academic profiles by gender**
Agriculture and environmental studies, 2007



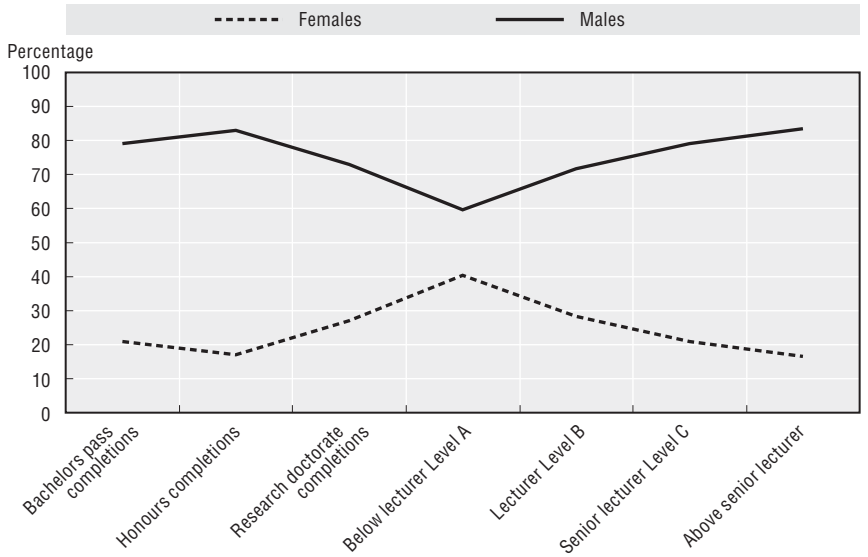
Source: DEEWR Selected Higher Education Student Statistics 2007; Department of Education, Science and Training, *Special Report*, FTE Staff in AOU Groups, 2007.

Figure 3. **Academic profiles by gender**
Engineering and related technologies, 2007



Source: DEEWR Selected Higher Education Student Statistics 2007; Department of Education, Science and Training, *Special Report*, FTE Staff in AOU Groups, 2007.

Figure 4. **Academic profiles by gender**
Information technology, 2007



Source: DEEWR Selected Higher Education Student Statistics 2007; Department of Education, Science and Training, *Special Report*, FTE Staff in AOU Groups, 2007.

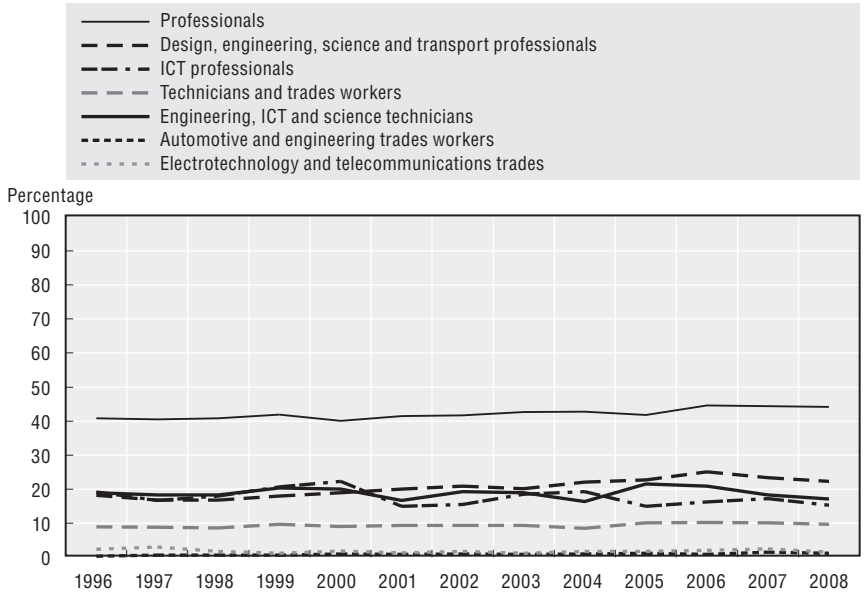
government or not-for-profit sectors commensurate with their knowledge and skills base. However, the absence of reliable data that tracks mobility of the scientific workforce between universities, industry and government means it is much harder to evaluate whether there is net attrition or simply a wide range of graduate and postgraduate outcomes. Available data from government agencies and industry does however suggest attrition, as the pattern of feminised lower ranks and male dominated senior ranks is replicated.

Scientific careers

At the Australia 2020 Summit organised by the Australian government it is reported that there was extensive discussion on improving the career structure in research, with the need for research to be better recognised as a "profession":

> The importance of having the "best and brightest" being more disposed to return to Australia and work here was seen as critical. Development of a more secure career structure for researchers was seen as a way of achieving this. The current lack of a clear career path was seen as an issue for retention of young researchers, with issues including ensuring job security and meeting changing demands for work–life balance to take account of family needs. […] There is a clear need for a shift in how we

Figure 5. **Occupational categories and sub-categories**
By percentage of females, 1996-2008



Source: ABS Labour Force, Australia, Detailed, Quarterly, May 2009.

> see research as a career path and not just as being tacked onto the end of a science degree. Well-established mentorship is seen as important. Better recognition for research will encourage recruitment and retention, including **retention of women at senior levels, which remains poor.** (Australian government, 2008, p. 161, emphasis added)

The representation of women in the science workforce is dependent on the proportion of female students who undertake science degrees in universities and qualifications in the vocational education and training (VET) sector. There are many potential employment opportunities for students who undertake a career in science and technology. The report by Stevens-Kalceff and colleagues (2007) on the academic profile of the School of Physics at the University of New South Wales shows that a typical physics male follows a linear career path from PhD through the academic ranks from A to E. Conversely, a typical physics woman follows a non-traditional career path (Figure 6) due to several factors such as career breaks, not having undertaken a postdoctoral appointment, and family responsibilities (Stevens-Kalceff *et al.*, 2007). These factors ultimately lead to low levels of female participation and to limited representation of women at the more senior levels of the academic and research sectors.

Figure 6. **The non-traditional academic career progression**



Source: Stevens-Kalceff *et al.*, 2007.

The fact that women are still under-represented in the academic and research sectors of science and technology leads to the conclusion that there are barriers or obstacles that prevent a woman from progressing through an appropriate career path. Several metaphors have been introduced in an attempt to describe such barriers or obstacles. The "glass ceiling" and the "leaking pipeline" are two such metaphors.

The term "glass ceiling", first coined in 1986 by *The Wall Street Journal's* Carol Hymowitz and Timothy Schellhardt, describes an invisible barrier which women encounter as they progress through the ranks of their career. In more recent times, the efficacy of the glass ceiling metaphor has been questioned as it implies a single barrier in a linear career which women are unable to overcome.

Berryman (1983) first conceptualised the "leaking pipeline" metaphor, which has become the commonly accepted paradigm describing the attrition of women along their career path in science. In more recent times, the leaking

pipeline metaphor has been argued to be an oversimplified representation of the attrition of women from science (Soe and Yakura, 2008). Again, this metaphor fails to convey the complexities women encounter in their academic and research careers.

To quote Eagly and Carli:

> ... times have changed, and the glass ceiling metaphor is now more wrong than right [because] it describes an absolute barrier at a specific high level in organisations. The fact that there have been female chief executives, university presidents, state governors, and presidents of nations gives the lie to that charge. At the same time, the metaphor implies that women and men have equal access to entry and mid-level positions when in fact they do not. The image of a transparent obstruction also suggests that women are being misled about their opportunities, because the impediment is not easy for them to see from a distance. But some impediments are not subtle. Worst of all, by depicting a single, unvarying obstacle, the glass ceiling fails to incorporate the complexity and variety of challenges that women can face in their leadership journeys. In truth, women are not turned away only as they reach the penultimate stage of a distinguished career. They disappear in various numbers at many points leading up to that stage (2007, p. 64).

In a paradigm shift, Eagly and Carli (2007) propose a new metaphor, that of a labyrinth, which conveys the idea of women's complex journey towards a worthy goal through a passage that is neither simple nor direct but which "requires persistence, awareness of one's progress, and a careful analysis of the puzzles that lie ahead". In their view, "routes exist for women who aspire to top leadership, but [those routes] are full of twists and turns, both unexpected and expected. Because all labyrinths have a viable route to the centre, it is understood that goals are attainable. The metaphor acknowledges obstacles but is not ultimately discouraging" (2007, p. 64).

In keeping with Eagly and Carli's paradigm, the obstacles that women face in their careers (the factors that affect career progress and research output such as career breaks, lack of time for research, teaching loads that are prohibitive of research) resemble the twists and turns of a labyrinth. As one female scientist reflects:

> It seems to me that, as a female, as you work your way further along the career path in science, it seems to get harder, rather than easier: there are less female role models at every stage; the "game" becomes more competitive and complex and involves not just what you know, but who you know (or who you are buddies with); the job requires you to work long hours (while these are not fruitless, they are strenuous and often not possible if you have out-of-work commitments); you are required to

> publish consistently excellent results in order to stay competitive with the field; and this is all on top of the complexities of finding the time to start and raise a family, and time for general life-work balance (Hatchell and Aveling, 2008, p. 11).

The reasons for the low representation of women in science and technology can be separated into two broad categories. First, horizontal segregation of women in the various science disciplines based on perceptions regarding women's innate ability in science and mathematics, societal attitudes towards gender stereotypes and gender equality, and job security and employability of science graduates. Second, vertical segregation, generated by the organisational culture of the workplace through practices that disadvantage women such as work load, promotions policies and practice, sex discrimination, lack of female role models, mentors and networks, family responsibilities and so on.

We know from the literature (Probert, 1998; Hobson *et al.*, 2003; Stevens-Kalceff *et al.*, 2007) that women make strategic decisions about career and life options at critical points in their careers, particularly at the conclusion of doctoral studies – which appears to be a "tipping point" (Bell and Bentley, 2005). These may be seen as "situationally" temporary but turn out to be career limiting (Hobson, 2003). As Probert (1998, p. 52) points out in her critique of proponents of human capital theory, women do not always have a choice in any meaningful sense: "If they wish to pursue an academic career as vigorously as men do, they may feel that this is incompatible with responsible family life. In this sense women have to make a choice between work and family which men can avoid, and this renders the scope for 'choice' incommensurate." Following Eagly and Carli (2007) these "choices" are made many times over as women negotiate complex career labyrinths.

Success and esteem: some are more equal than others

There are a number of measures, in addition to publications, that may be employed to ascertain scientific excellence or esteem, or at least recognition of scientific excellence through peer assessment. In the FASTS report four measures were considered: Australian Research Council (ARC) Discovery and Linkage grants; the ARC Federation Fellow Scheme; admission to the learned academies of Science and Technological Sciences and Engineering; and the newly introduced ARC Future Fellows scheme.

Analysis of data on ARC grant schemes reveals some clear trends. All grant schemes have seen significant increases in applications over recent years. Success rates for male and female applicants are generally comparable. However, participation rates for women are lower than for males – in the case of the Discovery and Linkage grants, significantly so.

Similarly, women make up only 8.5% of ARC Federation Fellows, the fellowships designed to attract world-class researchers and world-class research leaders to key positions. The Federation Fellow scheme has been in place since 2001, and there has been little annual variation in this figure.

In 2008, the Australian Government announced the creation of a new scheme, the ARC Future Fellowships. The scheme is intended to promote research in areas of critical national importance by giving outstanding researchers incentives to conduct their research in Australia. The aim of ARC Future Fellowships is to attract and retain the best and brightest mid-career researchers. Results of the first Fellowship round have been announced. Women constituted 29% of applicants (283 of 975) and secured just under 29.5% of the Fellowships. But the largest number of successful female applicants (36 of 59) was clustered in the lowest band (Salary Level 1).

In the learned Academy of Science women constitute only 7% of Fellows. The Academy has 426 Fellows of whom 30 are women; just 1% more than five years ago. In the learned Academy of Technological Sciences and Engineering 6% of Fellows are female. There are 45 female Fellows from a Fellowship of 788, compared with 5% five years ago.

In this context, it is encouraging to note that currently a number of key leadership roles are occupied by women: the Chief Scientist, the NSW Chief Scientist, the CEO of CSIRO and the CEO of the ARC to name but a few. Nonetheless, without significant female seniority in the sector, this profile of leadership – arguably based on individual achievement – is fragile.

It is salutary in this regard that the Leadership Census (2008, pp. 3-5) launched by the Equal Opportunity for Women in the Workplace Agency (EOWA) reveals that across all indicators, the proportion of women to men on corporate boards and in executive leadership roles in the ASX200 companies has declined since 2006. At the time of the Census, women chaired only four boards and held only 8.3% of board directorships (125 seats out of 1 505), down from 8.7% in 2006. On these measures, Australia has now fallen behind the United States, Canada, the United Kingdom and South Africa.

Conclusion

The outset of this paper documented the current concern around female *versus* male participation in higher education. What is fascinating in terms of the higher education policy agenda is that this is, in some ways, as evidenced above, based on a superficial analysis of gender profiles in the academic and research environments – analysis that fails to acknowledge persistent patterns of gendered participation and success. This generates the risk of premature abandonment of an agenda half prosecuted.

It is also intriguing that the gains that have been recorded for women in education in the past two decades are reported as if illustrating the success of one equity group over others – a threat to men and boys in education. Indeed if there is cause for cautious optimism, it is through the examples of extraordinary individual achievement of women in the academy and in research and the systemic change that has occurred at the secondary school level and in some disciplines at the undergraduate and postgraduate levels in terms of women's participation and success. This should provide evidence that such change can be achieved if a clear agenda is set and pursued – an agenda that will be critical if we are to achieve more ambitious social inclusion, goals and diversity outcomes.

Indeed the next steps to secure full participation, productivity and success of women may be more demanding than the strategies of "accommodation" of women (such as through provision of maternity leave and child care) to established (male) patterns of work that we have seen over the past two decades. As Australian Sex Discrimination Commissioner Elizabeth Broderick has noted recently, the solution does not lie in making it necessary for women to work like men.

> It lies in both genders being able to work, and be remunerated, equally in the industries in which they wished to work, without being judged by or buying into stereotypes about gender and particular industries. It also lies with the opportunity to work flexibly so that both genders can involve themselves successfully in paid work and care (HEREOC Media Release, 1 December 2009).

The FASTS report argues that the agenda now requires serious attention to changing the "context", to reframing scientific career paths in tandem with institutional cultures and decision making (including composition of decision-making bodies).

The author:

Prof. Sharon Bell
Senior Program Developer
LH Martin Institute for Higher Education Leadership and Management
University of Melbourne
Victoria 3010
Australia
E-mail: s.bell@unimelb.edu.au

The funding source supporting the research presented was the Australian Government Department of Innovation, Industry, Science and Research.

References

Athena Project (2005), "UK Assets: 6 500 plus UK Scientists Can't be Wrong", the Athena Surveys, *www.sci-soc.net/NR/rdonlyres/FDCB6A81-8E09-4840-86CE-4C3523919FCF/335/AAAS05CFpresentation.pdf*, accessed 19 June 2009.

Australian Bureau of Statistics (ABS) (2001), *Australian Standard Classification of Education (ASCED)*, ABS Catalogue No. 1272.0, from *www.ausstats.abs.gov.au/ausstats/free.nsf/0/2D08D604CC032093CA256AAF0080766F/$File/12720_2001.pdf*, accessed 12 June 2009.

Australian Government (2008), *Australia 2020 Summit – Final Report*, Department of the Prime Minister and Cabinet, *www.australia2020.gov.au/final_report/index.cfm*, accessed 10 August 2009.

Bebbington, D. (2002), "Women in Science, Engineering and Technology: A Review of the Issues", *Higher Education Quarterly,* Vol. 56, No. 4, pp. 360-375.

Bell, S. *et al.* (2009), *Women in Science in Australia: Maximising Productivity, Diversity and Innovation,* FASTS, Canberra.

Bell, S. and R. Bentley (2005), "Women in Research", AVCC National Colloquium of Senior Women Executives, *www.universitiesaustralia.edu.au/documents/policies_programs/women/Bell-BentleyAVCC.05.pdf*, accessed 10 May 2009.

Berryman, S. (1983), *Who will Do Science? Minority and Female Attainment of Science and Mathematics Degrees: Trends and Causes,* Rockefeller Foundation, New York.

Carrington, K. and A. Pratt (2003), "How Far Have We Come? Gender Disparities in the Australian Higher Education System", *Information, Analysis and Advice for the Parliament*, p. 7, *www.universitiesaustralia.edu.au/documents/policies_programs/women/Gender_Disparities_20Report_Jun03.pdf*, accessed 27 May 2009.

Castleman, T. *et al.* (1995), *Limited Access: Women's Disadvantage in Higher Education Employment*, National Tertiary Education Union, Melbourne, pp. 46-48.

Cervantes, M. (2006), "OECD Background Paper", Chapter 1, *Women in Scientific Careers: Unleashing the Potential*, OECD Publishing, Paris.

Department of Education, Employment and Workplace Relations (DEEWR) (2008), "Students 2007, Selected Higher Education Statistics", *Higher Education Statistics, www.dest.gov.au/sectors/higher_education/publications_resources/profiles/Students_2007_full_year_.htm*, accessed 12 June 2009.

DEEWR (2009) "Staff 2008: Selected Higher Education Statistics", *Higher Education Statistics*, *www.dest.gov.au/sectors/higher_education/publications_resources/profiles/Staff_2008_selected_higher_education_statistics.htm*, accessed 12 June 2009.

Eagly, A. and L. Carli (2007), "Women and the Labyrinth of Leadership", *Harvard Business Review*, September, pp. 63-69.

Equal Opportunity for Women in the Workplace Agency (EOWA) (2008), *Australian Women in Leadership Census*, *www.eowa.gov.au/Australian_Women_In_Leadership_Census/2008_Australian_Women_In_Leadership_Census/Media_Kit/EOWA_Census_2008_Publication.pdf*, accessed 17 July 2009.

Hatchell, H. and N. Aveling (2008), "Gendered Disappearing Acts: Women's Doctoral Experiences in the Science Workplace", *Australian Association for Research in Education*, Queensland.

Hobson, J., G. Jones and E. Deane (2005), "The Research Assistant: Silenced Partner in Australia's Knowledge Production?", *Journal of Higher Education Policy and Management*, Vol. 27, No. 3, pp. 357-366.

Kingston-Smith, A. (2008), "Wisdom, Justice and Skill in Science, Engineering and Technology: Are the Objectives of the Athena Project Mythology?", Society for Experimental Biology Bulletin, *www.sebiology.org/publications/Bulletin/March08/Athena.html*, accessed 13 July 2009.

Langford, P. (2006), "Attracting and Retaining Staff Engaged in Science, Engineering and Technology: A Review of National and International Best Practice", *www.dest.gov.au/NR/rdonlyres/3EA5DE36-96D0-4CF6-BC95-AE6695D1B942/25868/AttractandRetainSETEmployeesLiteratureReviewFinalr.pdf*, accessed 10 May 2009.

Massachusetts Institute of Technology (MIT) (1999), "A Study on the Status of Women Faculty in Science at MIT", *http://web.mit.edu/fnl/women/women.pdf*, accessed 23 May 2009.

National Academies: Committee on Science, Engineering, and Public Policy (2007), "Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering", Washington DC, The National Academies Press, *www.nap.edu/catalog.php?record_id=11741#toc*, accessed 22 November 2007.

National Science Foundation (NSF) (2008), "Graduate Students and Postdoctorates in Science and Engineering: Fall 2006", *www.nsf.gov/statistics/nsf08306*, accessed 12 June 2009.

National Science Foundation (NSF) (2009), "ADVANCE: Increasing the Participation and Advancement of Women in Academic Science and Engineering Careers", National Science Foundation, *www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf0941*, accessed 13 July 2009.

Organisation for Economic Co-operation and Development (OECD) (2006), *Women in Scientific Careers: Unleashing the Potential*, OECD Publishing, Paris.

Probert, B., P. Ewer and K. Whiting (1998), "*Gender Pay Equity in Australian Higher Education*", The National Tertiary Education Union, Melbourne.

Queensland Government Smart Women – Smart State Taskforce (2006), "Supporting Women's Participation in Emerging Industries: Science, Engineering and Technology", *www.women.qld.gov.au/work-and-life/smart-state-strategy*, accessed 28 September 2009.

Soe, L. and E. Yakura (2008), "What's Wrong with the Pipeline? Assumptions about Gender and Culture in IT Work", *Women's Studies*, Vol. 37, No. 3, pp. 176-201.

Stevens-Kalceff *et al.* (2007), "Maximising Potential in Physics: Investigation of the Academic Profile of the School of Physics at the University of New South Wales", *http://phys.unsw.edu.au/phys_about/EQUITY/Maximising%20Potential%20in%20Physics.pdf*, accessed 12 May 2009.

UNESCO Institute for Statistics (2009), *UNESCO Global Education Digest: Comparing Education Statistics Across the World*, UNESCO, Montreal.

University World News (2009), Issue No. 0098, 25 October 2009, *www.universityworldnews.com*, accessed 29 October 2009.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Quality assurance in higher education as a political process

by
Michael L. Skolnik
University of Toronto, Canada

The procedures commonly employed for quality assurance in higher education are designed as if the endeavour were a technical process, whereas it may be more useful to view it as a political process. For example, quality assurance requires making choices among competing conceptions of quality, and in so doing privileges some interests over others. Moreover, some stakeholders tend to be given a greater voice than others in the design and implementation of quality assurance. The author concludes that rather than denying the political nature of quality assurance, it would be better to accept Morley's claim that quality assurance is "a socially constructed domain of power", and design procedures for it in a way that is appropriate for a political process. It is suggested that employing the "responsive model" of evaluation could make quality assurance more effective in improving educational quality. In the responsive model, evaluation is deemed to be a collaborative process that starts with the claims, concerns and issues put forth by all stakeholders.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Assurance qualité de l'enseignement secondaire évaluée en tant que processus politique

par
Michael L. Skolnik
Université de Toronto, Canada

Les méthodes généralement utilisées pour évaluer l'assurance qualité de l'enseignement secondaire s'apparentent plus à un processus technique, alors qu'il serait plus bénéfique d'analyser cette question d'un point de vue politique. Par exemple, l'assurance qualité implique de faire des choix parmi les diverses conceptions existantes du mot qualité et, ce faisant, elle privilégie certains intérêts plutôt que d'autres. En outre, certaines parties prenantes ont plus d'influence que d'autres au niveau de la conception et de la mise en œuvre de l'assurance qualité. L'auteur en conclut qu'au lieu de renier la nature politique de l'assurance qualité, il serait préférable d'accepter la thèse exposée par Morley *selon laquelle l'assurance qualité est « un domaine de pouvoir construit sur un modèle social », et de mettre en place à cet égard des procédures conformes à un processus politique. Il est suggéré que l'utilisation du « modèle sensible » pourrait rendre l'assurance qualité plus efficace, notamment en termes d'amélioration de la qualité de l'enseignement. Dans le cadre du modèle sensible, l'évaluation est perçue comme un processus de collaboration avant tout basé sur les réclamations, les préoccupations et les problèmes mis en avant par l'ensemble des parties concernées.*

Introduction

Quality assurance in higher education has become a major phenomenon worldwide. For example, the International Network for Quality Assurance Agencies in Higher Education (INQAAHE), which started in 1991 with eight members, now has more than two hundred members (International Network, 2010). Not only has quality assurance become widespread, but its impacts can be significant. As a consequence of the Research Assessment Exercise in the United Kingdom in the 1990s, the highest rated department received four times the funding of the lowest rated department for the same volume of research (Morley, 2003, p. 22). As quality assurance has become more pervasive and influential, there have been reports of faculty resistance to quality assurance processes in some countries (Harley, 2002; Morley, 2003; Anderson, 2006; Hoecht, 2006). As of yet, however, there has been relatively little critical analysis of the foundations and impacts of quality assurance in higher education, including the way that these processes privilege some interests over others. Harvey and Newton note that there has been "a reluctance in higher education as a discipline to critique the ideology of quality evaluation" (Harvey and Newton, 2004, p. 156).

In this article, I explore one fundamental question about quality assurance in higher education (hereinafter referred to as QA): is QA a technical process or a political process? In formulating this question, I use the term "political" in the way that it is typically explained in political science textbooks. For example, according to Danziger, politics is "the exercise of power"; "the public allocation of things that are valued"; "the resolution of conflict"; and "the competition among individuals and groups pursuing their interests" (Danziger, 1994, p. 5). Thus, a process could be described as political if it displays some or all these characteristics of politics. A cursory familiarity with quality assurance processes suggests the presence of these characteristics.

Quality assurance agencies often have the formal or effective power to confer or deny the authority that is necessary for an academic programme to be offered or to be successful, and in the course of exercising that power they can practically dictate how the programme is to be designed. The decisions of quality assurance agencies may have a great influence on the allocation of resources among institutions and programmes, as in the example of the Research Assessment Exercise in the United Kingdom noted earlier. The element of Danziger's definition of politics that is most likely to escape notice

is the resolution of conflict. Frequently QA processes are designed in such a way as to deny the existence of, or to suppress, conflict rather than to provide a mechanism for its resolution.

I have deliberately worded the question that is the focus of this article in the most provocative way by casting it in "either-or" terms. In fact, rarely if ever is a QA process likely to be purely technical or purely political. Rather, QA processes can be viewed on a continuum in regard to the relative prominence of technical and political dimensions. The location on this continuum is likely to vary considerably depending upon the nature of the activity being assessed, the context of the assessment, and the way that the assessment process is structured. The assessment of activities for which there is a consensus on the specific learning outcomes and how to measure them may be regarded as a primarily technical process. On the other hand, the assessment of activities for which there are wide differences of opinion regarding what the goals should be and how to measure success is likely to have a substantial political dimension.

The notion that quality assurance in higher education is at least in part a political activity is not new. Harvey and Newton observed that quality evaluation is not "a neutral measuring process", but rather, it is "imbued with politics" (Harvey and Newton, 2004, p. 156). While a few other observers, such as Morley (2003), also have called attention to the political nature of QA, most of the literature about QA concentrates on the technical aspects of the process and by implication treats QA as if it were a purely technical process. This article is an attempt to add to the discussion of the political nature of QA by examining in some detail the factors that contribute to that political nature and by addressing the implications for the practice of QA.

Given the world-wide scale of QA, it is impossible to generalise across all the mechanisms and processes. In this article, I give particular attention to QA practices in Canada, and more particularly in Ontario where I have served on the board of one agency and have studied the operation of another. In addition to Canada, I cite examples and research findings from the United States, the United Kingdom, and Australia. I employ the methodology of critical analysis as Birnbaum (1983) defined that term. Drawing upon my own experience and the experience, observations and findings of others reported in the literature on QA, I have attempted to identify patterns and themes pertaining to the political nature of QA and its implications for practice.

The body of this article is divided into three sections, followed by a brief conclusion. The first section explains what is meant by the term quality assurance in higher education and briefly describes the arrangements for quality assurance in Canadian higher education. The second and longest section elaborates on what it means for QA processes to be political and discusses three factors that may contribute to the political nature of QA. The

third section begins by noting that the typical QA process looks as if it had been designed to solve a technical problem and neglects the political dimension of QA. Drawing upon the literature in the field of evaluation, the article then outlines the characteristics of a different approach to QA that is more appropriate to the political dimension of the project and could make quality assurance more effective in improving educational quality.

Quality assurance in Canadian higher education

Quality assurance is a term that refers to the monitoring, evaluation or review of higher education in order to establish stakeholder confidence that it fulfils expectations or meets minimum requirements (Martin and Stella, 2007, p. 34). The object of QA may be a programme or an entire institution. Internal QA occurs when an institution reviews itself or some of its own programmes. External QA is conducted by an external agency, usually either a government-related agency or an association of institutions, programmes, or relevant occupational groups. The focus of this article is external QA, and unless otherwise noted, the abbreviation QA will be used to refer to just this.

The most common forms of quality assurance are quality assessment, quality audit, and accreditation. In quality assessment, the external agency directs the process and is responsible for the evaluation of quality. In the quality audit approach, the external agency does not actually evaluate institutional or programme quality, but confines itself to evaluating the procedures and criteria that the institution or academic unit uses and the methods by which those who are responsible for a programme assure themselves of its quality. While the quality audit approach has the potential to give more autonomy to institutions in making judgments about quality than the quality assessment approach, much depends upon how prescriptive the audit requirements are. For example, if these include precise standards for educational processes, and specify the particular data that the institution must use in making its "own" quality assessments, then the external agency may exert considerable influence on the shape and outcomes of "internal" quality reviews.

Accreditation is a variant of quality assessment in which an external, non-governmental body determines whether an institution or programme meets pre-determined minimum quality standards. The main distinction between accreditation and quality assessment is that the former is concerned only with whether an institution or programme meets minimum standards, whereas the latter involves making "graded judgments about academic quality levels" (Dill *et al*., 1996, p. 21).

Perhaps the most noteworthy feature of QA in Canadian higher education is the absence of any national system of institutional quality assessment or accreditation. Quality assurance systems in Canadian higher education operate

mainly at the provincial or regional level and focus primarily on programmes rather than institutions. The main exceptions to these generalisations involve a relatively small number of national or North America-wide systems for accreditation of professional programmes or of highly specialised private post-secondary institutions such as bible colleges.

Three of Canada's ten provinces – Alberta, British Columbia and Ontario – have established their own quality assurance bodies and three other provinces – New Brunswick, Prince Edward Island and Nova Scotia – have jointly established an agency, the Maritime Provinces Higher Education Commission. The Commission, among other functions, serves the QA needs of the three partner provinces (Leyton-Brown, 2005). One of the responsibilities common to these agencies is to conduct quality assessments for all proposed new degree programmes of private post-secondary institutions and of institutions whose home base is outside the province in which the programme is to be offered. Provincial governments consider these quality assessments when deciding whether to allow the applicant institutions to offer the degree programmes in their province. If approval is granted, these agencies do periodic quality assessments of the programmes on a continuing basis. Although an institutional review is part of the application process, all approvals from the government are programme-specific.

Except in Ontario, another responsibility of the agencies just described is to conduct assessments of the academic content and quality of all new programmes proposed by the public universities, which comprise the vast majority of universities in Canada. In Ontario, the government-established QA agency, the Post-secondary Education Quality Assessment Board, has no role in QA for the public universities that have statutory authority to award degrees. These institutions have developed their own collective, self-regulatory system for quality assurance (Monahan, 2004). One component of this system, which Leyton-Brown describes as "the outstanding example of academic self-regulation at the programme level in Canada" is the appraisal process of the Ontario Council on Graduate Studies (Leyton-Brown, 2005, p. 235). This council conducts appraisals of all proposed new graduate programmes of the public universities, and the member universities have agreed that they will not offer a new graduate programme that does not get a satisfactory quality rating. Subsequently all graduate programmes are reappraised on a seven-year cycle.

For university undergraduate programmes in Ontario, there is an audit process in which an agency established by the public universities assesses the QA procedures of the individual institutions. Following a comprehensive review of the QA system for graduate programmes in Ontario universities, a 2007 report recommended a shift from the assessment approach to the audit approach and the creation of a unified framework for QA of graduate and undergraduate programmes (Van Loon, 2007). That recommendation has been

accepted by the parties concerned, and efforts are now under way to move to a consolidated quality audit system for Ontario's public universities. A quality audit system for Ontario's colleges was recently developed jointly by the association of colleges and the government ministry that regulates the colleges. Quality audit systems for ongoing programmes in provincial university sectors now exist as well in Quebec and in the provinces that are party to the Maritime Provinces Higher Education Council.

The political nature of quality assurance in higher education

There are three factors in particular that contribute to the political nature of QA in higher education. These are the differences of opinion among stakeholders as to what constitutes quality in higher education; the pressures toward conformity within academe; and imbalance of influence among different stakeholders in QA.

Differences of opinion as to what constitutes quality in higher education

There is no widely accepted definition of quality in higher education. Attempts to define quality have followed two main approaches, one dealing with philosophical concepts, the other focused on tangible phenomena thought to reflect quality. An example of the former is the categorisation of "conceptualisations" of quality offered by Harvey and Green (1993). Harvey and Green suggest that these conceptualisations can be grouped into five "discrete but interrelated ways of thinking about quality" (Harvey and Green, 1993, p. 11). These ways are described as exceptional, perfection or consistency, fitness for purpose, value for money and transformative. The alternative approach is typified by the classification of "conceptions" of quality suggested by Astin (1980). Astin offers five conceptions of quality: mystical, reputational, resources, outcomes and value-added. The latter four conceptions relate to phenomena that are in principle measureable, and in fact measures for all four exist. The mystical conception differs from the others in that it reflects the view that quality simply cannot be measured because the activities of higher education are too complex, variable and subtle. Whereas some might regard the four measureable groups of phenomena in Astin's typology as surrogates for quality, the fact that Astin presents them as "conceptions" of quality illustrates the lack of consensus not just on a specific definition of quality, but as to what constitutes a definition. At issue is whether it is helpful start the quest for a definition of quality in higher education with abstract philosophical principles or proceed directly to what Harvey and Green refer to as "criteria that reflect common-sense aspects of quality" (Harvey and Green, p. 29). While Harvey and Green express a preference for the former approach, they acknowledge that some prefer the latter.

Whether we regard the types of categories that Astin proposed as truly consisting of different conceptions of quality or as Harvey and Green would view them, sets of criteria that reflect common-sense aspects of quality, these are the kind of descriptors of quality that feature prominently in real world debates about quality. Harvey and Green note that different criteria for judging quality are seized upon by different interest groups on the basis of which best serve the interests of their group, and Astin makes a similar observation regarding his conceptions of quality. In like vein, Martin and Stella concluded that "the concept of quality is much disputed in higher education and often used by stakeholders to legitimise their specific vision or interests" (Martin and Stella, 2007, p. 30).

Outside the United States, the QA movement in higher education is a relatively recent phenomenon. Interest in QA developed earlier in the United States than in most other countries because of the historically limited governmental control over higher education in many states, and thus the ease with which new degree-granting institutions, particularly privately owned and controlled ones, could be established. The early emphasis in QA in the United States was on trying to prevent fly-by-night promoters of higher education from exploiting students. In countries such as Canada, where government approval was necessary for institutions to offer degree programmes, and governments funded those institutions, it was assumed that the institutions that conducted degree programmes were responsible ones with sufficient resources to offer credible programmes (Harris, 1976). If governments thought at all about the quality of higher education – and there is not much indication that this was of concern to governments in much of the world prior to the 1970s – they thought this job could be left to the universities themselves.

The big change in QA that occurred in the last third of the twentieth century in Canada and many other countries was from confining concern about quality to small, new, unknown and undercapitalised institutions to extending that concern to whole publicly funded systems of higher education that included well known, large and seemingly well-resourced institutions (Clark *et al.*, 2009). Several factors were responsible for the new interest that governments had in the quality of the central institutions in higher education systems rather than in just the peripheral elements of those systems.

The two most important factors contributing to the QA movement that began around the 1970s, though the precise date varies among jurisdictions, were economic. One related to the large and growing amounts of public expenditures that were going to higher education as a result of the enrolment expansion that got under way in the last third of the twentieth century. The other important factor, which along with demography provided the impetus for the enrolment expansion, was the growing belief that more and better higher education was essential for economic growth. If governments were

going to spend large sums of money on an enterprise that was deemed critical to national and regional economic growth and security, then it was natural for governments to take pains to ensure that the money was being spent effectively in service of its goals.

The idea that universities were increasingly going to be held accountable for their use of public funds was not lost upon higher education leaders (Corry, 1970). Educators in many jurisdictions realised that if they did not implement QA regimes on their own, governments would likely do it for them. Also, government-imposed QA regimes might not be as respectful of academic values as academics themselves could be if they were in charge. In particular, by establishing QA processes themselves, higher education leaders could **define quality in a way that best served their interests**. And what served their interests best was to define quality primarily in terms of resources, *i.e.* student-faculty ratio, average class size, expenditure per student, laboratory space, library holdings, etc. As Astin noted, "When educators feel the need to develop operational measures of institutional quality, they typically equate quality with an institution's educational resources" (Astin, 1980, p. 2). This tendency was evident in the annual briefs submitted to the government by the Council of Ontario Universities during the 1980s that purported to show the harmful effects of government funding policy on academic quality. In making this connection, quality was depicted exclusively in terms of resources (Skolnik, 1986).

As quality assurance in higher education developed and evolved, the ideas of quality that had the greatest influence and impact until practically the end of the twentieth century were those that defined quality in terms of resources, faculty research productivity, admissions selectivity and conformity to conventional educational practices and requirements. The explanation for the historic dominance of definitions of quality based upon these factors, especially resources, is that for quite some time academic leaders controlled QA, either because they were the ones to initiate it, or because even when governments promoted the establishment of QA systems, they deferred to academic leaders in the design of those systems.

One of the attractions of the resources conception of quality for educational leaders is that it makes quality a function of how much revenue the institution has. Thus, a QA process which embraces the resources conceptualisation of quality provides constant pressure on government to increase its financial support of public universities.

For the same reasons that higher education leaders like the resources view of quality, it is anathema to governments. Governments tend to prefer the learning outcomes conceptualisation of quality. Its appeal lies in the fact that it addresses the issue that is of paramount interest to parents and the

public: whether "students will receive a standard of education that provides both the technical knowledge required for practice and the general knowledge and skills that are essential for full participation in society" (Massaro, 2010). Educators tend not to like the learning outcomes view of quality, because they don't want to be held accountable for outcomes over which they feel that they have, at best, quite limited control.

In part, the story of the politics of QA over the past four decades in Canada and some other countries can be told in terms of the growing interest of the public and the government in the learning outcomes view of quality relative to the resources and other input and process-oriented views which had been dominant previously. In the seventies and eighties in some jurisdictions, higher education leaders were able to satisfy government concerns for accountability merely by establishing mechanisms of quality assurance. Governments initially didn't concern themselves with how quality was defined, only that some form of QA was in place. More recently though, they have taken an interest in how quality is defined and, in particular, the weight given to outcomes. In the 1990s governments in some Canadian provinces began to require post-secondary institutions to publicly report their performance in regard to some indicators that reflected outcomes, such as employment and graduation rates, and tied a portion of institutional funding to such indicators. In the United States this change in thinking was reflected in the recommendation by a 2006 Commission appointed by the Secretary of Education, that regional accreditation agencies should give learning outcomes the central place in accreditation reviews and give less attention to resources and educational processes (Commission Appointed by the Secretary of Education, 2006, p. 25). Not surprisingly, the recommendation to give less weight to what academics have traditionally regarded as the appropriate indicators of quality aroused a strident reaction from some accreditation agencies.

Besides the resources view, the other views of quality that traditionally have held sway in academe are those that emphasize admissions selectivity, educational processes and faculty research performance. As each of these views privileges some interests over others, debates about them can be viewed through a political lens of competition between different interests and ideologies. For example, those who favour the admissions selectivity conception of quality demonstrate an elitist view about admissions in contrast to those who place a high value on increasing accessibility and equity in higher education (Astin, 1985). Those who want to define quality in terms of conventional educational processes and requirements are in the camp of "traditionalists", as opposed to the "reformists" who urge the adoption of innovations that would make higher education more accessible and relevant (Trow, 1973).

The somewhat counter-intuitive view that the quality of the education that students are receiving can be judged by the research performance of the institutions and departments in which the students study serves to advance the research interests of those institutions and departments but not necessarily the education of their students. In Ontario the research-performance conception of educational quality has been championed by those who believe that all universities should emulate the research university paradigm, and opposed by those who favour a more heterogeneous university system in which some institutions concentrate on educating undergraduates (Clark *et al.*, 2009). Proponents of the research performance conception of educational quality maintain that strong research performance is a necessary condition for good teaching. However, the considerable research that has been done on the relationship between faculty research performance and teaching effectiveness has failed to reveal evidence of a positive correlation between these two realms of faculty activity (Hattie and Marsh, 1996; Marsh and Hattie, 2002; Halliwell, 2008). On the basis of their studies of this relationship, Hattie and Marsh concluded that "the common belief that research and teaching are inextricably entwined is an enduring myth" (Hattie and Marsh, 1996, p. 529). The rationale for the perpetuation of this myth in the face of contradictory evidence is that it serves to justify the research preoccupation "of existing interests and institutions" (Crimmel, 1984, p. 192).

A newer view of quality that has been the subject of increasing attention within the past decade focuses on the nature of the student experience. This view is based on the assumption that the more time that students engage in educationally useful activities, the better their learning and personal development (Kuh, 2003). Examples of such activities include writing papers, discussing assignments with instructors and asking questions in class. The National Survey of Student Engagement, in which students are asked about the frequency with which they engage in certain behaviours and about their perceptions of campus life, has been used in the United States and Canada to reflect this view of quality. The choice between the student engagement view and older views of quality may be the subject of some controversy if, as seems quite likely, the student engagement view tends to support interests other than the older views. For example, in some of the key benchmarks for student-faculty interaction and academic challenge, the larger, more research intensive universities in Canada tended to score the lowest, while smaller teaching-oriented institutions tended to score the highest (*Maclean's on Campus*, 2009).

The preceding paragraphs have illustrated some of the different views about what constitutes quality in higher education and how these views serve different interests and ideologies. Diversity in views about quality is not surprising because ideas of quality are both personal and social constructions that vary from

stakeholder to stakeholder (Ratcliff, 2003). How is this diversity of views as to what constitutes quality to be handled in a QA process? One way would be to invite input from all stakeholders in the design and implementation of the QA process and to strive for consensus among stakeholders. The more common way is for those who have the power to control the QA process to impose their views and values on the process. Either approach – conflict resolution or exercise of power – makes QA a political process.

QA as an instrument for promoting conformity in academe

The idea that there are strong pressures toward conformity in academe is not new. As the author noted in an earlier publication, "concerns about intellectual conformism and suppression of unorthodox thought and method are as old as organised intellectual activity itself and evoke such names as Bruno, Copernicus and Gallileo" (Skolnik, 1989, p. 620). Winchester observed that "intolerance for new and different ideas, although it has no place in the university, is often found there" (Winchester, 1986, p. 272). Kuhn maintained that science tends to operate largely within an accepted paradigm of knowledge and tends to be suspicious if not hostile to ideas that are outside that paradigm (Kuhn, 1962). Barber observed that a major theme in biographies and autobiographies of scientists is the frequency with which their new ideas were met with intolerance and often suppressed (Barber, 1962).

Prior to the development of QA, the pressure toward conformity within academe operated through a variety of mechanisms and forums, including scholarly journals, academic presses, research funding councils, departmental promotion and tenure committees, awards committees and so on. These pressures were manifested in a dispersed manner through multiple channels. QA processes constitute a more centralised, visible, and potent mechanism for promoting conformity in academe.

New or unorthodox paradigms, theories, research methods and pedagogical practices may be good or bad. However, it is difficult for most people to assess something that is new or unorthodox on its own terms rather than in relation to the older or orthodox. In fact, probably the greatest challenge in QA is to devise common standards for quality when assessing phenomena that are very different, for example a virtual university compared to a traditional university. The tendency is to equate quality with conformity to convention, and consequently to conclude that the more a programme diverges from conventional format or paradigm the poorer its quality. With its capability to stick the epithet of poor quality to its subjects, QA may be unprecedented in its power to discourage innovation in higher education.

A good example of how QA can penalise the unorthodox is provided by the studies that Harley and Lee have done on the Research Assessment Exercise (RAE) in the United Kingdom within the discipline of Economics (Harley and Lee,

1997; Lee and Harley, 1998; Harley, 2002; Lee, 2007). Lee and Harley portray the field of economics as consisting of different factions that compete for status, influence and resources. They refer to the factions, differentiated by ideology and methodology, as mainstream and non-mainstream, the latter consisting of institutional, radical, post-Keynesian, Sraffian and Marxian economics. Because of their dominant position, the mainstream economists were able to capture the process through which the panel of assessors for economics in the RAE was appointed. Lee and Harley provided data that show that panel members tended to publish primarily in mainstream journals and cite mainly mainstream journals, and they argued that the panel members were not qualified to judge the quality of work published in non-mainstream journals. They argued further that in their assessments, the panel members gave a higher weight to articles published in mainstream than in non-mainstream journals.

Lee and Harley concluded that the quality of mainstream economics departments tended to be rated higher than that of non-mainstream departments. This conclusion was not based on empirical evidence about quality, but because panel members approached their task with "the paradigm-bound view that the quality of non-mainstream research is largely inferior to mainstream research" (Lee and Harley, 1998, p. 198). The research by Lee and Harley shows how a QA process can be captured by the dominant camp within a discipline and enables the dominant camp to stifle competing camps more effectively than through the more dispersed processes of peer review of individual work, one item at a time, that Kuhn described.

Alternatively, in highly centralised QA processes there is a possibility of disciplinary bias when individuals from some disciplines have the ultimate power to make decisions about the quality of programmes in other disciplines. In an earlier publication, the author analysed the case of the graduate programmes appraisals process in Ontario in which a single group of academics make quality judgments for all programmes (Skolnik, 1989). It was concluded that the process reflected primarily the paradigm and conventions of the natural sciences, and that it discriminated against some other disciplines.

Over the period studied, the range in the percentage of programmes receiving less than the highest quality rating was from zero in the natural sciences to 30% in the humanities. Within the humanities, fields that had least in common with the natural sciences did the worst: half or more of the programmes in classics, religion and education received less than the highest rating. It was noted also that "the extraordinary success of the natural science programmes is remarkable in view of the widespread concerns of the university community during this period about inadequate funding and the expectation that this could result in serious quality problems for those areas which require expensive equipment and research support" (Skolnik, 1989, p. 632). Unless faculty in the humanities were less competent than their peers in the sciences, the most

plausible explanation for this finding was that what were purported to be measures of quality were in reality measures of the extent to which programmes conformed to the model of graduate programmes that was favoured in the sciences. This conclusion was consistent with Winchester's observation that at least until that time "research" had come to be understood to mean "doing things rather like one would if the subject matter was physics or chemistry" (Winchester, 1986, p. 282). A decade later, a committee that reviewed the graduate programmes appraisals process in Ontario expressed concern about the centralisation of quality judgments across disciplines, and it recommended a realignment of the respective influence of the central appraisals body and its subject matter consultants (OCGS Review Committee, 1999).

The role of faculty in QA

The role and power of different stakeholders in QA varies considerably. Often, the ordinary faculty member has a very circumscribed role and no power. Typically in QA processes in Canada, the role of faculty is confined to the onerous task of providing data to the assessment body, and then awaiting a judgment that may have far-reaching and significant effects on his or her work and career. Faculty have little opportunity for input into the way that the process is designed, its purpose, or the criteria that are used in the assessment; nor is there much opportunity for appeal of the results. Apparently, these conditions apply to some extent in some other countries as well. For example, Hoecht noted that the Quality Assurance Agency in the United Kingdom "does not appear to be involved in a consultative open debate in its policy-making and policy implementation" (Hoecht, 2006, p. 546). The marginalisation of the faculty role in QA stands in contrast to the norms in many countries regarding the role of faculty in academic governance. Yet, today QA is arguably among the most important of higher education governance functions.

In view of the marginalisation of faculty that occurs in some QA processes, it is perhaps not surprising that some of the relatively few surveys and interview-based studies that have been done of faculty perceptions of QA show that substantial proportions of faculty have negative views about their quality assurance processes (Harley, 2002; Morley, 2003; Anderson, 2006; Hoecht, 2006). According to Anderson, the negative views of Australian academics first identified in the early 1990s still persisted (Anderson, 2006). Harley reported that the majority of academics in her study were "outrightly hostile" to the Research Assessment Exercise in the United Kingdom, either because they felt themselves personally disadvantaged, or because of "distortions that it was felt to have introduced into academic life" (Harley, 2002, p. 203). Hoecht found that QA was perceived by faculty as a form of control and an encroachment on their professional autonomy, and addressed quality only at a superficial level (Hoecht, 2006). In view of comments made earlier on possible implications of a

shift from the assessment to the audit approach in QA, it is noteworthy that some of the research on faculty perceptions of QA in the United Kingdom summarised here was done during the period when the audit approach was employed, and in those studies many faculty reported that they felt that the audits were a form of external control.

Common themes in these surveys of faculty were: feelings of marginalisation and powerlessness; that the processes were intended to serve managerial ends rather than to improve quality; and that some interests were being favoured over others. Morley noted that, paradoxically, the QA movement was perceived to be undermining quality. She observed that "the compliance culture and command economy in higher education threatens to produce self-policing, ventriloquising apparatchiks, as opportunity for cultural agency is reduced" (Morley, 2003, p. 162). Consistent with these perceptions of faculty, Harvey and Newton concluded that "the rhetoric and documentary preambles in many countries refer to quality evaluation as a process of improvement, yet all the emphases are on accountability, compliance and, in some cases, control…" (Harvey and Newton, 2004, p. 151).

The marginalisation and disempowerment of faculty that often is part of the QA process should be cause for concern if, as some observers maintain, faculty commitment to QA is essential to quality higher education (Dill *et al.*, 1996). Newton may well be right when he says that "… genuine quality enhancement can only be fully sustained if it is premised on the energies and initiatives of frontline academics" (Newton, 2000, p. 153). But if he is right, how does one reconcile this imperative with the perception that QA is oriented toward "correction and rehabilitation" of faculty (Morley, 2003, p. 50)?

An alternative evaluation paradigm

The arguments in this article support the conclusion of Harvey and Newton cited earlier that quality assessment is not "a neutral measuring process", but is a process that is "imbued with politics" (Harvey and Newton, 2004, p. 156). Thus, rather than following a method that denies the inherent political nature of quality assurance, as is frequently the case in QA in higher education, would it not be more appropriate to employ a method that recognises this reality? There is an extensive scholarly and professional literature on evaluation that offers some useful guidance in this respect. However, the practice of QA has developed largely in isolation from that body of literature, and in fact, there is little cross referencing between the two. When viewed from the perspective of the literature on evaluation, the typical method employed for QA looks rather outmoded and deficient (Skolnik, 2000). Morley noted that "the methodology of quality assessment is perceived [by the subjects in her study] as taking the academy backwards in terms of sophistication of analysis" (Morley, 2003, p. 160).

A helpful perspective for considering possible reforms of QA is provided by Guba and Lincoln's depiction of four generations through which evaluation has evolved since its origin in the early twentieth century in elementary and secondary schooling (Guba and Lincoln, 1989). The first three of these generations were: evaluation as measurement; evaluation as description; and evaluation as judgment. While there were unique issues in each of these three orientations, a problem common to all was the failure to recognise the competing values, perceptions and interests of different stakeholders. Another characteristic was managerialism, the tendency to view evaluation from the perspective of the manager. Guba and Lincoln note that a consequence of this orientation was that the manager was placed outside the evaluation, whereas management might be a significant part of the problem. Also, the managerial orientation tended to disempower other stakeholders.

Guba and Lincoln depict fourth generation evaluation as attempting to remedy the problems of the earlier generation models. The fourth generation model embraces value-pluralism and takes as its starting point the claims, concerns and issues put forth by **all** stakeholders. It employs a constructivist as opposed to a positivist paradigm. It changes the role of the evaluator from that of technical expert or judge to that of mediator and negotiator. It makes evaluation a **collaborative** rather than a top-down process. Some of these characteristics of fourth generation evaluation smack of postmodernism, and models of evaluation with these characteristics have sometimes been referred to as postmodern evaluation (Mabry, 2002, p. 14).

Guba and Lincoln's fourth generation model of evaluation has many similarities to the "responsive" model of evaluation developed by Robert Stake (Stake, 1980). The essential feature of this approach to evaluation is responsiveness to key issues and problems identified by stakeholders during the evaluation process, rather than responsiveness to programme theory or stated goals (Stake, 2004, p. 89). Stake describes responsive evaluation as "a general perspective in the search for quality", and as "an attitude more than a model or recipe" (Stake, 2004, p. 86). While Guba and Lincoln portray their fourth generation evaluation as a naturalistic alternative to criterion-based evaluation, Stake seeks to combine responsive and standards-based evaluation. However, Stake acknowledges that interpretive thinking – which is the hallmark of responsive evaluation – does not blend easily with criterial thinking.

Looking at quality assessment in higher education through the lens of Guba and Lincoln's four generations of evaluation, it appears that as QA developed in the 1970s and 1980s, the judgment model predominated. The judges were drawn primarily from an academically conservative elite and imposed their own fairly narrow view of academic quality that emphasised resources, admissions selectivity, educational process criteria and research reputation. As the accountability movement gained steam in the 1990s,

emphasis shifted to the measurement model. Within the past decade the interest in institutional performance indicators has shifted to the measurement of student learning. Ironically, that shift takes us back full circle to the origins of the field of evaluation in the measurement of student learning, then in the elementary and secondary schools, now in universities and colleges.

Although it has been identified as a model that could be used for QA (Conrad and Wilson, 1985), the responsive model of evaluation has not been used much in this field. The continued persistence of the earlier models of evaluation in QA, over the responsive model, is curious given the growth of constructivist and postmodern influences in academe. The dearth of use of the responsive model suggests that the individuals who exercise the greatest power within and over higher education are largely not constructivists or postmodernists. They are more likely to have a neo-liberal orientation (Olssen and Peters, 2005). The lack of interest in the responsive model may also reflect the strength of the natural sciences, a branch of academe where constructivism has made little impact, in higher education policy making.

Conclusion

Insofar as QA is to a large extent a political process, it is healthier to recognise the political dimension of QA than to try to sweep it under the carpet. Likewise, it is important to be clear about the purpose of QA, specifically whether it is intended to serve merely as a ritual in which to "feed the [accountability] beast" (Newton, 2000, p. 155), to enforce a particular ideology or set of policies, or to foster improvement in education. If there is a genuine desire to recognise the diverse views regarding quality and to strive for educational improvement, then the quality assessment process should be designed in a way that will further these ends. The literature on evaluation can be helpful in finding such a design. In particular, serious attention might be given to the responsive model of evaluation, which provides a framework for involving faculty and other stakeholders in QA in a meaningful, collaborative way.

It is worth stressing that inherent in the responsive model of evaluation is the idea that the input and influence of ***all*** stakeholders should be sought and facilitated. While it is common for faculty to be excluded from a significant role in quality assurance processes, it would be a mistake to swing to the opposite situation wherein faculty control QA, and other stakeholders, *e.g.* students, employers and the public, are excluded. Some faculty may try to limit the operational definition of quality to resources and other traditional conceptions of quality, whereas there is much to be said for the relevance and value of newer conceptions like learning outcomes and student engagement. The appropriate way to view quality may vary considerably from one context to another. The healthiest way to deal with the political nature of QA is to

accept the diversity of views and interests of all stakeholders and work toward the reconciliation of those diverse views and interests. Perhaps foremost among the standards for academic audits should be the inclusiveness and transparency of the QA process.

The author:

Michael L. Skolnik
Professor Emeritus
Department of Theory and Policy Studies
Ontario Institute for Studies in Education of the University of Toronto (OISE/UT)
252 Bloor Street West, 6th floor
Toronto, Ontario M5S 1V6
Canada
E-mail: mskolnik@oise.utoronto.ca

No specific funding source was called upon to support the research presented.

References

Anderson, G. (2006), "Assuring Quality/Resisting Quality Assurance: Academics' Responses to Quality in some Australian Universities", *Quality in Higher Education*, Vol. 12, No. 2, pp. 161-173.

Astin, A. (1980), "When Does a College Deserve to be called 'High Quality'?", *Current Issues in Higher Education*, Vol. 2, No. 1, pp. 1-9.

Astin, A.W. (1985), *Achieving Educational Excellence*, Jossey-Bass Publishers, San Francisco.

Barber, B. (1962), "Resistance by Scientists to Scientific Discovery", in B. Barber and W. Hirsch (eds.), *The Sociology of Science*, Free Press, pp. 539-556, New York.

Birnbaum (1983), *Maintaining Diversity in Higher Education*, Jossey-Bass Publishers, San Francisco.

Clark, I.D. *et al.* (2009), *Academic Transformation: The Forces Reshaping Higher Education in Ontario*, Queen's University School of Policy Studies and McGill-Queen's University Press, Montreal.

Commission Appointed by Secretary of Education Margaret Spellings (2006), *A Test of Leadership: Charting the Future of US Higher Education*, US Department of Education, Washington DC.

Conrad, C.F. and R.F. Wilson (1985), *Academic Programme Reviews: Institutional Approaches, Expectations, and Controversies*, ASHE-ERIC Higher Education Research Report No. 5, Association for the Study of Higher Education, Washington, DC.

Corry, J.A. (1970), *Farewell the Ivory Tower*, McGill-Queen's University Press, Montreal.

Crimmel, H.H. (1984), "The Myth of the Teacher-Scholar", *Liberal Education*, Vol. 70, No. 3, pp. 183-198.

Danziger, J.N. (1994), *Understanding the Political World*, second edition, Longman, New York.

Dill, D.D. *et al*. (1996), "Accreditation and Academic Quality Assurance", *Change*, Vol. 17, No. 24, September/October.

Guba, E.G. and Y.S. Lincoln (1989), *Fourth Generation Evaluation*, Sage Publications, Newbury Park, CA.

Halliwell, J. (2008), *The Nexus of Teaching and Research: Evidence and Insights from the Literature*, Higher Education Quality Council of Ontario, Toronto, ON.

Harley, S. (2002), "The Impact of Research Selectivity on Academic Work and Identity in UK Universities", *Studies in Higher Education*, Vol. 27, No. 2, pp. 187-205.

Harley, S. and F.S. Lee (1997), "Research Selectivity, Managerialism, and the Academic Labor Process: The Future of Nonmainstream Economics in UK Universities", *Human Relations*, Vol. 50, No. 11, pp. 1427-1460.

Harris, R.S. (1976), *A History of Higher Education in Canada,* University of Toronto Press, Toronto, ON.

Harvey, L. and D. Green (1993), "Defining quality", *Assessment and Evaluation in Higher Education*, Vol. 18, No. 1, pp. 9-34.

Harvey, L. and J. Newton (2004), "Transforming Quality Evaluation", *Quality in Higher Education*, Vol. 10, No. 2, pp. 149-165.

Hattie, J. and H.W. Marsh (1996), "The Relationship between Research and Reaching: A Meta-Analysis", *Review of Educational Research*, Vol. 66, No. 4, pp. 507-542.

Hoecht, A. (2006), "Quality Assurance in UK Higher Education: Issues of Trust, Control, Professional Autonomy, and Accountability", *Higher Education*, Vol. 51, No. 4, pp. 541-563.

International Network of Quality Assurance Agencies in Higher Education (2010), "About INQAAHE", *www.inqaahe.org/main/about-inqaahe*, accessed 7 January 2010.

Jenkins, A. (2007), *Towards Sane UK National Policies: or Learning from some US policies*, Presentation at the Marwell 2007 colloquium on International Policies and Practices for Academic Enquiry, *http://portal-live.solent.ac.uk/university/rtconference/2007/resources/alan_jenkins.pdf*.

Kuh, G.D. (2003), "The National Survey of Student Engagement: Conceptual Framework and Overview of Psychometric Properties", *http://nsse.iub.edu/pdf/conceptual_framework_2003.pdf*, accessed 16 February 2009.

Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Lee, F.S. (2007), "The Research Assessment Exercise, the State and the Dominance of Mainstream Economics in British Universities", *Cambridge Journal of Economics*, Vol. 31, No. 2, pp. 309- 325.

Lee, F.S. and S. Harley (1998), "Economics Divided: The Limitations of peer review in a Paradigm-Bound Social Science", in D. Jary and M. Parker (eds.), *The New Higher Education: Issues and Directions for the Post-Dearing University*, Staffordshire University Press, Stoke-on-Trent.

Leyton-Brown, D. (2005), "Demystifying Quality Assurance", in C.M. Beach, R.W. Boadway and R.M. McInnis (eds.), *Higher Education in Canada*, McGill-Queen's University Press, pp. 229-248, Montreal.

Mabry, L. (2002), "Postmodern Evaluation – or Not?", *American Journal of Evaluation*, Vol. 23, No. 2, pp. 141-157.

Maclean's on Campus (2009), "University Students Grade Their Schools", by *T. Keller and S. Farran*, *http://oncampus.macleans.ca/education*, accessed 16 February 2009.

Marsh, H.W. and J. Hattie (2002), "The Relation Between Research Productivity and Teaching Effectiveness: Complementary, Antagonistic, or Independent Constructs?", *The Journal of Higher Education*, Vol. 73, No. 5, pp. 603-641.

Martin, M. and A. Stella (2007), *External Quality Assurance in Higher Education: Making Choices*, UNESCO, International Institute for Educational Planning, Paris.

Massaro, V. (2010), "Cui bono? The Relevance and Impact of Quality Assurance", *Journal of Higher Education Policy and Management*, Vol. 32, No. 1, pp. 17-26.

Monahan, E.J. (2004), *Collective Autonomy: A History of the Council of Ontario Universities, 1962-2000*, Wilfrid Laurier University Press, Waterloo, ON.

Morley, L. (2003), *Quality and Power in Higher Education*, Society for Research into Higher Education/Open University Press, Buckingham, UK.

Newton, J. (2000), "Feeding the Beast or Improving Quality? Academics Perceptions on Quality Assurance and Quality Monitoring", *Quality in Higher Education*, Vol. 6, No. 2, pp. 153-163.

OCGS Review Committee (1999), *Review of the Ontario Council on Graduate Studies*, Council of Ontario Universities, Toronto.

Olssen, M. and M.A. Peters (2005), "Neoliberalism, Higher Education, and the Knowledge Economy: From the Free Market to Knowledge Capitalism", *Journal of Education Policy*, Vol. 20, No. 3, pp. 313-345.

Ratcliff, J.L. (2003), "Dynamic and Communicative Aspects of Quality Assurance", *Quality in Higher Education*, Vol. 9, No. 2, pp. 117-131.

Skolnik, M.L. (1986), "If the Cut is So Deep, Where is the Blood? Problems in Research on the Impact of Financial Restraint", *Review of Higher Education*, Vol. 9, No. 4, pp. 435-455.

Skolnik, M.L. (1989), "How Academic Programme Review Can Foster Intellectual Conformity and Stifle Diversity of Thought and Method in the University", *The Journal of Higher Education*, Vol. 60, No. 6, pp. 619-643.

Skolnik, M.L. (2000), "Does Counting Publications Provide any Useful Information About Academic Performance?", *Teacher Education Quarterly*, Vol. 27, No. 2, pp. 15-26.

Stake, R.E. (1980), "Programme Evaluation, Particularly Responsive Evaluation", in W.B. Dockrell and D. Hamilton (eds.), *Rethinking Educational Research*, Hodder and Stoughton, London.

Stake, R.E. (2004), *Standards-Based and Responsive Evaluation*, Sage Publications, Thousand Oaks, CA.

Trow, M. (1973), *Problems in the Transition from Elite to Mass Higher Education*, Carnegie Commission on Higher Education, reprint, Berkeley.

Van Loon, R. (2007), *The Ontario Graduate Programme Appraisal Review*, Council of Ontario Universities, Toronto.

Winchester, I. (1986), "The Future of a Mediaeval Institution: The university in the twenty-first century", in W.A. Nielson and C. Gaffield (eds.), *Universities in crisis: A Mediaeval Institution in the Twenty-First Century*, Institute for Research on Public Policy, Montreal.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

System accreditation: an innovative approach to assure and develop the quality of study programmes in Germany

by
Tanja Grendel and Christoph Rosenbusch
Centre for Quality Assurance and Development,
Johannes Gutenberg University, Germany

"System accreditation" is a new approach developed for German universities to conduct the mandatory accreditation of all their study programmes. A pilot project at Johannes Gutenberg University in Mainz is playing an important role in paving the way for this alternative to prevailing programme accreditation. This article describes how system accreditation, an innovative approach towards organisational adaption to national regulations, was conceived and how it functions. Based on the experience of Johannes Gutenberg University, the article explores the potential of system accreditation to improve quality assurance and the development of study programmes. System accreditation faces three global challenges: that of creating an integrated approach, establishing a solid evidence base and fostering the effectiveness of evaluation efforts.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

L'agrément combiné : une approche innovante de l'assurance qualité et du développement des programmes d'études

par
Tanja Grendel et Christoph Rosenbusch
Centre pour le développement de l'assurance qualité,
Université Johannes Gutenberg, Allemagne

L'agrément combiné est un nouveau schéma de conduite de l'agrément obligatoire de tous les programmes d'études développés par les universités allemandes. Un projet pilote, mené à l'Université Johannes Gutenberg de Mayence, joue un rôle d'éclaireur important dans cette manière alternative d'obtenir l'agrément des programmes. Cet article décrit la conception et le fonctionnement de l'agrément combiné, une approche innovante d'adaptation aux règlementations nationales. Basé sur le cas de l'Université Johannes Gutenberg, il explore le potentiel de l'agrément combiné dans l'amélioration de l'assurance qualité et le développement de programmes d'études en relation avec trois défis principaux : créer une approche intégrée, établir une solide base de données probantes et promouvoir l'efficacité des efforts d'évaluation.

The accreditation of study programmes in Germany

Over recent decades, regulations governing higher education in Germany, as in many other countries, have put a growing emphasis on accountability and quality issues (for an early account see Neave, 1988). In parallel to this development, higher education institutions (HEIs) are expected to take on greater autonomy while the state contents itself with steering from a distance (for more details on the German situation, see Kehm and Lanzendorf, 2006; Kehm, 2007a). The quality assurance of study programmes became a legal obligation for German universities in 1998.[1] One of the main outcomes of this development is that nowadays all study programmes in higher education need to be accredited externally. The accreditation process is meant to secure a minimum level of quality standards in the area of teaching and learning. Certification for the respective study programmes is granted for a duration of five years, after which the programmes need to be re-examined and accredited for a further five years.

The accreditation process follows certain formal and subject-specific rules, but in general three requirements need to be met. These consist of a sound study programme concept with clear qualification goals, instruments and mechanisms of quality assurance; the study programmes need to be feasible within the allotted amount of time; and programme requirements for students must be transparent.[2]

The mandatory certification/accreditation of study programmes is conducted under the aegis of the Accreditation Council. This decision-making body consists of representatives from HEIs and federal states' authorities, industry, the student body and international experts. The Accreditation Council sets the standards for the accreditation process and authorises accreditation agencies to carry out programme accreditations. There are currently six such agencies in operation in Germany.

Over recent years, more and more voices have been critical of the process of programme accreditation (see, for example, Kehm, 2007b, p. 88 *et sqq*.; Schmidt and Horstmeyer, 2008, p. 42 *et sqq*.). The main areas of concern are:

- There is a lack of coherence between programme accreditation and other initiatives of quality assurance and development in universities.
- The structural and conceptual development of the university as a whole is neglected.

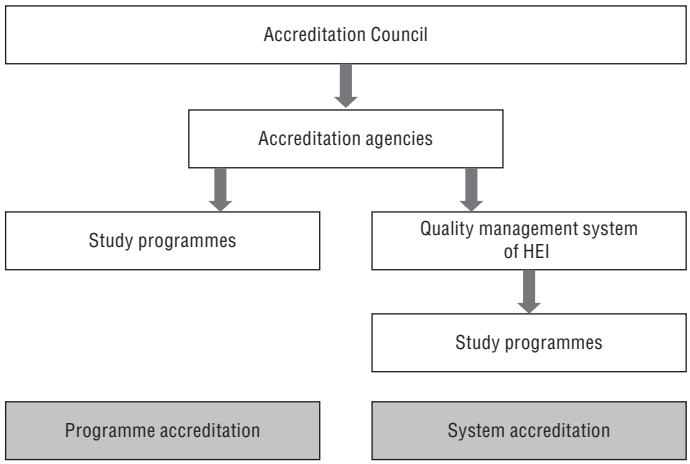
- Programme accreditation entails high costs for universities.
- The accreditation agencies have difficulty in coping with the massive numbers of new bachelor and master programmes.[3]
- Criticism is being formulated by departments about the lack of consistency in the agencies' decisions.

The first two concerns constitute the main conceptual objections to the prevailing programme accreditation. Although one of the criteria for the assessment of study programmes is the existence of instruments and procedures of quality assurance, in reality the results are not integrated systematically into the accreditation process. Furthermore, the exclusive focus of programme accreditation on single study programmes means that related structural and strategic developments in the university are bypassed. This is a significant shortcoming given that universities increasingly attempt to establish themselves as "organisational" or "strategic" actors (Krücken and Meier, 2006; Whitley, 2008).

System accreditation as an alternative approach

It is in this context that, since 2004/05, HEIs have been looking for alternative approaches to accredit their study programmes. One of the alternatives that is gathering more and more support is system accreditation. The main difference between this approach and the prevalent mode of programme accreditation is that the quality management of the whole university is externally accredited, rather than just the individual study programmes. Thus, responsibility for the accreditation of all study programmes lies with the Quality Assurance and Development Unit of the HEI (Figure 1).

Figure 1. **General schema of programme and system accreditation**



According to its proponents, the main advantages of system accreditation are:

- University autonomy (especially with regard to developing a competitive profile) is enhanced.
- Related strategic aspects (research quality, organisational development, general development plans of the university) are included into the quality assurance process.
- Quality assurance and development instruments are integrated into a comprehensive quality management system for study programmes.
- Administration is streamlined, entailing lower costs.[4]

A prerequisite for the realisation of system accreditation is a comprehensive quality management system. This system must comply with the European Standards and Guidelines for Quality Assurance in Higher Education (ESG), the specifications of the Kultusministerkonferenz (KMK)[5] and the Accreditation Council's standards.[6] Moreover, institutional accreditation of a quality management system has to be renewed after a period of seven years.[7]

At the present time, only very few German HEIs have the structure and experience necessary to apply for accreditation of their quality management systems. One of these is Johannes Gutenberg University Mainz (JGUM), where system accreditation was co-developed and implemented as a pilot project.

Working conditions and quality management approach at JGUM

JGUM is one of the largest universities in Germany. It is currently conducting research into a wide range of academic fields and offers study programmes for nearly 35 000 students.[8] The university employs almost 500 professors, as well as some 2 700 research associates spread among 11 departments in a total of 150 institutes. Nearly 1 800 employees are non-academics. JGUM offers over 150 study courses;[9] these cover a wide spectrum of university subjects and include music, fine arts, sports and science. More than 30 collaborative research centres work in various fields of research activity.

The quality management of scientific work at JGUM is guided by the university's strategic vision and consists of different monitoring and consulting initiatives. The department of university statistics collects general data (mainly on capacities and finance), while many departments of the central administration (*e.g.* the central Bologna Commissioner or the Department of Teaching and Learning) offer specialised consultancies for the scientific units.[10] A central function in quality management at JGUM is under the responsibility of the Centre for Quality Assurance and Development (ZQ). ZQ's main task is to gather differentiated evidence in the field of research and teaching quality as well as organisational development in JGUM's departments and institutes.[11] On

this basis, ZQ offers consultancies to the scientific units as well as to the central university management.[12]

ZQ's role and work can best be characterised by four main features:

i) To date, ZQ has conducted a huge number of internal and external evaluations at JGUM in the areas of teaching and learning, research and organisation.[13] Over the years, ZQ's knowledge about JGUM's scientific institutions has grown, as have personal contacts. As a result, a climate of mutual trust between ZQ and the scientific units has developed. This is a considerable asset for the establishment of a sustainable and effective quality management system. As related research consistently shows, if evaluation is to be successfully designed and implemented and evaluation results properly utilised, a good communication base is essential (see, for example, Owen and Rogers, 1999, p. 105 *et sqq*.).

ii) ZQ's structural position guarantees it significant independence as well as providing a good basis for co-operation with the central management of the university and the scientific units. ZQ is not subordinated to the university president, but governed by a special committee of the academic senate. This committee oversees general matters relating to quality management at JGUM, such as the extent to which course evaluation surveys are mandatory. Furthermore, it fulfils the function of an ombudsman for eventual complaints from the scientific units. At the same time, ZQ is solely responsible for data collection as well as for the development measures it suggests to the scientific units. This status is essential if qualitative aspects are to be valued and placed above and beyond the political logic and constraints of day-to-day decision making in universities.

iii) ZQ services the departments and institutes and the university management. At the same time, it makes efforts to stay abreast of state-of-the-art scientific knowledge in the field of quality management in higher education and to develop its own instruments accordingly. This is made possible mainly by attracting third party-funded research projects and because of the multi-professional skill mix of its staff.

iv) ZQ considers that quality management has to be founded on a system, or model, of quality and organisational development. The ZQ quality model is based on four distinct quality dimensions. The customary three quality dimensions of input/structure quality, process and outcome quality, identified by Donabedian (1980), are supplemented by the dimension of goal quality. As regards the promotion of quality teaching, ZQ considers that the research potential and teaching quality of a scientific unit should be developed in concert. Different quality aspects are assessed with a considerable range of instruments and indicators. Departments are supported by course evaluation surveys, *alumni* surveys, workload studies,

surveys of first-year students and longitudinal analyses of students' study courses. Standardised procedures have to be adjusted occasionally to special working conditions and information needs in different scientific areas. Furthermore, the quantitative data is selectively supported by qualitative analysis gained in internal and external evaluations.

The certification of study programmes within the institutional framework of system accreditation

At JGUM individual departments and institutes are free to decide whether they want their study programmes to be accredited through ZQ or an external agency. The attractiveness of system accreditation for the departments and institutes lies mainly in the quality of the monitoring instruments and results as well as in ZQ's competence and experience in the field of organisational and quality development in scientific organisations. If required to, ZQ can help the scientific units to build up organisational structures which promote continuous quality development and to develop appropriate measures in function of evaluation results.

If a study programme is to be accredited for the first time, the institutes or departments concerned need to develop an initial concept for it, in co-operation with the central Bologna Commissioner. This concept has to be approved by the faculty council, whereupon it is passed on to the central university management. Next, and in function of existing faculty development plans in the field of teaching and research, and reports on past evaluations (both internal and external), it is decided whether the concept should be worked out further along the lines proposed by the institute or department.[14] New concepts are checked by the department of teaching and learning with regard to formal aspects (*e.g.* the parameters of the Accreditation Council) and by ZQ with regard to learning and teaching quality. The assessment of scientific quality includes the involvement of external reviewers, namely scientists from the field as well as representatives from the student body and the industrial sector.

Subsequently, ZQ transmits the information provided by the institutes or departments as well as a written statement (usually containing recommendations) to the senate committee for teaching and learning. The Committee subsequently presents the proposal for the new study programme to the academic senate for approval.

In the case of reaccreditation, the process is fairly similar.[15] The main difference is that, during the course of system accreditation, continuous quality monitoring is gradually established. This permits a systematic analysis of the quality of study programmes on different levels. For example, bachelor programmes are accompanied by surveys of first-year students and *alumni*. Furthermore, subject-wide course evaluation surveys are conducted

over a cycle of three semesters. If necessary, workload studies or qualitative group discussions with representatives from different status groups allow a close examination of certain issues and help to complete the picture.

ZQ's role in the process is to provide survey methods, to conduct the surveys and to edit the results on behalf of the scientific units. ZQ can also support the scientific units when it comes to interpreting the data collected and defining adequate measures of quality development. This support is often essential due to the fact that hasty conclusions from monitoring data constitute one of the basic impediments to successful quality development (see, for example, Kis, 2005).

Progress towards integrated, evidence-based and effective quality development?

System accreditation integrates external demands for a minimum of quality standards for study programmes as well as the university's internal quest for quality development (see Schmidt, 2009, p. 167). Potentially, both needs can benefit substantially from it.

On the one hand, the externally requested accreditation process has a more solid evidence base as the data from internal evaluation procedures is systematically taken into account. The comprehensive approach, covering the four quality dimensions of study programmes over a period of five years, leads to a more sophisticated understanding of relevant quality aspects. Thanks to the broad range of indicators and analytical methods provided by ZQ, different dimensions of teaching quality can be analysed. It is possible to evaluate diverse aspects such as: study organisation and communication/co-operation among scientists and between scientists and students (process quality); the scientific units' didactic competencies and equipment (input or structure quality); completion rates, competence gains and students' first jobs (outcome quality); and the content and consistency of the study programme (goal quality).

On the other hand, system accreditation adds a substantially higher level of commitment to HEIs' internal quality management efforts because accreditation of study programmes is a necessity for them. This commitment can be used to systematically engage people at the collective level into continuous and evidence-based quality improvement of their study programmes. This is a very important feature for universities as "professional bureaucracies" (Mintzberg, 1992) and "loosely coupled systems" (Weick, 1976). The status of universities as professional bureaucracies and loosely coupled systems, which essentially underlines the substantial autonomy of scientific units and individual scientists alike, is undoubtedly appropriate for working in the field of science and education. At the same time, these organisational characteristics cause recurrent problems with regard to collective organisational and strategic action. System

accreditation offers the potential to foster collective, goal-directed efforts in the field of quality development; this is crucial to the establishment of effective quality development above the individual level.

As mentioned before, ZQ consults with the scientific units when interpreting data and when formulating appropriate measures of quality development. Measures for improving quality in the area of teaching and learning can be conducted on different levels: at the level of the individual teacher, the study programme or the department. It is interesting to note that for the individual teacher, learning through advanced vocational training in didactics is also proposed and co-ordinated by ZQ. Measures on a collective level include re-designing study programmes, single courses or modules, as well as organisational development. The level of study programmes in particular is becoming increasingly important for the development of teaching quality, as new bachelor and master programmes tend to expand over traditional borders of scientific disciplines and institutes.

Continuous monitoring, which is a standard feature of system accreditation, helps to put teaching quality in the context of a quality (PDCA: plan, do, check, act) circle. Based on the initial concept (plan), study programmes are implemented (do); their quality is systematically evaluated (check) and based on the evidence gathered; adequate measures of development and refinement are initiated (act). System accreditation definitely gives a better picture of the quality of study programmes. However, its main benefit may be that it strengthens the delicate nexus between evaluation and governance, creating the environment for effective quality management. This step (from check to act) can be seen as one of the most crucial and challenging problems for systematic, evidence-based and effective quality development in higher education. (For more details see, for example, Grendel, Schmidt and Springer, 2006.)

Conclusion

The argument above demonstrates that system accreditation holds significant potential for integrated, evidence-based and effective quality assurance and development of study programmes. This new approach to accreditation integrates external demands for quality assurance and the university's internal quest for quality development and thus helps to overcome one of the main conceptual shortcomings of prevalent programme accreditation. Furthermore, at JGUM, an integrated approach to the quality management of study programmes is enhanced by a comprehensive quality model. This quality model, along with a broad range of instruments which continuously assess relevant quality indicators, significantly strengthens the evidence base for accreditation and quality management alike. Finally, the

effectiveness of quality management efforts is enhanced by the scientific units' greater commitment. In this context, consultancy services offered by a central quality agency (in this case ZQ) can help the scientific units to establish the necessary structures and competencies to effectively take advantage of the knowledge generated during the course of system accreditation.

The authors:

Dipl.-Soz. Tanja Grendel
Zentrum für Qualitätssicherung und -entwicklung
Johannes Gutenberg-Universität Mainz
Forum 4
55099 Mainz
E-mail: *Tanja.Grendel@zq.uni-mainz.de*

Christoph Rosenbusch, M.A.
Zentrum für Qualitätssicherung und -entwicklung
Johannes Gutenberg-University Mainz
Forum 4
55099 Mainz
E-mail: *Christoph.Rosenbusch@zq.uni-mainz.de*

This article is based on a long-term project funded by different sources. The principal sources of funding are Johannes Gutenberg University and the government of the "Bundesland" Rhineland-Palatinate.

Notes

1. Federal law for the higher education sector (*Hochschulrahmengesetz*), paragraph 6.
2. For a detailed list, see the resolution of the Accreditation Council at: *www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse_AR/08.02.29_Kriterien_Studiengaenge.pdf*.
3. The Bologna Process required a quite extensive reform of German study programmes. The prevailing *diplom* and *magister* programmes had to be transformed into bachelor and master programmes. As a result, nearly all study programmes had to be rearranged and accredited at the same time.
4. First estimates based on the experience of JGUM suggest that the costs for system accreditation of study programmes are about half that of prevalent programme accreditation.
5. The Kultusministerkonferenz consists of representatives from the federal states of Germany. It responds to and regulates requests for the accreditation of study programmes.
6. The Accreditation Council specifies several formal requirements and criteria for the assessment of quality management systems. For more details, see: *www.akkreditierungsrat.de*.

7. As a part of this institutional review, certain aspects of all the HEI's study programmes are analysed. Furthermore, 15% of all study programmes are reviewed in depth. For more details, see: *www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse_AR/08.02.29_Regeln_Systemakkreditierung.pdf*.
8. Data for the winter term, 2007/08.
9. After the establishment of bachelor and master programmes in all fields of study.
10. Another part of the university's quality management system which should be mentioned is the department of organisational development. It concentrates mainly on the optimisation of administrative processes and on furnishing a management information system for JGUM.
11. Some departments (*Fachbereiche*) include very heterogeneous scientific institutes, *e.g.* for sports science as well as politics and media studies. In these departments, quality discussions at the department level are clearly limited.
12. ZQ also offers and co-ordinates advanced vocational training on aspects of teaching and research quality (mainly higher education didactics and the promotion of young researchers).
13. After having operated for a few years on a project basis, ZQ was founded as a permanent institution in 1999.
14. Research-oriented master programmes in particular are checked for their compliance with the university's research profile.
15. If examination regulations are not modified, the senate committee for teaching and learning and the senate do not need to be consulted on whether or not to approve the continuation/reaccreditation of the study programme.

References

Donabedian, A. (1980), *The Definition of Quality and Approaches to its Assessment*, Explorations in Quality Assessment and Monitoring, Health Administration Press, Michigan.

Fähndrich, S. and T. Grendel (2007), "Systemakkreditierung an der Johannes Gutenberg-Universität Mainz – Voraussetzungen und Ziele", *VHW Mitteilungen – Informationen und Meinungen zur Hochschulpolitik*, No. 4, pp. 13-16.

Grendel, T., U. Schmidt and E. Springer (2006), "Steuerung und Steuerungswissen im Hochschulsystem", in F. Hamburger and S. Hradil (eds.), *Steuerungswissen im Bildungssystem, Mainzer Beiträge zur Hochschulentwicklung*, Vol. 10, pp. 119-148.

Kehm, B. and U. Lanzendorf (2006), "Germany: 16 Länder Approaches to Reform", in B. Kehm and U. Lanzendorf (eds.), *Reforming University Governance: Changing Conditions for Research in Four European Countries*, Kassel University Press, pp. 135-186.

Kehm, B. (2007a), "The Evaluative State and Higher Education Policy in Germany", in J. Enders and F.A. van Vught (eds.), *Towards a Cartography of Higher Education Policy Change. A Festschrift in Honour of Guy Neave*, Center for Higher Education Policy Studies (CHEPS), Enschede, Vol. 1, pp. 139-148.

Kehm, B. (2007b), "Struktur und Problemfelder des Akkreditierungssystems in Deutschland", *Beiträge zur Hochschulforschung*, Vol. 29/2, pp. 78-97.

Kis, V. (2005), *Quality Assurance in Tertiary Education: Current Practices in OECD Countries and a Literature Review on Potential Effects*, OECD Publishing, Paris.

Krücken, G. and F. Meier (2006), "Turning the University into an Organizational Actor", in G.S. Drori, J.W. Meyer and H. Hwang (eds.), *Globalization and Organization: World Society and Organizational Change*, Oxford University Press, Oxford, pp. 241-257.

Mintzberg, H. (1992), *Structures in Fives: Designing Effective Organizations*, Prentice Hall, New York.

Neave, G. (1988), "On the Cultivation of Quality, Efficiency and Enterprise: An Overview of Recent Trends in Higher Education in Western Europe, 1986-88", *European Journal of Education*, Vol. 23, pp. 7-23, Blackwell Publishing Ltd.

Owen, J.M. and P.J. Rogers (1999), *Program Evaluation: Forms and Approaches*, Sage, London.

Schmidt, U. and J. Horstmeyer (2008), "Systemakkreditierung: Voraussetzungen, Erfahrungen, Chancen am Beispiel der Johannes Gutenberg-Universität Mainz", *Beiträge zur Hochschulforschung*, Vol. 30/1, pp. 40-59.

Schmidt, U. (2009), "Evaluation an Deutschen Hochschulen: Entwicklung, Stand und Perspektiven", in T. Widmer, W. Beywl and C. Fabian, *Evaluation Ein Systematisches Handbuch*, VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 163-169.

Weick, K.E. (1976), "Educational Organizations as Loosely Coupled Systems", in *Administrative Science Quarterly*, Cornell University, Vol. 21/1, pp. 1-19.

Whitley, R. (2008), "Constructing Universities as Strategic Actors: Limitations and Variations", in L. Engwall and D. Weaire (eds.), *The University in the Market*, Vol. 84, Portland Press, London, pp. 23-37.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

The intended and unintended effects of the Bologna reforms

by
Sybille Reichert
Reichert Consulting: Policy and Strategy Development in Higher Education, Switzerland

Bologna reform eulogies and protests tend to focus on the benefits and shortcomings of the new two-tier curricula, their implementation and orientation. In this article, an assessment of the Bologna reforms is made in terms of their larger and less widely discussed systemic and institutional effects – which go far beyond the original reformers' intentions. Apart from the introduction of new degree structures, the two Bologna reform dimensions which have been most readily adopted and dynamically implemented are the overhaul of Europe's quality assurance system and the recent reforms of doctoral education. In contrast, the visionary goals of using learning outcomes and competencies as the structuring principle of all curricula in order to ensure greater transparency and reliability, and of promoting student-centred learning, have only been adopted by few countries and institutions. However, the Bologna reforms have also had a range of unintended effects on systems and institutions that often go unnoticed when discussing their impact on European higher education. These include redefining the relationship between institutional profiles, strengthening central institutional leadership and mobilising horizontal communication within institutions.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Les réformes de Bologne : effets prévus et imprévus

par
Sybille Reichert
Reichert Consulting : Développements politique et stratégique dans l'enseignement supérieur, Suisse

Les éloges de la réforme de Bologne et les protestations dont elle a fait l'objet ont tendance à se concentrer sur les avantages et les inconvénients du nouveau cursus réparti sur deux cycles, sur sa mise en œuvre et sur l'orientation. Cet article constitue une évaluation des réformes de Bologne en termes de leurs conséquences, plus importantes mais moins débattues, au niveau systémique et institutionnel, qui vont bien au-delà des intentions d'origine des auteurs de la réforme. Au-delà de l'introduction de nouvelles structures de cycle, les deux dimensions de la réforme de Bologne qui ont été adoptées le plus facilement et mises en œuvre avec la plus grande dynamique sont la restructuration des systèmes d'assurance de la qualité en Europe et les récentes réformes de l'enseignement doctoral. En revanche, les objectifs visés par l'utilisation des résultats de l'enseignement et des compétences comme le principe de structuration de tous les cursus afin d'assurer une plus grande transparence et une meilleure fiabilité, et de promouvoir un enseignement centré sur l'étudiant ont uniquement été adoptés par de rares pays et institutions. Toutefois, les réformes de Bologne ont également entraîné une série de conséquences imprévues sur les systèmes et les institutions qui passent souvent inaperçues lorsque l'on parle de leur impact sur l'enseignement supérieur européen. Ces conséquences incluent la redéfinition des rapports entre les profils institutionnels, le renforcement de la position centrale des institutions et la mobilisation horizontale de la communication au sein des institutions.

Introduction

Ten years after the adoption of the Bologna Declaration, the Bologna reforms are still debated as fervently as in the initial years of their development. Student protests in Austria, France, Germany, Greece and Spain are being waged in the name of Bologna; disgruntled professors criticise its implementation in internal and public meetings and the media. At the same time, policy makers and higher education representatives in Europe and abroad applaud the scope and depth of the changes. Ten years ago it would have been beyond most people's expectations that the Bologna reforms would manage to mobilise so many academics and students into discussion, protest, criticism, fervent support or engaged appeals, and into spending hundreds of hours in meetings to orchestrate changes in courses, standards and procedures.

When a group of ministers of education signed the Bologna Declaration in 1999, their goals were seen by many stakeholders as another instance of unreliable rhetoric which has become a feature of many European ministerial meetings on education: no ties attached, no contractual obligations, no previous discussions with finance ministers, no actions to follow. But in the case of Bologna, the sceptics were proven wrong. Twenty-nine ministers committed to a common agenda of concerted national reforms of their higher education systems in order to create more transparent and comprehensible degree structures in a common European Higher Education Area. This engagement soon turned into action plans, debates with the higher education sector, new regulations and orchestrated institutional reforms across entire countries. A few early adopters forged ahead with the formulation of new laws and national action plans, and made others feel obliged to follow up on their own declarations to avoid loss of face. Rhythmed by their biennial meetings, a remarkable inter-ministerial peer pressure game emerged. Regular monitoring brought about naming and shaming, and sector consultation added sense, details and additional dimensions to the broad strokes of reform goals. Out of the Bologna Declaration grew the Bologna reforms, welcomed by some, hated by others. Everyone in the higher education sector, from policy makers to professors and administrators, was kept busy in consultations on legislation at national level, in curricular and administrative restructuring at institutional level, and in devising new or revising old quality assurance processes at all levels.

So what are we to make of this past decade of far-reaching national reforms and their unprecedented degree of European orchestration? What are the major achievements to date? Where have the reforms fallen short of the expectations vested in them by the most convinced reformers? What unintended effects can one observe, good or bad? And where should the reform process be heading if it wants to live up to its fundamental goal of improving the quality and international attractiveness of higher education in Europe?

When making any summary assessments and observations regarding the effects of the Bologna reforms, access to relevant data constitutes a challenge. While the official stocktaking reports are updated every two years (as a basis for the ministers' meetings), they provide generally aggregated data but lack details about institutional realities. Moreover, their presentation of national situations has, in some cases – and at least until 2007 (the 2009 report being more scrutinising) – glossed over some higher education institutions' awkward facts to avoid national loss of face.

To gain insights into institutional realities and challenges, the best comparative sources are the European University Association (EUA)'s *Trends* reports, last published in 2007, which base their assessments on surveys of a large part of the sector (over 800 institutions). Unfortunately, the next report will only be published in March 2010, and its comparative institutional data comparing realities across Europe is at least two years old. The only other comparative study which has generated its own data through surveys and interviews (Witte, 2006) is even older. A more recent comparative study looking at some of the effects of the reforms through student surveys was carried out by the European Student Union (2008). And in some countries, such as the Netherlands (Westerheijden *et al.*, 2008), Germany (Winter, 2009) and Switzerland (CRUS, 2008; Dell'Ambrogio *et al.*, 2009), studies of the Bologna implementation, institutional realities and effects are available now. They contain much more detailed analyses than can be found in the official national stocktaking reports.

However, although the Bologna process is applied by 46 countries, very few of them have drawn up such reports: the institutional implementation is still too young in most countries to evaluate the new curricula and their effects as the key ingredient of the reforms. Also, they do not offer any trans-national comparative data. An extensive comparative study of the effects of Bologna implementation in the 46 participating countries is currently being undertaken by a European consortium (including CHEPS University of Twente, INCHER at the University of Kassel and ECOTECH in the United Kingdom). This study will fill the gap, but results will only be published in 2010. For the time being, only the above-mentioned data sources – as well as the many institutional visits the author has undertaken during the course of her work – could reasonably be used as a base for observations and assessments.

To assess the successes and shortcomings of the Bologna reforms to date, one needs to distinguish between two types of effects: those that were intended (even though such intentions may have been vague), planned, monitored and widely discussed among policy makers and higher education representatives, and those that emerged unintentionally, mostly unnoticed and rarely appreciated. It will be argued later in this paper that these latter effects are perhaps even more profound and transformative than the intended actions that followed the Bologna objectives.

The intended effects of the Bologna reforms: their achievements and shortfalls

A new European policy arena

First and foremost, the Bologna Process has produced a new form of trans-national policy development which triggers off multiple national processes of decision making. It sets soft norms and places them in a context of trans-national pressure to comply, combining top-down initiatives with bottom-up agenda setting. Moreover, it intertwines national and European policy issues in mutually reinforcing effects and, through its organisational process, catalyses and accelerates national higher education reforms. While the intention of this process was to create a European Higher Education Area, some effects (as will be discussed in the section below on unintended effects) went well beyond the intention of creating a common arena for policy exchange.

The idea of bringing together national policy discussions in a European forum of exchange had already begun to take form in the 1990s through various initiatives of the European Commission and the regular European Council of Ministers meetings. The Bologna Process went beyond this forum, however, by virtue of its committed nature, its closely knit follow-up process and its links to defined national actions. Instead of the Commission developing policy issues in a common context to elicit national responses, Bologna introduced a nationally driven European process in which common issues were defined by combining top-down initiatives with bottom-up agenda-setting. While the reforms seemed imposed at first – given that the ministers took the first initiative with the Bologna Declaration in 1999 – the actual interpretation and details of the initially vague agenda were co-defined by representatives of the higher education sectors working through representative organisations. At European level, these included the European University Association (or its predecessor CRE before 2002), EURASHE (the European association of non-university higher education institutions). At national level, higher education associations fed their input into the "Bologna Follow-up" process.

From the Prague Communiqué (2001) onwards, the higher education sector contributed actively to the definition of the Bologna reform agenda, adding areas of reform activity such as the doctoral education reforms, or re-emphasising other aspects such as the autonomy needed to modernise universities. Likewise, the quality assurance dimension of the Bologna reforms was defined in close co-operation with the quality assurance agencies (through the European Association for Quality Assurance in Higher Education – ENQA). The higher education sector was represented through the European University Association, and the students represented by the European Students' Union. The latter emphasised, for example, the importance of higher education as a public value and that of the social dimension of European higher education (including questions of access and stipends). These bodies also repeatedly reminded the political authorities about the vision of flexible learning paths as a goal of the curricular reforms. Thus, reform issues were added or emphasised by the sector itself, and originally vague concepts were spelt out in terms of more operational reform plans.

The act of defining a higher education policy agenda at European level and translating the European reform agenda into multiple national reform contexts also resulted in a confluence of reform agendas. At European level, the Bologna objective of creating a common Higher Education Area was soon linked with the Lisbon agenda which sought to create a common European Research Area (ERA), as well as to increase research and innovation competitiveness and investment in research and development. Existing channels of communication were used by policy entrepreneurs from both policy domains. The modernisation of universities and the reforms of research training served as a bridge between both areas and allowed for mutual policy influence. Likewise, at national level, local reform issues were linked to the Bologna reforms. In the end, it is hard to tell whether the Bologna reforms pushed the national reforms or *vice versa*, but a mutual reinforcement effect can be noted in most contexts (see the section below on unintended effects).

From minimal compliance to far-reaching institutional change

Close to the self-imposed deadline of 2010, the top reform objectives appear, at first glance, to have been implemented in most signatory countries, at least in all those which had signed the Declaration by 2003. The cornerstone of the reforms, the introduction of a two-tier bachelor/master degree structure, seems to have been largely achieved. According to the last European University Association *Trends* report (Crosier *et al.*, 2007), 83% of higher education institutions (out of a sample of roughly 1 000) indicated that they had the new two- or three-cycle structure (including the doctoral level) in place. By 2010, the vast majority of students will be enrolled in programmes reflecting the Bologna structures.

In many countries curricular reforms remain rather formal and superficial in nature, in comparison to the more ambitious expectations vested in them originally. In principle, the introduction of more learner-centred teaching and the definition of flexible curricula in terms of competences and skills should have improved the quality and international attractiveness of European higher education; however, this has not yet been achieved. To bring about these improvements, better student-staff ratios, didactic training and other staff development measures would have been needed, as well as substantial investments which few countries were ready to provide. However, in most countries, many institutions' budgets are already stretched given the increase in their counselling activities which has not been off-set by greater funding.

Furthermore, in many countries, higher education institutions undertook far-reaching curricular restructuring without abandoning the old expectations of knowledge scope and range, sometimes pressing as many courses into the short programmes as the territorial negotiations between professors would allow. Hence students may have less time for independent study than before, less room for individualised optional courses, even less flexibility to integrate mobility time within their programme. This is in total contradiction with the essential Bologna goals. While ministers can evade this reform aim and see their country appear in saturated green in the Bologna Follow-up monitoring reports, students are losing hope and patience with the realities of their everyday student life, and some are even staging large-scale protests.

Behind this generalised comment, however, one should distinguish a wide range of institutional realities. These vary not only between but also within countries and even within institutions, ranging from reforms that amount to mere cosmetic surgery to deep institutional reforms addressing past problems such as fragmented curricula or ineffective teaching methodologies. As the *Trends* studies have shown (Reichert and Tauch, 2005; Crosier *et al.*, 2007), many institutions have actually adopted the reforms as an opportunity to address a range of changes which they had not been able to promote as easily without such external pressure. For a sizeable minority of institutions the reforms have meant introducing new teaching approaches and curricular design principles, more options and flexible interfaces between programmes, more opportunities for interdisciplinary study and research, and new internal quality assurance processes. At many institutions, the new master level became the focus of institutional positioning by developing new interdisciplinary programmes in areas of strengths, or by addressing new professional profiles. In contrast, other institutions (even when working in the same regulatory framework) chose to implement the reforms in a minimalist fashion by simply regrouping existing courses and creating a pro-forma *cesura* between the bachelor and master levels, with the expectation that most

students would go straight from one level to the other without any change of direction or institution.

Similarly, the definition of learning outcomes for each course or qualification profiles for whole programmes had mixed results. It was seen as a useful opportunity for reflection and reorientation by some institutions and merely as a bureaucratic exercise by others.

These institutional choices are attributable to three factors:

1. The willingness of national policy makers and administrators to develop the reform details and processes in close consultation with institutions, taking account of their aims and concerns. This attitude was crucial in helping them to develop their sense of ownership, in defining reform contents that were perceived to be meaningful to higher education leaders, teachers and students, and in providing information and support adapted to institutional needs.
2. The resources invested in the reform process and measures. Given the proposed breadth and depth of the reforms, these were essential if provision was to be made for the additional staff to support curricula with more options, to allow for teaching in smaller groups, as well as for promoting competencies. In 2005, only seven countries had actually committed additional funds to help institutions implement the reforms. Even fewer (*e.g.* Norway and Switzerland) invested additional resources to support the more staff-intensive approaches to teaching and curricular design which the Bologna reforms had proposed.
3. The willingness and ability of rectors and vice-chancellors to lead the institutional reform process. Given the scope of the reform project and the fact that it concerned all academic units, institutional leadership and central management support were vital to the success of the reforms. One may even look at Bologna implementation as a sort of leadership test for many rectors, with a challenge factor added at the more decentralised and conservative institutions. At departmental level, substantial resistance was encountered. Rectors had to strike the right balance between persuasion and swift decisions and actions. The ability to create a coherent approach to quality improvement was a challenge in itself. The resulting resistance was exacerbated by the fact that other reforms were often introduced at the same time. In particular, there were simultaneous changes which introduced new forms of higher education governance (these preceded or accompanied the Bologna reforms in many countries and increased institutional autonomy as well as central leadership); they were received with mixed feelings at departmental level.

4. Problem awareness. As many interviews with institutional actors in the framework of the *Trends* reports showed, the success of Bologna reform implementation depended strongly on overall institutional perceptions of the need for reform. Wherever the need for improvement in related areas had already been identified, the Bologna reforms were much more easily adopted and developed into an institutional improvement programme (Reichert and Tauch, 2005; Crosier *et al.*, 2007).

Limited progress on the transition to student-centred teaching

The varied successes of the curricular reforms are mirrored in the mixed reactions of the higher education sector to the idea of introducing student-centred teaching and learning. It should be noted that this objective was only integrated into the explicit European reform agenda as a second step, even though some reform visionaries and many students had attributed the highest importance to it from the beginning. Moreover, since this objective required more favourable student-staff ratios than most institutions currently boast, only few governments adopted an explicit reference to introducing more student-centred teaching approaches in their reforms.

In fact, only few institutions have made the transition to student-centred or competency-based teaching and learning a key part of their curricular reforms. This is because no countries were ready to invest the full range of resources needed to realise such far-reaching pedagogical change. It was only under pressure from student representatives in the European Bologna Follow-up Group that the objective of greater student-centred teaching received more political and rhetorical attention in recent years, although financial backing was rare. Most recently (November 2009), the European Students' Union launched a project to shed more light on the realities of current teaching conditions and competency-based teaching, in the hope of raising awareness of the need for more student-centred teaching. However, there is still little national money to "walk the talk" and only few university teachers or leaders are willing to make this change their primary cause. One would hope that the 2009 student protests will exert some pressure on politicians to take these concerns seriously, but the medium-term budget restrictions after the financial crisis do not help the cause.

A similar fate befell the accompanying transparency instruments, the European Credits Transfer System (ECTS) and the Diploma Supplement. Implementation appears to be far-reaching if one looks at the number of institutions that have formally introduced these instruments across their institutions. A vast majority of institutions are now using the ECTS as their credit accumulation system and providing Diploma Supplements to make the student records more readable for employers. However, more often than not, these instruments are not being applied according to the way they were

designed in the recent Bologna context, *i.e.* by including explicit reference to learning outcomes and/or competency profiles. Frequently, contact hour-based course credits were simply multiplied to calculate the new ECTS, without any attention being paid to actual workload or to the contribution of the module to clearly defined learning outcomes, as had been intended.

At national level, the Qualifications Framework – which ministers agreed to develop in 2005 (Bergen Communiqué) to describe levels of competencies at different degree levels for employers or students – has only recently been developed. Moreover, the combination of all these competency-based instruments into a continuous tapestry – which should foster competency and learning outcome orientation at all levels – still remains a vision for the future. Nevertheless, some of the new national qualifications frameworks are clearly designed to offer more reliable opportunities for transition between different types of higher education institution. Articulation between different types of institution and levels of higher education has become a key concern for Ireland, Germany and Switzerland. It has even spilled over to other continents, where qualifications frameworks have been revised in light of European developments: this is the case for the Australian Qualifications Framework 2009. Interest in some of the European methods is also growing (as is the case in the United States (Tuning USA Conference, April 2009; AACU Conference, January 2010). Significant efforts still need to be made before there will be fluid transition between different types of higher education institutions. However, the new transparency which prevails is making it more difficult for institutions to close their doors to applicants from other types of institution which lack sound arguments relating to the reality of competencies.

Signs of changing attitudes

However disappointing the achievements may seem to date to all those who had invested high hopes in the Bologna Process as an accelerated solution for urgent reforms, one should not jump to pessimistic conclusions. There are increasing signs of a slow but remarkable attitudinal change in teaching and curricular design. Within institutions, an increasing number of departments or faculties are becoming interested in more profound reflection on the competencies their programmes are meant to promote and on the methods needed to assess them. More and more departments are discussing student-centred teaching concepts, defining learning outcomes for their programmes and courses, experimenting with competency-based curricula and new assessment methods. Many institutions have intensified their efforts (*e.g.* through programme boards) to entertain a more regular dialogue with employers on the relevance of their curricula and the competencies of their graduates. To consider societal or employer needs in the context of course development is no longer seen as selling out academic values, except in a few

academic niches. While "employability" may still be regarded as a provocative term in many European academic circles and is often associated with short-term perspectives, the need to prepare students for long-term adaptability to changing labour markets is indeed being taken seriously as an aim of higher education programmes.

Engagement in this deeper reform agenda started with a few early adopters, mostly located in north-west Europe, and is spreading east and southwards, in many cases facilitated by European conferences and national best-practice workshops. An external indication of such a shift is the fact that the availability and quality of counselling and tutoring services has expanded significantly, as has the attention to student needs in services and internal quality assurance processes (Crosier *et al.*, 2007). Moreover, even where institutions had initially engaged in reforms somewhat reluctantly, many now perceive benefits, or at least potential, in terms of greater flexibility within and between programmes and institutions.

The most dynamic areas of the Bologna reforms: quality assurance and doctoral reforms

Other than the central question of curricular reform, the Bologna reforms have developed greatest momentum in the area of quality assurance. Profound reforms, increased compatibility and heightened attention to robust internal and external quality assurance may be regarded as the most resounding successes of the Bologna Process so far. Beyond the introduction of common standards for external quality assurance (QA), the idea of constant self-improvement and voluntary integrated institutional quality processes beyond simple accountability is spreading across Europe, as institutional data from the *Trends* reports and many quality audit reports suggest.

The key focus of the Bologna reforms, however, has been on processes relating to external QA, common QA standards, guidelines and comparability across Europe, as well as exchange and mutual understanding between QA agencies. The European Standards and Guidelines for QA (2005) include a common approach (often called misleadingly a "methodology"). This comprises regular institutional or programme reviews involving self-evaluations and external peer reviews, leading to improvement measures, as well as student participation in QA processes and on evaluation panels. The fact that all higher education institutions and systems will have to apply these common elements of QA through their own national QA processes will have a profound impact on institutional quality awareness and processes. To be accepted onto the European Register of QA Agencies (the mechanism introduced in 2008 to ensure that quality agencies comply with the common European guidelines) is already proving to be more demanding than even some well-established QA or accreditation agencies would have anticipated.

Since external quality assurance, agency formation and new QA processes have kept higher education institutions well occupied, the concern with institutional quality culture, which the EUA has attempted to emphasize, has not been as highly prioritised. In light of new governance structures and recently increased autonomy in many European countries (CHEPS *et al*., 2007), issues of accountability and external quality control seem to dominate the quality discussion. In future, the Bologna reforms will have to face the challenge of how to consolidate the added value of quality assurance reforms while keeping the spirit of regular voluntary self-improvement alive. Higher education institutions are caught in an increasingly dense Web of external reviews. As funding sources diversify, accountability demands are also diversifying. Quality assurance is becoming more compliance and accountability-driven, with less time and energy and motivation left for genuinely improvement-oriented quality assurance processes. In the next decade, QA fatigue may be the biggest challenge for the more mature QA systems.

Given Bologna's focus on teaching and learning, other QA issues (in the wider sense of the term) have not been addressed in spite of their important impact on institutional and individual behaviour. These include, for example, the impact of popular international rankings on concepts of quality measurement as well as on institutional behaviour (Hazelkorn, 2008) which also raises concerns about policy priorities and resource allocation in the future. Higher education policy makers across Europe will be challenged to strike the right balance between the demands of research quality and international competition, on the one hand, and the many other necessary functions of higher education and their quality demands, including the need to cater for larger and more diverse groups of knowledge users and producers (Reichert, 2009). At the most recent Bologna ministerial meeting, the challenge of institutional diversity was explicitly mentioned. We can therefore expect that Europe's key policy makers will make use of the successful communication channels which have been built up in the context of Bologna for these wider policy issues, not so much to develop common guidelines or standards, but to exchange ideas and good practice.

The other reform area in which remarkable momentum has been achieved is that of doctoral education (added to the Bologna reform agenda via the Berlin Communiqué in 2003). This area was given policy prominence because of the increased attention on Europe's international research competitiveness. In particular, the attractiveness of European universities to young researchers has become a key concern in times of increasing brain circulation or potential brain drain. But beyond the concern with the attractiveness of the European Research Area, recent attention has also focused on the quality and relevance of doctoral education, as part of the wider Bologna concern with societal relevance and employability. This

becomes apparent in the competency orientation of newly developed doctoral programmes which often seek to address diversified research careers by fostering research-related transferable skills (such as research project management skills, team skills, presentation skills). Reform efforts concerning graduate programmes and graduate schools have also resulted from a heightened awareness of the need to improve the quality of teaching, supervision and tutoring, reflecting the systematic quality assurance perspective which Bologna has helped to foster. This is also reflected in the rising popularity of graduate schools, which place importance on supervision quality, doctoral committees, reliable doctoral agreements, opportunities for interdisciplinary exchanges, relevant skills training and career development.

All in all, the intended effects of the Bologna reforms have been slowest where hopes were highest, namely in the planned widespread reforms of teaching quality and approaches (competency and learning outcome orientation), since these involve attitudinal changes rather than mere changes of procedures and structures. In contrast, the intended effects have been strongest and fastest in what may at first have appeared to be more secondary areas of activity. Most change took place in areas where the reforms were clearly led by concerned actors and sectors: quality assurance awareness, good practice and mutual trust have spread faster and more deeply than many of us predicted in 1999. These good surprises are due to the persistent engagement of QA agencies through ENQA, the higher education sector (through the European University Association and the European Association of Higher Education Institutions) and students (through the European Students' Union). Doctoral reforms have developed a momentum and institutional initiative which few would have predicted in 2003. Again, intense sector engagement (through EUA and others) contributed to the high level of activity in this domain, supported by the urgent interest in improved international competitiveness of university research training in some national systems.

Unintended effects: system shifts and institution building

In addition to the effects of explicit Bologna actions, one should highlight the deeper and less predictable side-effects which the Bologna reforms have brought about, both at system and institutional levels. While these effects would not have occurred without the breadth and depth of the Bologna agenda, they have not been consciously pursued and are rarely noted. Nevertheless they may have changed higher education in Europe even more profoundly.

Among the unintended effects of the Bologna reform process, four system-level ones are of particular interest to higher education researchers and policy makers.

Increased transnational policy osmosis

The Bologna Process has created such a dense network of policy makers and institutional leaders exchanging good practice or policies that new ideas and choices are much more easily and rapidly transferred from one national or institutional context to another than in the past. While solutions are adapted to local constraints and conditions, issues and policy options are exchanged between different national contexts. In fact, policy discussion between countries has reached such intensity that it makes the European higher education landscape begin indeed to resemble a common European Higher Education Area. Bologna has enabled frequent transnational exchanges and policy consultation between European rectors' conferences, university presidents and relevant agencies facing similar systemic transitions or policy dilemmas. Such policy exchanges go far beyond the scope of the Bologna action lines, encompassing such questions as research evaluation schemes, schemes to promote vertical differentiation among higher education institutions, or experiences with institutional mergers and consortia.

Bologna's role as catalyst for other higher education reforms

The Bologna rhetoric in favour of urgent qualitative improvements and international competitiveness and the far-reaching systemic implications of its curricular and quality objectives have not only mobilised everyone in higher education but also acted as a catalyst – or even framework – for other national reforms. Many national reform agendas went well beyond the Bologna action plan, but were presented as inevitable ingredients of an increasingly international higher education arena, and closely associated with the Bologna agenda. In addition, the simultaneous Lisbon agenda with its emphasis on innovation reinforced the relevance theme even further. Such reforms comprised greater institutional autonomy, new governance structures with stronger institutional leadership and more direct stakeholder influence, as well as increased proportions of performance-based funding. Examples of deeper system changes which were attached to the Bologna agenda can be found in Austria, the Netherlands, Denmark, Norway, some of the German *Länder*, Belgian Flanders, Portugal, Slovakia and (recently) Spain. A recent study on governance in Europe funded by the European Commission documents these convergences (CHEPS *et al*., 2007).

Renegotiated institutional profiles and "vocational drift"

Through the new bachelor's degrees emphasis on graduates' employability and labour market relevance, the Bologna reforms have strengthened the professional orientation of higher education. Consequently, they have also strengthened the position of professionally oriented

institutions such as University Colleges, *Hogeschols* or *Fachhochschulen*, as can be seen for instance in Norway, Germany, the Netherlands and Switzerland. Due to the growing importance given to innovation by many countries as well as the EU (Lisbon agenda), the non-traditional functions of universities have risen in public recognition and status to such an extent that even some traditional universities are pushing the innovation and professional orientation of their research and teaching programmes.

Given the rise of public attention to widely used international rankings, the so-called research universities are increasingly seen as a primary element of an internationally oriented higher education system. Nevertheless, given that greater attention is being paid to the relevance and contributions of research to innovation and societal challenges, the role of research which is driven simply by curiosity and traditional academic research training has been weakened. Likewise, research training is seen as an important preparation for non-academic research and employment, as illustrated by the attention to transferable research skills training. Considerable research training reform efforts which widen the scope and utility of research training can be found in Germany, Switzerland, Norway, Finland, Belgium and France. In light of these more utilitarian attitudes to higher education, universities are becoming less traditional, and the few remaining ivory towers are being dismantled.

At the same time, professionally oriented university colleges are expanding their applied research activities, with the help of state funds and new funding instruments, and up-grading the research profiles of their staff. Thus, the dividing lines between the old institutional types (universities *vs*. Fachhochschulen) are no longer defined by the presence or absence of research or research training, but rather by the (weakening) basic research orientation of the universities *versus* the business responsiveness of university colleges (Reichert, 2009). Increasingly, it is expected that the dividing lines between different types of institution will be blurred, and smooth transitions between them will characterise the overall systems (helped by the new qualifications frameworks). This is despite the fact that separate institutional types with their different mission mixes are still politically desirable in most dual and even integrated higher education systems in Europe.

Whether such institutional differentiation and complementarity will survive in the long term will depend on differentiated financial and career incentives for the different institutional orientations (Reichert, 2009). Bologna's relevance agenda has pushed these questions to the foreground and has set the stage for a new conflict between institutional types and profiles over resources and support.

Mass versus elite degrees

The *Trends* studies on the implementation of the Bologna reforms show that the new two-tier curricula have often been conceived essentially as a mass-oriented bachelor's and a selective or elite master's programme. While many systems continue to regard the master's as the main university degree and so-called consecutive masters have been put in place to guarantee such continuity, the nature of the first three years has clearly changed. This is partly due to the overburdened programme structure into which too many of the five-year course ingredients have been crammed. The more elective offerings have covered research project work, and independent study has often been pushed back to a secondary level. Staff-intensive research training has been restricted to the master's level. In addition, many new specialised and explicitly selective master programmes have been developed to help position the research strengths of the respective institution. The master's level is becoming a marketing arena for higher education institutions where institutional profiling and selectivity are given free reign. Whenever bachelor's degree education is accessible to all holders of high school diploma, and when master's programmes are selective, there are signs of institutions differentiating their mass and élite cohorts along bachelor's and master's lines. Some less well-resourced national systems (such as that of Hungary, Slovakia and the Czech Republic) already allow only the best qualified students to continue into master's programmes. The extent to which bachelor's degrees become the main degree – at least in some fields – will ultimately be determined by cost pressures and labour market reactions to the new degrees.

At institutional level, the Bologna reforms have had three major unintended effects.

Strengthened central leadership

The Bologna reforms have required institution-wide consultation and decision-making processes of an unusual breadth and depth. Deliberations had to cover all subject areas, delving into the *raison d'être*, organisation and contents of programmes; they raised over-arching concerns with common guidelines and approaches to new degrees as well as to the quality processes underpinning their creation and future management. To steer such a review and reform process over a short time required strong central leadership and support. Moreover, formal, strong central leadership was made possible for the first time in some countries. While these reforms were not only generated as part of the Bologna reforms, they reinforced each other. With greater autonomy and increased accountability demands, central administrations had to face a range of new tasks. At the same time, the competitiveness rhetoric resulted in more competitive funding arrangements which tended to

promote the need for strategic positioning and priority setting. While many of these changes were not explicitly part of the Bologna measures, they were experienced by academics and students as an effect of the overall reform agenda. For many, Bologna has become the symbol of changed higher education attitudes, of a different academic culture and landscape. Remarkably, in several countries the new governance structures were seen as an erosion of academic freedom, egalitarian values and democratic culture, resulting in controversy from students and academics (*e.g.* France and Spain).

Increasing stakeholder influence

Associated with the central concern of social and labour market relevance and relevant competency profiles, the Bologna reforms have contributed to increased communication with stakeholders. Even though dialogue was already well developed in some subject areas such as engineering, medicine, economics, business management and law, and in all professionally oriented higher education institutions, the Bologna reforms have broadened its scope and raised its status. Stakeholder influence has also been strengthened by the introduction of new governance structures such as stakeholder boards which serve as executive or advisory bodies in higher education institutions. This is the case in Austria, Denmark, Norway, many of the German *Länder*, Spain and finally the Netherlands, where stakeholder boards are already well established. While the direct influence of external stakeholders still raises scepticism at most traditional institutions, the idea of consulting outside partners on programme orientation is now more widely accepted among academics from all types of institutions.

Increased horizontal communication

The Bologna reform process has, in many instances, contributed to institution building beyond the often disjointed array of separate faculties. Given that the reforms involved a simultaneous review and overhaul of programmes in all institutions, many of them made use of this opportunity to develop closer interfaces between disciplines, pathways from several disciplinary areas to an interdisciplinary programme at master level, or to establish new interdisciplinary programmes. Often, new curricular structures contributed to softening the boundaries between faculties or departments, which often act as obstacles to institutional innovation. Sometimes, institutional incentives, such as strategic reserves, have been established or expanded to foster interdisciplinary programmes or research. Especially at the new master's and doctoral level, such new interfaces were sought and institutionalised in new graduate programmes or schools or new interdisciplinary "centres of excellence". Examples of such developments abound in France, Germany, Switzerland, the Netherlands and the Scandinavian countries. They

have contributed to the creation of a matrix structure in institutions, cutting across and often weakening existing territorial and traditional decision-making boundaries. Moreover, enhanced communication and deliberation across internal administrative boundaries were needed to review and expand quality assurance processes and develop transparency instruments, such as ECTS, diploma supplements or guidelines for qualification profiles. This has resulted in side benefits such as new opportunities for internal co-operation and information flow (Reichert and Tauch, 2005).

Conclusion

Ten years after its inaugural Declaration, the Bologna reform process has led to an unprecedented wave of systemic and institutional change. However, only some of these changes were intended effects of the Bologna agenda. Among the intended reforms, implementation has been remarkably swift and efficient in the new two-tier curricular structures, the commonly aligned quality assurance systems and the widespread doctoral education reforms. However, whether these reforms will be effective enough to achieve the overall goal of quality improvement remains to be seen.

With respect to the deeper visions of change, more time and resources will be required. Student-centred teaching approaches and curricula that are organised around agreed and assessable definitions of learning outcomes and competencies presuppose attitudinal changes as well as staff development measures. These new approaches need appropriate reward structures which have to be developed through continuous academic leadership and persuasion. Such new teaching methodologies also require better student-staff ratios than most programmes in continental Europe currently offer. Nevertheless, given that so many institutions have embarked on this reform path and that teaching quality has become a key, widespread institutional concern, Europe may well witness an unparalleled leap forward in teaching performance and quality over the next decade.

It is doubtful, however, given the limited resources available, whether such a major change in attitudes and teaching methodologies will occur in a co-ordinated and consistent way across the whole of Europe. It also remains to be seen whether the vision of a Europe in which a substantial proportion of students and graduates will study or work abroad, with full recognition of their prior studies and qualifications, will become a reality within the next decade. This would presuppose a more substantial investment in higher education than most states are currently prepared to provide. But beyond such hopes for the realisation of Bologna's original vision and potential, one should also note the many positive effects on institutional development that Bologna has helped to achieve, even though they were not part of the original vision.

The author:

Sybille Reichert
Germania Str. 41
8006 Zürich
Switzerland
E-mail: sybille@reichertconsulting.ch
www.reichertconsulting.ch

No specific funding source was called upon to support the research presented.

References

AACU Annual Conference (2010), *www.aacu.org/meetings/annualmeeting/index.cfm*.

Center for Higher Education Policy Studies (CHEPS) *et al.* (2007), *The Extent and Impact of Higher Education Governance Reform Across Europe. Final report to the Directorate-General for Education and Culture of the European Commission*, European Commission.

Conférence des Recteurs des Universités Suisses (CRUS) (2008), *Rapport final 2004-07 de la CRUS sur l'état d'avancement du renouvellement de l'enseignement des hautes écoles universitaires suisses dans le cadre du processus de Bologne*, Bern.

Crosier, D., L. Purser and H. Smidt (2007), *Trends V: Universities Shaping the European Higher Education Area*, European University Association, Brussels.

Dell'Ambrogio, P., J. Rinaldi and J.F. Stassen (2009), *Étudier après Bologne : le point de vue des étudiant-e-s. Résultats de l'enquête nationale menée auprès des étudiant-e-s sur les conditions d'études dans les universités en 2008*, CRUS and VVS, Bern.

European Association for Quality Assurance in Higher Education (2005), *Standards and Guidelines for Quality Assurance in the European Higher Education Area,* European Association for Quality Assurance in Higher Education, Helsinki.

European Students' Union (2007), *Bologna with Student Eyes*, European Students' Union, Brussels.

European University Association (EUA) (2003, 2005, 2007), *Trends* reports. These reports are published every two years and constitute the most in-depth monitoring of the process. They take account of institutional realities (since 2003) all over Europe, surveying a sample of around 1 000 higher education institutions of all types. All three reports can be downloaded at *www.eua.be/trends-in-european-higher-education*.

Hazelkorn, E. (2008), "Learning to Live with League Tables and Ranking: The Experience of Institutional Leaders", *Higher Education Policy,* Vol. 21, No. 2, Palgrave Macmillan, pp. 193-215.

Reichert, S. (2009), *Institutional Diversity in European Higher Education*, European University Association, Brussels.

Reichert, S. and C. Tauch (2003), *Trends III: Progress Toward the European Higher Education Area*, European University Association, Brussels.

Reichert, S. and C. Tauch (2005), *Trends IV: European Universities Implementing Bologna*, European University Association, Brussels.

Stocktaking Reports of the Bologna Reform Process (2005), *www.bologna-bergen2005.no/Bergen/050509_Stocktaking.pdf;* (2007), *www.dcsf.gov.uk/londonbologna/uploads/documents/6909-BolognaProcessST.pdf;* (2009), *www.ond.vlaanderen.be/hogeronderwijs/bologna/actionlines/stocktaking.htm*.

Tuning USA Project and Conference (2009), *www.luminafoundation.org/newsroom/newsletter/Archives/2009-04.html*.

Westerheijden, D. *et al.* (2008), *Evaluation of the Bachelor-Master Structure and Accreditation in Dutch Higher Education: Aims, Ddynamics and International Reflections on the Working of the Bachelor-Master System and Accreditation in the Netherlands*, CHEPS.

Winter, M. (2009), *Das neue Studieren. Chancen, Risiken und Nebenwirkungen der Studienstrukturreform: Zwischenbilanz zum Bologna-Prozess in Deutschland*, Institut für Hochschulforschung Arbeitsbericht 1/2009, Wittenberg.

Witte, J. (2006), *Change of Degrees and Degrees of Change:. Comparing Adaptations of European Higher Education Systems in the Context of the Bologna Process*, CHEPS, Enschede.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

Knowledge as a common good: the societal relevance of scientific research

by
Lex M. Bouter
VU University Amsterdam, the Netherlands

Universities are, to a large extent, publicly funded. It is reasonable to expect that society should benefit as a result. This means that scientific research should at least have a potential societal impact. Universities and individual researchers should therefore give serious thought to the societal relevance of their research activities and report on them widely. Core questions they should be asking are: "Do we do the right things?" and "Do we do them right?". This implies that as well as indicators of scientific quality, attention should be given to indicators of societal relevance. These two considerations are examined in the context of current evaluation practices of academic research. Twelve indicators of societal relevance are proposed, focusing on both their socio-cultural and economic value. The examples given mainly concern the health and life sciences. This paper concludes with a discussion of the key challenges in evaluating the societal relevance of scientific research.

Higher Education Management and Policy
Volume 22/1
© OECD 2010

La connaissance comme « bien commun » : la pertinence sociétale de la recherche scientifique

par
Lex M. Bouter
L'université VU d'Amsterdam, Pays-Bas

Les universités sont, en grande partie, financées par l'État. Il est légitime de penser que la société puisse en bénéficier. Cela signifie que la recherche scientifique devrait au moins avoir un impact sociétal. Les universités et les chercheurs indépendants devraient, par conséquent, sérieusement considérer la pertinence sociétale de leurs activités de recherche et en tirer les conséquences à grande échelle. Les questions fondamentales qu'ils devraient se poser sont : « Faisons-nous les choses comme il faut ? » et « Les faisons-nous correctement ? ». Ceci implique qu'outre les indicateurs de qualité scientifique, l'attention devrait être portée sur les indicateurs de pertinence sociétale. Ces deux considérations sont analysées dans le contexte des pratiques actuelles d'évaluation de la recherche universitaire. Douze indicateurs de pertinence sociétale sont proposés, et l'attention est portée à la fois sur leur valeur socioculturelle et économique. Les exemples fournis concernent essentiellement les sciences de la santé et de la vie. Cet article s'achève sur une discussion portant sur les défis majeurs en matière d'évaluation de la pertinence sociétale de la recherche scientifique.

One of my champions is the British statistician Austin Bradford Hill (Bouter, 1993; Hill, 1952, 1965). In 1948 he introduced the randomised clinical trial to determine the effectiveness of medical treatments through randomised treatment allocation and arbitrary outcome assessment. At the time, this approach gave rise to heated debate and was criticised as unscientific and unethical. Sixty years later, Hill's position is no longer controversial. Today there is a consensus that, in cases where there is reasonable doubt, it is actually unethical not to conduct a clinical trial. The doctrine of evidence-based medicine is founded on this very principle. It centres on the notion that, while theory and fundamental research are essential, they nevertheless provide an insufficient basis for the application of scientific knowledge. This requires applied research in the relevant field. Hill formulated four questions that authors and readers of scientific articles should ask themselves:

- Why did you start?
- What did you do?
- What answer did you get?
- And what does it mean anyway?

I would like to discuss the last question in particular, and there are two key elements to my stance. The first is that researchers should reflect on the societal relevance of their work. The second is that universities should report on the work of their researchers and formulate concrete indicators of societal relevance.

This paper will begin by considering the relationship between scientific quality and relevance to society. It will then demonstrate the importance of focusing on societal relevance and its place within the traditions of VU University Amsterdam. Subsequently, it will identify a number of indicators of societal relevance and will propose ways for working with them. Lastly, it flags the challenges that lie ahead.

Quality and relevance

The dream of every modern-day university manager is a cockpit with a dashboard full of performance indicators. To be honest, didn't we all want to be pilots when we were growing up? I am such a manager and I know that my steering ability depends on the quality of those indicators. I also know that vision and wisdom are important when it comes to interpreting the readings on those displays and dials. With this in mind, we can make two demands of

research: that it is of high scientific quality and that it is relevant to society. The first demand is non-negotiable and is central both to the assessment of research proposals and the evaluation of academic research. The demand for societal relevance is less self-evident and gives rise to a great deal of discussion.

To begin with, relevance often depends to a large extent on the outcomes of a body of research, in addition to all kinds of circumstances beyond the control of researchers and universities. A good example of external circumstances benefitting favourably is a study on the use of back support belts to prevent injuries among the baggage handlers who load aircraft at Amsterdam's Schiphol Airport (Van Poppel *et al.*, 1998). This kind of work can cause back problems, but our study showed that back support belts were not effective in preventing these problems. "Negative" findings like this often make a study difficult to publish, but in this case we were fortunate. Our findings came out just as legislation in the United States to make back support belts compulsory was being drawn up. In record time, our study was printed in a leading journal and became the topic of heated debate in the media and during sessions at the US House of Representatives.

It would not be reasonable to expect every research project to have such a clearly identifiable social impact. But this example does show how scientific quality and societal relevance can go hand in hand. On the whole, I think there is at least a moderately positive correlation between quality and relevance. Of course, it should also be noted that bad research is never relevant. There should be no misunderstandings about that. In my view, quality and relevance are not interchangeable. I have no sympathy for attempts to boost societal relevance by making concessions on scientific quality.

Since universities are publicly funded, it is reasonable for society to expect something in return for this investment. First and foremost, this means that universities should train professionals who can make a difference. This is surely the best way to make our societal relevance clear to all. But it is research, not education, that I want to discuss here. At the very least, we should be entitled to demand that research has the potential to be relevant and lead to results which can be implemented. While this can be seen most clearly in applied research, I believe it applies also to fundamental research, although the relevance of the latter is much more difficult to predict and can sometimes take decades to emerge. In many cases it may not even materialise, but this should not undermine the researchers' good intentions in the first place.

Universities have to become transparent when it comes to the scientific quality and societal relevance of their research. In other words, are we doing the right things and are we doing them right? (Bensing *et al.*, 2003) The first question relates to taking up the challenge of society's problems. The second concerns both the quality of research and the relevance of its findings. Society

has the right to receive a clear answer to these questions. Besides, it is becoming increasingly difficult to make an impression by simply referring to the intrinsic importance of fundamental research.

Scientific quality

In many countries, the quality of scientific research is evaluated at fixed intervals and often the primary significance of such an evaluation concerns the policy of the university in question (Meta Evaluatie Commissie, 2007; Standard Evaluation Protocol, 2003). This is certainly true of the Netherlands. But there are cases where this evaluation process has far-reaching consequences for the budget allocated by the government. The United Kingdom's Research Assessment Exercise (RAE), which is held every four years, is a good example of this (HEFCE, 2008). This approach is still subject of much discussion, since the method used appears to work to the distinct disadvantage of interdisciplinary and applied research (Banatvala *et al*., 2005; Shewan and Coats, 2006).

The RAE also gives rise to strategic effects such as the temporary transfer of foreign colleagues who have an impressive list of publications to their credit. This is not the right way forward. However, I do believe that it is important to reward good behaviour and therefore I agree with those who advocate more dynamism in the funding of research (Commissie Dynamisering, 2006a, 2006b; Raad voor Medische Wetenschappen, 2005; Zuijdam, 2006). I also think that, within universities, there are good reasons for linking budget allocation to performance, at least to some extent. But any such allocation needs to be based on performance indicators which are simple to measure and difficult to manipulate.

Publications and citations are measurable aspects of scientific quality. The financial support obtained also makes a statement about the quality of the researcher and the research group. However, there are important cultural differences between disciplines as regards funding, publication and citations (Wouters, 1999). Citations are interesting because they show how great a contribution a specific publication has made to the acquisition of knowledge in a given field. The indicator which reflects the relative impact in comparison with the rest of the field is particularly informative (Moed, 2005; Van Raan, 1996; Moed *et al*., 1995). It is obtained by dividing a research group's average number of citations per article by the average number of citations of an article in the same field.

Table 1 shows a bibliometric analysis of the eight university medical centres in the Netherlands (CWTS, 2006). Apart from the VU University Medical Center, the identities of the various institutes have been concealed, but actually VUmc comes a close third behind Rotterdam and Utrecht. The final column shows that the Netherlands' university medical centres are cited 40% more often than the international average.

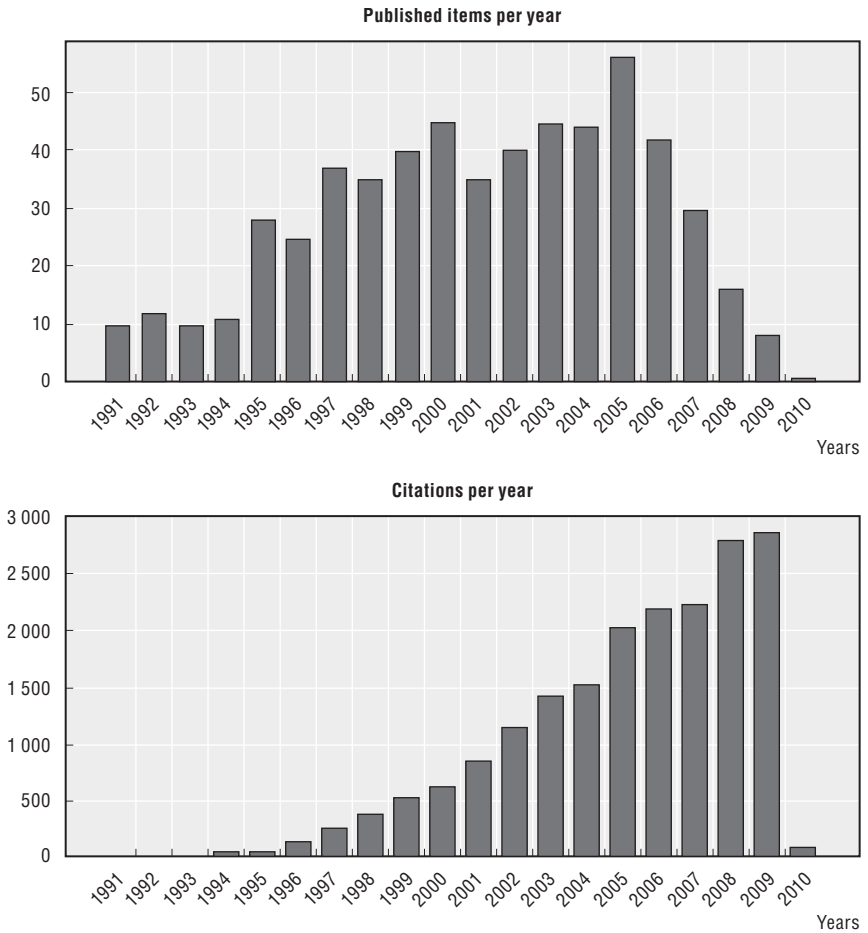
Table 1. **Output of eight University Medical Centres (UMCs) in the Netherlands, 1998-2005**

UMC	Publications	Citations per publication	CPP/FCSm
UMC a	9 034	15	1.59
UMC b	11 886	15	1.59
VUmc	7 711	14	1.52
UMC c	8 600	12	1.46
UMC d	8 937	15	1.44
UMC e	8 893	14	1.43
UMC f	9 023	12	1.27
UMC g	6 431	10	1.21
Total	**59 664**	**13**	**1.40**

The recently published Leiden Ranking shows that VU University Amsterdam and the VU University Medical Centre combined rate 35% above the world average when it comes to citation scores (CWTS, 2008; Council for the Humanities and Social Sciences Council, 2005). This puts us in 15th place in the European rankings and in fourth place among Dutch universities. While this is not bad, it also indicates that there is room for improvement. But even as we consider such matters, we must not lose sight of the limitations of this one-dimensional approach. As I see it, there are three.

First of all, this approach is largely dominated by the pure sciences. It does not work effectively for the arts and social sciences, where a different publication culture applies (Council for the Humanities and Social Sciences Council, 2005). Secondly, citation analyses rely heavily on past achievements. In most cases, between four and eight years pass between the start of a report and its publication. It then takes at least another year before the first citations start to appear. Third, only absolute citation scores are available for individual researchers, with all the disadvantages this entails. Nevertheless, the information in Figure 1 is increasingly being used when deciding on staff promotions to associate or full professor. It gives the number of publications and citations per year, the average number of citations per publication and the h-index. The significance of this last criterion is rising dramatically. The h-index is the number of articles which received h or more citations (Hirsch, 2005). In other words, if a researcher's articles are ranked in descending order on the number of citations, an h-index of 76 means that the 76 articles will have 76 citations or more. A singular index such as this is appealing in its simplicity but the strong influence of age and discipline means it is also potentially misleading. Researchers with the same h-index may also vary widely in terms of the number of publications to their credit.

Figure 1. **Publications and citations**



Notes: Results found: 581.
Sum of the times cited: 19 356.
Average citations per item: 33.31.
H-index: 76.

Source: ISI Web of Knowledge, *www.isiwebofknowledge.com*.

The discussion above makes it clear that finding good indicators of scientific quality is no easy task. Operationalising societal relevance is far more difficult still. But this does not strike me as grounds for giving up all attempts to achieve it.

The position of research

Research is becoming less and less the exclusive province of the universities (Wissema, 2005; Van Vught, 2004). Indeed, in both research and education,

there has been a dramatic rise in the level of international competition. Such developments call for decisive co-operation in a variety of changing contexts, and lead to the creation of new interdisciplinary scientific fields. In addition, governments are placing ever greater demands on quality, transparency and innovative ability. This requires a flexible, dynamic and entrepreneurial organisational structure. And all of this has to be combined with a greater, more readily demonstrable commitment to society.

Progressively, society has become more and more interested in science and academic endeavour, as evidenced by the 25th anniversary edition of the science section of leading Dutch newspaper *NRC Handelsblad* (NRC, 2007). It argues that innovation and new insights often occur unexpectedly and spontaneously. This suggests that there is good reason for having freedom and room for manoeuvre in scientific practice. I wholeheartedly concur with this position. Societal relevance is a difficult thing to predict. It is far simpler to assess impact in retrospect but even then, it can seldom be demonstrated beyond a reasonable doubt that a specific research project provided the missing piece of the puzzle (Oortwijn *et al*., 2007; AWT, 2005, 2007). The time frame also varies enormously: sometimes the societal impact of a study is readily apparent, but it often takes many years to make itself felt.

VU University Amsterdam has enjoyed a reputation for strong societal commitment since its foundation (Deursen, 2005). One manifestation of this commitment is the large number of political leaders among the university's former students. It can also be seen in the inspiring plans for a new campus in Zuidas, Amsterdam's dynamic new business district, in which the university's academic functions are interwoven with living, working and cultural activities. Academic citizenship and academic entrepreneurship are central to VU University Amsterdam's educational vision and its institutional development plan (Vrije Universiteit, 2007; Onderwijscentrum VU, 2006). This means that, in addition to excelling in scientific quality, it is also important to be outstanding in terms of societal relevance. It is therefore high time to make societal relevance a tangible entity, expressed in concrete terms. This is something we need to do for the outside world in the interests of accountability. But more importantly we need to do it for ourselves, to serve as a compass for our choices in the world of science.

Indicators of societal relevance

As I go on to discuss how indicators of societal relevance have been taking shape, I will make reference to health research, the sector with which I am most familiar. This familiarity stems partly from my involvement in commissions at the Royal Netherlands Academy for Arts and Sciences and the Council for Health Research, which focused on selecting indicators for this field of research

(Council for Medical Sciences, 2002; Raad voor Gezondheidsonderzoek, 2007). However, it is possible to imagine how my observations might be translated in terms of the arts, the exact sciences and the social sciences. Indeed, such translations are already available to some extent (Council for the Humanities and Social Sciences Council, 2005; AWT, 2007).

Clearly, the aims of health research ultimately lie in improving public health and healthcare. These should manifest themselves as improvements in people's life expectancy and quality of life. These are the outcomes that really matter. But, as we have already seen, their relationship with the research whose social relevance we want to evaluate is a complex one. It would therefore be unreasonable to come down too heavily on the researcher and his research for not fully realising the potential impact of the new knowledge obtained. Bearing this in mind, it is better to choose indicators at the product or process level, as shown in Table 2.

Table 2. **Indicators of societal relevance of scientific research**

Indicators of social value
Products:
• Specialised publications.
• Lay publications.
• Guideline or policy document.
• Service, method, technology.
Processes:
• Committees in professional or public domain.
• Public information services.
• Continuing education.
• Public opinion or political decision making.
Indicators of economic value
Products:
• Patents.
• Intellectual property.
• Start-up company.
Process:
• Committees in commercial domain.

Across the world, there have been many efforts to identify indicators of societal relevance (Bensing *et al.*, 2003; Council for the Humanities and Social Sciences Council, 2005; Council for Medical Sciences, 2002; Bensing and Oortwijn, 2006; Bouter and Knottnerus, 2000; Buxton *et al.*, 2000; Hanney *et al.*, 2004; Hicks, 2005; Kingwell *et al.*, 2006; Oortwijn *et al.*, 1998; Roper *et al.*, 2004; Spaapen *et al.*, 2007; UK Evaluation Forum, 2006; Wooding *et al.*, 2005). These efforts vary from very simple, expedient attempts to all-encompassing systems which turn the process of evaluation into a field of research in its own right. We have now acquired a considerable amount of experience in this area. Sometimes people

look almost exclusively at the economic value of the research (Ranga *et al.*, 2003; Blakemore and Davidson, 2006; Clairborne Johnston *et al.*, 2006). But many others consider this too limited (Van Oostrom, 2007). A number of institutes, including the NIVEL Institute for Health Services Research in Utrecht, the Netherlands, and our very own EMGO Institute, incorporate indicators of societal relevance in their annual reports (EMGO Institute, 2007; NIVEL, 2007).

As mentioned above, I distinguish between indicators that relate either to a product or a process (Table 2). Products are concrete and countable, and do not generally present too much of a problem when it comes to establishing the plausibility of the relationship between the research results and the unit evaluated. To establish indicators for processes is much harder.

As concerns products, it is a known fact that only 10% of Dutch doctors regularly read international scientific journals (Bouter and Knottnerus, 2000). For other professional groups, the situation is probably no different. If research results are to reach a group of professionals here in the Netherlands, **specialised publications** in Dutch, either in article or book form, will be an important channel of communication. Books and articles are also an effective way of spreading knowledge among the general public. Academics also find out more about areas related to their own field mainly by reading the science sections of the newspapers (Willems and Woudstra, 1993). But more and more, Internet is becoming their favourite medium.

Public relations and science communication are not part of the scientific researcher's core activities. But to ensure the quality of the **lay publications** provided, it is important that researchers are involved. Ideally, professional conduct should be based on weighing up all of the relevant and available scientific knowledge in a clear and balanced manner. This means that authoritative **guidelines and policy documents** constitute a suitable indicator. Sometimes scientific research leads to a new **service, method or technology**.

As concerns process indicators, I believe that researchers should also play an active role in the process of distributing and applying research results, although theirs will certainly not always be a leading role. The extent to which research groups fulfil this important role of distributing and applying scientific knowledge can be seen by looking at the membership of **committees in the professional or public domain**. Exactly which committees should be included in such a survey would have to be established for each discipline separately. Other concrete process indicators are the research-based contributions to **public information services** or the retraining and **continuing education** of professionals.

Sometimes research results make a noticeable contribution to **public opinion or political decision-making**. One illustration of this is the research that followed an outcry in the Dutch media and among Dutch MPs about the

fact that nursing home staff in Groningen had left elderly patients with dementia to die if they were no longer able to drink for themselves. The press was outraged. But after the case was researched, the headlines took a very different tone. It turned out that allowing patients to die by no longer administering fluids only took place after a painstaking decision-making process and led to a relatively peaceful death. The impact of a study will not always be this clear. Yet I still believe it is worthwhile to chart the media attention devoted to a given department or institute.

It is also important to know what the economic value of a scientific research project is, mostly in macroeconomic terms. In other words, we want to see how research contributes to the knowledge economy. In this domain too, a distinction can be made between indicators at product level and at process level. First of all we can assess the amount of **patents**, as well as the sale of **intellectual property**. However, it is also important to remember that, as a rule, universities tend to make their research results public and therefore accessible to all. As far as I am concerned, that should definitely remain so. Knowledge is a common good. Sometimes research can lead to a **start-up company**, and in its early years it is usually closely associated with the university. A prime example of this construction is the company that developed a new improved voting aid – the Electoral Compass – together with the Dutch national daily *Trouw*, winning a national journalism prize in the process. Recently the Electoral Compass was used for the US presidential elections. In such circumstances, the researchers are often part-time entrepreneurs. Business initiatives of this kind provide very tangible evidence of the economic value of research. However, there are also risks attached to this dual role of researcher and entrepreneur. Independence can become compromised, the conclusions can become distorted, and public resources can be appropriated for private gain. I believe we should be more forthright in debating this darker side of academic entrepreneurship. A suitable process indicator of economic value may be membership of relevant **committees in the commercial domain**.

Challenges

The evaluation of the societal relevance of scientific research is still in its infancy. There is still plenty of room for discussion about the validity of the indicators, the optimum level of detail and weighing up the relative importance of its various aspects. Indeed, it is essential that such discussion takes place. However, it is clearly still too early to adopt a strong quantitative approach.

Societal relevance should be the focus of attention at both the start and the end of the empirical cycle (Groot, 1961). The primary motivation of many academics is intellectual curiosity and of course there is nothing at all wrong with that. However, the process of selecting a research topic and formulating

a research question should always be accompanied by a reflection on the expectations in terms of relevance to society. This reflection may serve as a compass, guiding the choices yet to be made. At an early stage in their work, researchers should take into account the burning questions being asked by stakeholders, whether they be citizens, patients, companies or politicians. Of course, their freedom will be restricted by the available expertise and the resources available for research. But I am nevertheless convinced that researchers in this phase should have a clear ambition to carry out research that has real societal relevance. Once the project has been completed, it is the responsibility of the researchers to disseminate the results across the various scientific forums and, where relevant, among professionals, politicians and the general public.

Universities have to take the societal relevance of research seriously and report on it in terms of concrete performance indicators. Before this can be achieved, the indicators need to be developed in greater detail. I would therefore like to take this opportunity to invite the academic community to take up this challenge in a creative and constructive manner. A clearly identifiable focus on societal relevance also gives a powerful signal to students and young researchers. It shows that the academic process is not simply about chalking up publications and collecting citations. This message can be reinforced by rewarding researchers who have made a special contribution in this regard.

To resume, we should think more carefully about how we articulate societal issues in research that can provide useful answers. I believe that every researcher should not only ask himself "Am I doing things right?", but also "Am I doing the right things?". And above all, we should be asking "What does it mean anyway?".

The author:

Prof. Lex M. Bouter
Rector
VU University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
E-mail: lm.bouter@dienst.vu.nl

No specific funding source was called upon to support the research presented.

References

Adviesraad voor het Wetenschaps- en Technologiebeleid (2005), *De Waarde van Weten: De Economische Betekenis van Universitair Onderzoek, AWT-advies*, No. 62, AWT, Den Haag, pp. 1-44.

Adviesraad voor het Wetenschaps- en Technologiebeleid (2007), *Alfa en Gamma Stralen: Valorisatiebeleid voor de Alfa- en Gammawetenschappen, AWT-advies*, No. 70, AWT, Den Haag, pp. 1-54.

Banatvala, J., P. Bell and M. Symonds (2005), "The Research Assessment Exercise is Bad for UK Medicine", *Lancet*, Vol. 365, pp. 458-460.

Bensing, J. and W. Oortwijn (2006), "De maatschappelijke Waarde van Maatschappelijk Onderzoek", *Kwaliteit in Beeld*, Vol. 2, pp. 7-9.

Bensing, J.M. *et al.* (2003), "Doing the Right Thing and Doing it Right: Toward a Framework for Assessing the Policy Relevance of Health Services Research", *International Journal of Technology Assessment in Health Care*, Vol. 19, pp. 604-612.

Blakemore, C. and J. Davidson (2006), "Putting a Value on Medical Research", *Lancet*, Vol. 367, pp. 1293-1295.

Bouter, L.M. (1993), *Extramurale Epidemiologie: Gewoon en Bijzonder*, Vrije Universiteit, Amsterdam.

Bouter, L.M. and J.A. Knottnerus (2000), "Beoordeling van de Maatschappelijke Relevantie van Toegepast Gezondheidsonderzoek: Het Belang van Publiceren in Nationale Vaktijdschriften als Ruwe Indicator", *Nederlands Tijdschrift voor Geneeskunde*, Vol. 144, pp. 1178-1183.

Buxton, M. *et al.* (2000), "Assessing Benefits from Department of Health and National Health Services Research and Development", *Public Money and Management*, Vol. 20, October-December, pp. 29-34.

Centre for Science and Technology Studies (CWTS) (2006), *Bibliometric Study on Dutch Academic Medical Centers 1997-2005, Research Report to the Deans of Medical Sciences (DMW) in the Netherlands*, CWTS, Leiden, pp. 1-78.

Centre for Science and Technology Studies (CWTS) (2008), *www.cwts.nl/cwts/LeidenRankingWebSite.html*.

Claiborne Johnston, S. *et al.* (2006), "Effect of a US National Institutes of Health Programme of Clinical Trials on Public Health and Costs", *Lancet*, Vol. 367, pp. 1319-1327.

Commissie Dynamisering (2006), *Investeren in Dynamiek: Deel 1*, Den Haag, pp. 1-50.

Commissie Dynamisering (2006), *Investeren in Dynamiek: Deel 2*, Den Haag, pp. 1-112.

Council for Medical Sciences (2002), *The Societal Impact of Applied Health Research: Towards a Quality Assessment System*, KNAW, Amsterdam, pp. 1-51.

Council for the Humanities and Social Sciences Council (2005), *Judging Research on its Merits*, KNAW, Amsterdam, pp. 1-36.

Deursen, A.Th. van (2005), *Een Hoeksteen van Het Verzuild bestel: de Vrije Universiteit 1880-2005*, Uitgeverij Bert Bakker, Amsterdam, pp. 1-502.

EMGO Institute (2007), *Annual Report 2006*, EMGO, Amsterdam, pp. 1-248.

Groot, A.D. de (1961), *Methodologie: Grondslagen van Onderzoek en Denken in de Gedragswetenschappen*, Mouton, Den Haag, pp. 1-423.

Hanney, S. *et al.* (2004), "Proposed Methods for Reviewing the Outcomes of Health Research: The Impact of Funding by the UK's 'Arthritis Research Campaign'", *Health Research Policy and Systems*, Vol. 2, pp. 4.

Hicks, V. (2005), "Developing a CIHR Framework to Measure the Impact of Health Research", CIHR, Ottawa, pp. 1-4.

Higher Education Funding Council for England (HEFCE) *et al.* (2008), Research Assessment Exercise, *www.rae.ac.uk*.

Hill, A.B. (1952), "The Clinical Trial", *New England Journal of Medicine*, Vol. 247, pp. 113-119.

Hill, A.B. (1965), "The Environment and Disease: Association or Causation?", *Proceedings of the Royal Society of Medicine*, Vol. 58, pp. 295-300.

Hirsch, J.E. (2005), "An index to Quantify an Individual's Scientific Research Output", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, pp. 16569-16572.

Kingwell, B.A. *et al.*, for the National Health and Medical Research Council Evaluations and Outcomes Working Committee (2006), "Evaluation of NHMRC Funded Research Completed in 1992, 1997 and 2003: Gains in Knowledge, Health and Wealth", *Medical Journal of Australia*, Vol. 184, pp. 282-286.

Meta Evaluatie Commissie Kwaliteitszorg Wetenschappelijk Onderzoek (2007), *Trust, but Verify*, KNAW, Amsterdam, pp. 1-21.

Moed, H.F., R.E. de Bruin and Th.N. van Leeuwen (1995), "New Bibliometric Tools for the Assessment of National Research Performance: Database Description, Overview of Indicators and First Applications", *Scientometrics*, Vol. 33, pp. 381-422.

Moed, H.F. (2006), "Bibliometric Ranking of World Universities", CWTS, Leiden, pp. 1-40.

Moed, H.F. (2005), *Citation Analysis in Research Evaluation*, Springer, Dordrecht, pp. 1-346.

Netherlands Institute for Health Services Research (NIVEL), *NIVEL Jaarboek 200: De Gezondheidszorg Tegen Het Licht* (2007), NIVEL, Utrecht, pp. 1-107.

NRC Handelsblad (2007), *Wetenschap en Onderwijs 25 jaar: Jubileumnummer*, NRC, 24 February, Rotterdam, pp. 1-24.

Onderwijscentrum VU (2006), *Onderwijsvisie VU: Achtergronden, Meningen, Ervaringen*, Papyrus, Diemen, *www.onderwijscentrum.vu.nl*.

Oortwijn, W.J., H. Vondeling and L. Bouter (1998), "Use of Societal Criteria in Priority Setting for Health Technology Assessment in the Netherlands", *International Journal of Technology Assessment in Health Care*, Vol. 14, pp. 226-236.

Oortwijn, W.J. *et al.* (eds.) (2007), *The Returns from Health Care Efficiency Research in the Netherlands: Final Report of Project 945-15-001*, ZonMw, Den Haag, pp. 1-41.

Oostrom, F., Van (2007), *Markt en ziel: Jaarrede voor de Verenigde Vergadering van de Koninklijke Nederlandse Academie van Wetenschappen gehouden op 21 mei 2007*, KNAW, Amsterdam, pp. 1-16.

Poppel, M.N.M. van *et al.* (1998), "Lumbar Supports and Education for the Prevention of Low Back Pain in Industry: A Randomised Controlled Trial", *Journal of the American Medical Association*, Vol. 279, pp. 1789-1794.

Raad voor Gezondheidsonderzoek (2007), *Onderzoek dat Ertoe Doet: De Responsiviteit van Universitair Medische Centra op Vraagstukken in Volksgezondheid en Gezondheidszorg*, RGO, Den Haag, pp. 1-63.

Raad voor Medische Wetenschappen (2005), *Erop of Eronder: Financiering van (bio)Medisch Wetenschappelijk onderzoek*, KNAW, Amsterdam, pp. 1-58.

Raan Van, A.J.F. (1996), "Advanced Bibliometrics as Quantitative Core of Peer Review based Evaluation and Foresight Exercises", *Scientometrics*, Vol. 36, pp. 397-420.

Ranga, L.M., K. Debackere and N. von Tunzelmann (2003), "Entrepreneurial Universities and the Dynamics of Academic Knowledge Production: A Case Study of Basic Versus Applied Research in Belgium", *Scientometrics*, Vol. 58, pp. 301-320.

Roper, S., N. Hewitt-Dundas and J.H. Love (2004), "An *ex ante* Evaluation Framework for the Regional Benefits of Publicly Supported R&D Projects", *Research Policy*, Vol. 33, pp. 487-509.

Shewan, L.G. and J.S. Coats (2006), "The Research Quality Framework and its Implications for Health and Medical Research: Time to Take Stock?", *Medical Journal of Australia*, Vol. 184, pp. 463-466.

Spaapen, J., H. Dijstelbloem and F. Wamelink (2007), *Evaluating Research in Context: A Method for Comprehensive Assessment, Second Edition*, COS, Den Haag.

Standard Evaluation Protocol 2003-2009 for Public Research Organisations (2003), VSNU, NWO and KNAW, Den Haag, pp. 1-43.

UK Evaluation Forum (2006), *Medical Research: Assessing the Benefits to Society*, Academy of Medical Sciences, Medical Research Council and Wellcome Trust, London, pp. 1-46.

Vught, F. van (2004), *De Academische Reputatierace*, Universiteit Twente, Enschede, pp. 1-22.

Vrije Universiteit, Amsterdam (2007), *Instellingsplan VU 2007-2011*.

Wissema, J.G. (2005), *Technostarters en de Derde Generatie Universiteit*, DUP Satellite, Delft, pp. 1-26.

Wooding, S. *et al.* (2005), "Payback Arising from Research Funding: Evaluation of the Arthritis Research Campaign", *Rheumatology*, Vol. 44, pp. 1145-1156.

Wouters, P. (1999), "The Citation Culture", University of Amsterdam.

Willems, J. and E. Woudstra (1993), "The Use by Biologists and Engineers of Non-Specialist Information Sources and its Relation to their Social Involvement", *Scientometrics*, Vol. 28, pp. 205-216.

Zuijdam, F. (2006), "De Universiteiten Zitten in Een Spagaat", *Spui*, Vol. 22, pp. 4-5.

Information for authors

Contributions to the *Higher Education Management and Policy* Journal should be submitted in either English or French and all articles are received on the understanding that they have not appeared in print elsewhere. Articles submitted for publication in the Journal are refereed anonymously by peers.

Selection criteria

The Journal is primarily devoted to the needs of those involved with the administration and study of institutional management in higher education. Articles should be concerned, therefore, with issues bearing on the practical working and policy direction of higher education. Contributions should, however, go beyond mere description of what is, or prescription of what ought to be, although both descriptive and prescriptive accounts are acceptable if they offer generalisations of use in contexts beyond those being described. Whilst articles devoted to the development of theory for its own sake will normally find a place in other and more academically based journals, theoretical treatments of direct use to practitioners will be considered.

Other criteria include clarity of expression and thought. *Titles of articles should be as brief as possible.*

Presentation

Electronic submission is preferred. ***Three copies*** of each article should be sent if the article is submitted on paper only.

Length: should not exceed 15 pages (single spaced) including figures and references (about 5 000 words).

The first page: before the text itself should appear centred on the page in this order: the title of the article and the name(s), affiliation(s) and country/countries of the author(s).

Abstract: the main text should be preceded by *an abstract of 100 to 200 words* summarising the article.

Quotations: quotations over five lines long should be single-spaced and each line should be indented.

Footnotes: authors should avoid using footnotes and incorporate any explanatory material in the text itself. If notes cannot be avoided, they should be endnotes, at the end of the article.

Tables and illustrations: tabular material should bear a centred heading "Table". Presentations of non-tabular material should bear a centred heading "Figure". The source should always be cited. Please provide original source files.

Addresses of author(s), including e-mail, should be provided at the end of the article.

The funding source used to support the article should be provided at the end of the article.

References in the text: Vidal and Mora (2003) or Bleiklie *et al.* (2000) in the case of three or more authors.

At the end of the article, references should be listed in alphabetical order under the heading "References". Examples of the reference style used in the Journal are:

- For periodicals: Kogan, M. (2004), "Teaching and Research: Some Framework Issues", *Higher Education Management and Policy*, Vol. 16, No. 2, pp. 9-18.
- For books: Connell, H. (ed.) (2004), *University Research Management – Meeting the Institutional Challenge,* OECD Publishing, Paris.

The names of up to three authors can appear; for more than three authors just cite the first name followed by *et al.*

The covering letter

This should give full addresses and telephone numbers and, in the case of multi-authored papers, indicate the author to whom all correspondence should be sent.

Complimentary copies

Each author will receive two complimentary copies of the Journal issue in which his/her article appears.

Articles submitted for publication should be sent to the editor:

Professor Vin Massaro
Professorial Fellow
LH Martin Institute for Higher Education Leadership and Management
Melbourne Graduate School of Education, The University of Melbourne
153 Barry Street
Carlton, Victoria 3010, Australia
E-mail contact: imhe@oecd.org

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(89 2010 01 1 P) ISSN 1682-3451 – No. 57217 2010

Higher Education Management and Policy

Journal of the Programme on Institutional Management in Higher Education

The state of learning outcomes assessment in the United States
George D. Kuh and Peter T. Ewell 9

Defining and monitoring academic standards in Australian higher education
Hamish Coates 29

Women in science: the persistence of gender in Australia
Sharon Bell 47

Quality assurance in higher education as a political process
Michael L. Skolnik 67

System accreditation: an innovative approach to assure and develop the quality of study programmes in Germany
Tanja Grendel and Christoph Rosenbusch 87

The intended and unintended effects of the Bologna reforms
Sybille Reichert 99

Knowledge as a common good: the societal relevance of scientific research
Lex M. Bouter 119

Subscribers to this printed periodical are entitled to free online access. If you do not yet have online access via your institution's network, contact your librarian or, if you subscribe personally, send an e-mail to **SourceOECD@oecd.org**.

Volume 22/1
2010

ISSN 1682-3451
2010 SUBSCRIPTION
(3 ISSUES)

OECD*publishing*
www.oecd.org/publishing

imhe

89 2010 01 1 P

