

Research quality, peer review and performance indicators

Penelope S. Murphy
University of Wollongong

Approaches to research assessment vary in international context

Approaches to assessing research quality, and their respective reliance on peer review and performance indicators, vary distinctly between different national contexts. Such differences of approach are related to differences in the characteristics of the national higher education systems with which they are associated, and to the broader political and social environments within which those systems are embedded. In Australia, with the advent in 1993 of the Committee for Quality Assurance in Higher Education, the assessment of research quality has been integrated, at the institutional level, with the assessment of higher education quality as a whole. The Australian approach to higher education quality assessment represented by the Committee for Quality Assurance manifests, on one hand, some distinctly different characteristics from the current British approach, while on the other, the Australian and British approaches share some strong similarities which distinguish them from other OECD countries and from the USA. International comparisons are informative in enabling an appreciation of the distinctive features of our own national system, and providing a basis for reflecting upon the implications of those features.

Characteristics of quality assessment in different national systems: Market, simulated market and non-competitive approaches

One of the characteristic differences in higher education systems in various national contexts is the extent to which they are based upon market relationships. A prime example of market-based higher education is that of the USA. The dominance of the market in American higher education was given a boost in 1972 by a decision to the money in the hands of students so that institutions would have to compete for students. The main role of Federal government in higher education is to provide financial assistance to students to remove financial barriers, enhancing equality of opportunity. Despite this, the basis for the competition for students has tended to be reputation, appearances and original features, rather than the quality of the student experience.¹ This demonstrates that although market control is probably more pronounced in the USA, it is by no means operating perfectly. In this system responsibility for quality assurance is shared by a variety of organisations which have evolved independently, operating at a variety of levels and commanding varying degrees of respect from institutions on the one hand and policy-makers on the other. The situation, according to Marchese, is very dynamic. For instance, institutional level accreditation, which was one of the earlier forms of quality assurance and has existed for decades, was under review in late January 1994, with a real possibility of its being replaced by some other form.

By contrast with the USA, in Britain and Australia, a much more conspicuous role is played by central government. In both countries government plays a significant role in contriving simulated market conditions for its higher education system, actively promoting competition between institutions in their simulated market. It is assumed that competition between institutions (*inter alia*) will serve

to enhance standards. Competition between institutions focuses attention upon the *institution's* role in the management of its activities, as distinct from the role of organisational units or individuals within the institution. Government intervention in promoting and emphasising institutional responsibility for management has given rise to a perception amongst some academic staff of a growing tendency towards 'managerialism' in universities (Marginson 1993).

In Britain the emphasis on inter-institutional competitiveness is reflected in the changed funding arrangements introduced by the Further and Higher Education Act of 1992, as well as a series of other changes introduced in the 1990s. There is, however, a division of powers in the assurance of higher education quality between government and the institutions. On the one hand, the Higher Education Funding Councils (which are responsible to government) conduct evaluation at the subject level between institutions. On the other, the Higher Education Quality Council (which is an institutionally funded body) conducts quality audits at the institutional level. Each of these organisations makes institutional visits and communicates with one another at least on some matters. But the two systems are not currently fully integrated. The assessments made by the Funding Councils have a direct bearing on the funding of institutions, whereas those of the Higher Education Quality Council only have an indirect one. The Research Assessment Exercise is conducted under the auspices of the Higher Education Funding Councils.

In Australia the White Paper on Higher Education in 1988 (Australia 1988) introduced policies aimed at promoting competitiveness between higher education institutions. Most particularly these were the abolition of the binary distinction between the universities and the former college sector, and their incorporation into the 'Unified National System' of higher education within which they were to compete for funding. In the process of preparing this White Paper the Government signalled its intentions to fund universities on the basis of performance indicators (White Paper), and commissioned research to develop the necessary indicators (Department of Employment, Education and Training, 1992). Although the government has since softened its position in this regard, the measurement of higher education performance has been given new life through being subsumed into the broader process of assessing quality. The promotion of competitiveness flowed through into research policy in 1989 with the Ministerial Statement (Australia 1989) which announced a policy of increased selectivity and concentration in research funding. Such a policy was, at least in part, a corollary of the abolition of the binary distinction, which the preceding year had more than doubled the potential pool of applicants for research funding.

The British and Australian approaches, with their governmental promotion of inter-institutional competition, are in vivid contrast to approaches to quality assessment in European countries. Concern for quality assessment in higher education is currently shown mainly by the northern countries in Europe. In these countries the issue is not primarily one of competition between individual institutions, but how to ensure the efficiency of the system as a whole (Goedegebuure *et al.* (eds.) 1990, p.16), as in Norway and Sweden. In others, concern is currently not so much market competitiveness, but rather internationalisation of higher education, of ensuring comparability of standards, reciprocity in recognising qualifications, and facilitating

the mobility of personnel amongst European Union (EU) and aspiring EU countries. Correspondingly, the role of the institution itself in the management of the system has received much less stress.

In summary, it appears that in a number of EU countries (excluding Britain), higher education quality assessment is envisaged as part of a broader process of integration of the EU as a political-economic unit. In Britain and Australia central government envisages higher education as playing a key role in positioning the nation in a world economic market, and is actively engaging in funding-based policies aimed at promoting its interaction with higher education institutions to that effect. In the USA the market exists *de facto*, and the more powerful users organise quality assessment systems to meet their own specific needs. Some of the distinctive characteristics of the British/Australian model emerge as a more comprehensive, integrated approach, more directly linked to funding, and in which Government and performance indicators play a significant role.

We may set these observations into a theoretical context as follows. According to Dill (1992), three successive models of higher education quality control are associated with the transition from elite to mass to universal higher education. They are epitomised as: clan (or collective, collegial) control, superseded by bureaucratic control (relying upon imposition of rules and regulations), superseded by market control (relying upon open competition amongst institutions for students, staff, resources and adaptiveness to new programs desired by the public).

Market control is discernible in the USA in a way that is perhaps yet hidden from other national systems as they have yet to achieve universal higher education. In some systems, particularly that of Australia, market control is also hidden by virtue of the traditional role of government. The elite status of universities within higher education systems has, as a general rule, been slower to erode in Britain and Australia than in the USA. In both Britain and Australia it was preserved by the 'binary' distinction. In Australia the White Paper of 1988 propelled Australian higher education forward in the mass stage, while in Britain the Further and Higher Education Act of 1992 had a similar effect. A corollary of 'massifying' higher education is an increased concern for its quality: both over variability within the much enlarged group of institutions which are now called universities, and over the threat to standards which rapid expansion is sometimes thought to pose. Australian higher education has undergone a period of rapid expansion, although there is some recent evidence of a change in that trend. Britain is planning for continued expansion, anticipating that by the year 2000 about one in three of all 18-19 year olds will enter higher education (Britain 1991, p.10). Some European countries such as the Netherlands and Germany have contained the expansion of their university sector by preserving differentiation between different types of institutions.

'Quality', in the evaluation of science, is not an objective characteristic of a group or an individual, but depends, as van der Meulen (1992, p. 39) has shown, upon the goals set by actors. These groups of actors constitute different frames of reference and therefore presuppose different concepts about what might constitute quality. Similarly, the different models of quality assessment which may be observed in different national contexts may be understood as constituting such different frames of reference. This means that we should look for the different values underlying the different models of quality assessment which may be observed in different national contexts.

If we understand each of these quality control models as an example of a different frame of reference for 'quality', we may observe that in clan control, the dominant values are those of the clan members, i.e. those of the academic peers, who form the 'clan'. In bureaucratic control, the dominant values belong to those who dominate the hierarchy. In market control a much larger role in determining what constitutes quality is played by the values of the many constituencies who demand higher education or its products. In higher education systems where government plays a significant

role in managing the assessment of quality (as in Britain and Australia), we may expect that the dominant values in the bureaucratic control model will be at least partly if not predominantly governmental values.

The significance of performance indicators as compared to peer review

The emergence and acceptance of performance indicators in general has occurred within a broader international, political trend. In most European countries (including Britain) conservative governments rose to power in the 1970s and 1980s (Maassen and van Vught, 1988, cited in Goedegebuure *et al.* (eds.) 1990, p. 15). The end of more or less unconditional funding of public higher education is attributed (Goedegebuure *et al.* (eds.) 1990, p. 15) to the 'value-for-money' approach of these governments, implying that public funding of higher education was increasingly becoming linked to the performance in higher education institutions. Despite being labelled as 'Labor', the last five successive terms of government in Australia have placed similar emphasis on value for money and performance measurement, consistent with the approaches of conservative governments in Europe.

Performance indicators need to be understood as 'the expression *par excellence* of a method of managing higher education' (Sizer *et al.* 1992, p. 133). Two key concepts underlie differences in the role of performance indicators in different national settings: on the one hand, the pursuit of 'equivalence' within the system, and on the other, the pursuit of greater 'variety'. European countries fall into both groups, with Norway and Sweden corresponding to the former, and the Netherlands and Britain forming examples of the latter. Denmark, according to Sizer *et al.* (1992), is in transition from the former to the latter. Systems which seek to establish equal opportunities for students and provide programmes or courses of equivalent quality in all institutions reflect the 'equivalence' concept. Those which recognise and promote the differences of needs of society reflect the 'variety' concept. In the past, Australian policies in higher education have tended to reflect the former more than the latter, but increasingly there have been manifestations of the latter in quality assurance policy, and particularly in research policy.

In pursuit of greater variety, governments seek to intervene more selectively or design mechanisms in which selectivity is a key feature (Sizer *et al.* 1992, p. 133). Furthermore

Where a relationship is involved by which a higher level within the system makes resources available, the desire to develop performance indicators will always originate in value for money considerations. The desire will be to allocate resources more selectively or to promote greater selectiveness in their allocation (Sizer et al. 1992, p. 138).

Where a system of quality control is based on the concept of variety and is directed at generating comparative quality ratings, governments will have a pressing need for performance indicators for rationalisation and resource allocation (Sizer et al. 1992, p. 142).

The distinction between management statistics on the one hand, and performance indicators on the other, rests principally on the use to which information is to be put (Sizer *et al.* 1992, p. 145-146). Such uses are potentially many and various, differing also at government and at institutional levels. The context and the mechanisms for use are therefore of critical importance. They are, however, often not made explicit or specified in any detail. Systems which subordinate the use of performance indicators to a peer review process provide such a contextual setting. Conversely, where indicators are unconstrained by peer review processes they are, in principle, available for use for whatever purposes may prove convenient.

Thus a quality assessment system which manifests 'peer review' as a dominant feature may be understood as one reflecting the prevalence of academic values, while one in which performance indicators are unconstrained by peer review may be understood as one reflecting the prevalence of non-academic values. Where peer

review is subsumed into the indicators, the system may be understood as a hybrid of the two sets of values, though tending (at least potentially) towards the non-academic.

The significance of the difference between peer review and performance indicators has been well appreciated by Goedegebuure *et al.* (eds.) (1990) who contrasted the cases of quality assessment in Dutch and British higher education by reference to the salience of peer review in the former and performance indicators in the latter. In the early 1980s governments in both countries expressed interest in the use of performance indicators as well as in a peer review approach, but after 1985 'in Dutch higher education the use of peer judgements was emphasised, while in the UK the use of performance indicators seemed to dominate the quality assessment efforts' (Goedegebuure *et al.* (eds.) 1990, p. 17). The contrast between these two approaches attracted a lot of interest from the higher education sectors of other European countries, and led to their use in the study by Goedegebuure *et al.* as exemplars of different approaches.

Sizer *et al.* drew a number of conclusions from their analysis, summarised by Yorke (1993) as follows:

- Government needs to discuss and agree with institutions how its policies will be implemented, how performance will be monitored, and the uses to which the evidence from monitoring are to be put.
- There is a need jointly to identify the indicators that are to be used, and to make clear their purposes and limitations.

Assessing research quality

The major national (and even continental) differences in approaches to higher education quality assessment, illustrated above, are also reflected in approaches to assessing research quality in different national contexts. The salience of peer review as compared to performance indicators similarly varies.

In the USA there are two main methods by which assessment of university research takes place. First, there is the assessment by research councils and other bodies of proposals for externally funded research. Although value judgements involving peer review are made of the potential value of the research, there are really at least two elements: the brilliance of the ideas and the proposal itself; and the track record of the researcher/research teams (usually groups). Second, research assessment may form part of program review. This encompasses research, teaching and community services, is undertaken at state level and also involves peer review. Program review is the main mechanism for assessing the productivity of a unit (as distinct from a researcher or team that does not conform with a unit). There is no direct link to funding in Program Review, and there is no formula linking these two kinds of assessment. In addition to these two methods, a very small number of universities have assessment of some individual staff at least on an infrequent basis.²

Another method undertaken intermittently, however, is the rating method developed by Allan M. Cartter for assessing the quality of graduate education. This method, known popularly as the 'Cartter Ratings', contributes supplementary information to the public domain which may or may not be used in assessments of research quality. It has no formulaic link to funding, and depends upon an analysis of judgements made by active scholars in each of 29 academic fields of study of the quality of graduate staff in the corresponding fields in other institutions (El-Khawas 1993).

In northern Europe, although there have been some government initiatives to promote a closer link between funding and quality assessment, there has been a fairly effective counter-tendency by institutions or their representative bodies. In the Netherlands, for example, a new system for quality 'care' was trialed in 1993. This quality 'care' is seen as serving two goals: quality maintenance and improvement through feedback to the research group; and management on the basis of quality through provision of quality assessments to executive boards of the faculties and universities. The fact that government can utilize the assessments to ascertain scientific potential and underwrite government policy is seen as a secondary effect

(Vereniging van Samenwerkende Nederlandse Universiteiten 1993). The new system relies heavily upon peer review and integrates research with teaching in a program review approach. It supersedes the former assessment of research in the context of Conditional Financing initiated in 1983 which distinguished research funding from general university funding. The Conditional Financing System introduced the requirement for research programmes to be approved by external bodies with a view to distinguishing excellent research from good research (Goedegebuure *et al.*, 1990, p. 25-26). The new system therefore represents a shift away from the more direct link between funding and assessment presupposed by its predecessor. Similarly, the Committee of Rectors of European universities (CRE) has taken the initiative to develop international institutional audits, and in so doing the rectors as a group are asserting the institutions' role in assessing quality and the separation of that assessment from direct funding considerations (van Vught and Westerheijden, 1993).

In Britain there have been a series of Research Assessment Exercises, the third and most recent of which was conducted for the Higher Education Funding Councils under the auspices of the former Universities Funding Council. These Exercises have been designed to enable the comparison of research performance between departments in different universities and ranking for funding purposes. They therefore constitute a powerful mechanism for promoting institutional competitiveness. Moreover the ratings generated by the Exercises have evolved into *de facto* indicators of performance, which in turn have been perceived by some as providing a stimulus to improve standards. The Exercises were systematic, relying heavily upon research output indicators, and involved a direct link with funding. The first of these Exercises was controversial, but successive exercises have gained greater acceptance through modification so as to incorporate a much greater input of peer review.

Interestingly, the White Paper 'Higher Education: A New Framework' opens its chapter on Research (chapter 3) with the statement that

the quality of research in the United Kingdom has achieved world-wide recognition. The Government remains committed to maintaining an internationally competitive research base through higher education institutions and the Research Councils. (Britain 1991)

This proposition implies that the excellence of British research is unquestionable and, counter to the tendency in some other European countries, suggests denial of a necessity to further demonstrate international acceptability. It also implies that the purpose of the Research Assessment Exercise is to decide which of the excellent research to fund, rather than to evaluate its quality.

In Australia the main method by which university research has been assessed has been through peer review of applications for research grants to funding bodies from individuals or research groups. Following the White Paper of 1988 the Relative Funding Model for universities was introduced which involved the formulaic redistribution of the research-related quantum of operating grants. The indicator (now known as the National Research Grants Index) used in the formula was the relative performance of each institution in winning competitive peer reviewed research grants from a specified selection of granting bodies. This Index was subsequently used for the formulaic distribution of research infrastructure funding from the Australian Research Council through the former Infrastructure Mechanism A grants scheme. Thus not only individuals and research groups but also whole institutions have been cast into a series of interrelated competitive relationships through the incorporation of grant-winning capacity into the systematic distribution of funding between institutions. Many institutions have mirrored this arrangement for the internal distribution of their discretionary research funds, some of which come to the institution through the process described (Harrold, R. (1992), Murphy and Hill, (1992)).

The emphasis on performance indicators in Australian higher education management at the system level, which we have described above, is also observable in research assessment. In the 1990s the Australian Research Council has sought to develop indicators of

research activity and quality, based less upon input and more upon research output. A Research Performance Indicators Survey (National Board of Employment, Education and Training 1993) identified seven indicators 'preferred' by the research community surveyed. A follow-up study in 1993 examined the quality assurance aspects of such quantitative indicators, focusing in particular on assuring the rigour of the peer review aspects embodied in those indicators (National Board of Employment, Education and Training, forthcoming). Such indicators are about to be implemented formally at the system level. It is envisaged that they may be used both at the level of assessment of grant applications and at the level of distribution of funding both within and between institutions, as well as other levels otherwise unforeseen.

The Committee for Quality Assurance, appointed by the Minister for Employment, Education and Training in 1993, has been superimposed on the assessment methods described above. In this new policy instrument, research forms an integral part of the institution's quality management responsibility. Compared to Britain, this represents a more coherent, integrated approach, for there is no division between government and institutions respectively in terms of responsibility for subject level assessment, on the one hand, and institution-level assessment on the other. In its first round, conducted under considerable time constraints, the assessments in general (and especially those of research) made by the Committee for Quality Assurance were necessarily quite superficial. In 1995 the Committee will, however, focus specifically on research. One possible use for the 'quality-assured' quantitative indicators discussed above might be in the conduct of the 1995 research assessment. The Committee for Quality Assurance, as a policy instrument, represents a direct link with funding even though it can be argued that the funding concerned does not form part of the institutions' base operating grants. As the announced policy of the former Minister Beazley was to limit distribution of the funds to not more than half the institutions, the Committee has inevitably acted as an instrument for promoting intense inter-institutional competition. The subsequent removal of that limitation by Beazley's successor, Minister Crean cannot in itself erase that effect.

In summary, it appears that government plays a more prominent role in British and especially Australian research quality assessment, and that much greater reliance is placed in these two countries upon standardised system-level performance indicators than in either the USA or northern Europe. The emphasis in Britain and Australia seems to be less upon evaluating the quality of research *per se*, than upon developing rationales for funding decisions. This is accounted for by the salience of funding as a governmental concern in these two systems. Such concern is readily comprehended in Australia, where there is very little funding derived from sources other than government, and where funding is essentially the only policy instrument available to the federal government in higher education policy. Compared to the other systems described, the Australian system seems to be the most exaggerated in terms of its integration, coherence, and reliance upon systematic performance measurement. The relatively small size of the Australian higher education system makes it feasible for such system-level indicators to be constructed and made available.

Some unanswered questions

It seems that in Australia responsibility for assessing the quality of research has been captured by government in a more coherent and systematic way than in many other comparable polities. Thus there is a more urgent need here than elsewhere for, following Sizer *et al.* and Yorke, 'government, itself and through its agencies, to engage the higher education institutions in productive discussion about aims, objectives and desired outcomes' with a view to achieving not so much agreement but 'a shared understanding of what the system is expected to achieve, and the criteria against which achievement may be assessed' (Yorke 1993).

Perceptions may differ as to the value of performance indicators

and research evaluation systems linked directly to funding (i.e. the value of coherence and government influence on universities). There is, however, very strong pressure in this direction in Australia and it is reflected in much broader government economic policies in both Britain and Australia.

The pressure for the institution to assume a role in assessing its own research function is already apparent in research management policies in most Australian universities (Murphy and Hill, 1992), and gradually it is being manifest in more systematic procedures for the collection of research data at the institutional level. Thus we may ask, what will happen to research if and when, under this pressure, responsibility for its quality management is captured by the institution? Will 'science' elude the university and move elsewhere, so as not to be subordinated to institutional management interests? Shapiro (1993) has proposed that 'the individual institution - at least in so far as it is research oriented - is less than previously the appropriate unit for allocation of costs'. This, he has argued, is because in the relatively advanced technological societies of the OECD countries,

the central issue of information is no longer ownership but access, and it is by no means clear how to either calculate or allocate the costs incurred in developing the information in the first instance. This is particularly a problem in research where (a) inter-institutional and inter-jurisdictional consortia are becoming so common a feature of frontier-level work and (b) we have yet to resolve at what prices basic research findings should cross national borders (Shapiro, 1993).

Another important question to address is what role, if any, can the assessment of research quality play in the management of institutional quality? This might require knowing what is meant by 'research assessment', variously, for the researcher, the research group, the organisational unit, the institution and the funding body. It might also require deepening our understanding of the role that research plays in the quality of the institution as a whole. It might require distinguishing between the quality of the research itself, and the management of the research process. The 'micro' end of the system, in Yorke's terms, at which quality enhancement may be effected, may require a different kind of performance data, which may be insufficiently robust for most externally-oriented purposes (Yorke 1993).

Can Australia afford, as in Britain, to assume that all its research is excellent, and that research management equates with the management of research funding? Or is there a place for Australian universities and, in particular, their researchers to consider how they might obtain better value for the research funding that is attracted? Perhaps, most importantly, institutions might need to consider how the quality-assured, quantitative indicators of research which will soon become available can be put to best use for the purposes of both the institution and its researchers, and what other approaches to assessment might be valuable within and for the institution itself.

References:

- Australia 1988, 'Higher Education: a Policy Statement', circulated by the Hon. J.S. Dawkins MP, Minister for Employment, Education and Training, Australian Government Publishing Service, Canberra, ACT.
- Australia 1989, 'Research for Australia: Higher Education's Contribution.' Statement circulated by the Hon. J.S. Dawkins, Minister for Employment, Education and Training, May 1989. Australian Government Publishing Service, Canberra.
- Australia 1993, 'Quality Assurance Program 1993 Guidelines for Higher Education Institutions.' Committee for Quality Assurance, Ministry for Employment, Education and Training, July 1993.
- Britain 1991, 'Higher Education. A New Framework'. White Paper. Department of Employment, Education and Training 1991, *Performance Indicators in Higher Education. Report of a Trial Evaluation Study*, (Chair: R.D. Linke), Australian Government Publishing Service, Canberra.
- Department of Employment, Education and Training 1993a, *National Report on Australia's Higher Education Sector*, Australian Government Publishing Service, Canberra.

Department of Employment, Education and Training 1993b, 'Recent Trends and Current Issues in Australian Higher Education.' Country Paper prepared for the International Conference on the Transition from Elite to Mass Higher Education, Sydney, 15 - 18 June.

Dill, D.D. 1992, *Quality by Design: Toward a Framework for Academic Quality Management. Higher Education: Handbook of Theory and Research*, Vol. VIII, Agathon Press, 5648 Riverdale Avenue, Bronx, NY 10471.

El-Khawas, E. 1993, 'Assessing Research Quality: Another Approach'. Paper presented to the Annual Conference of the Society for Research into Higher Education, University of Sussex, Brighton, 14-16 December.

Goedegebuure, L.C.J., Maassen, P.A.M., & Westerheijden, D.F. (eds.) 1990, *Peer Review and Performance Indicators. Quality assessment in British and Dutch higher education*. Uitgeverij Lemma B.V., Utrecht.

Harrold, R. 1992, 'Evolution of Higher Education Finance in Australia'. *Higher Education Quarterly*, 46:4, Autumn, pp. 321 - 337.

Harvey, L., Burrows, A. & Green, D. 1992, 'Criteria of Quality'. The QHE Project, Baker Building, The University of Central England In Birmingham, Perry Barr, Birmingham B42 2SU.

Higher Education Council 1992, *Achieving Quality*. National Board of Employment, Education and Training, Australian Government Publishing Service, Canberra, ACT 2600.

Marginson, S. 1993, *Education and Public Policy in Australia*. Cambridge University Press.

Murphy, P.S. and Hill, H. 1992, 'How Competitive Are Australian Universities?' Centre for Research Policy Report No. 4, University of Wollongong, March.

National Board of Employment, Education and Training 1993, Research Performance Indicators Survey. *Commissioned Report No. 21*, Australian Government Publishing Service, Canberra.

National Board of Employment, Education and Training (forthcoming), Quantitative Indicators of Australian Academic Research. *Commissioned Report No. 27*, Australian Government Publishing Service, Canberra.

Organisation for Economic Cooperation and Development (OECD) 1993, Background Report for the International Conference on the Transition from Elite to Mass Higher Education, Sydney, 15 - 18 June.

Shapiro, B.J. 1993, 'Mass Higher Education. Problems and Challenges.' Keynote Address, International Conference on the Transition from Elite to Mass Higher Education, Sydney, 15 - 18 June.

Sizer, J., Spee, A. & Bormans, R. 1992, 'The role of performance indicators in higher education', *Higher Education*, 24: 133-155.

van der Meulen, Barend J.R. 1992, *Evaluation Processes in Science. The Construction of Quality by Science, Industry and Government*. CIP-Gegevens Koninklijke Bibliotheek, Den Haag, Netherlands.

van Vught, F.A. and Westerheijden, D.F. 1992, 'Quality Management and Quality Assurance in European Higher Education: Methods and Mechanisms', Center for Higher Education Policy Studies, University of Twente, Enschede.

van Vught, F.A. and Westerheijden, D.F. 1993, 'European Rectors' Conference - Pilot Studies on Institutional Quality Audits'. Paper presented at IMHE Seminar on Quality Management and Quality Assurance in Higher Education, OECD, Paris 6 - 8 December.

Vereniging van Samenwerkende Nederlandse Universiteiten 1993, *Quality Assessment of Research - Protocol 1993*, VSNU, Leidseveer 35, Postbus 19270, 3501 DG Utrecht, Netherlands.

Yorke, M. 1993, 'Siamese Twins? Performance Indicators in the Service of Accountability and Enhancement'. Keynote address to the Annual Conference of the Society for Research in Higher Education, University of Sussex, 14 - 16 December.

Do performance indicators measure outcomes of education?

Lindsay D. Mackay
Deakin University

Introduction: The context

The current focus in higher education on quality, and the current interest in using performance indicators are reflections of the context in which higher education is operating.

That context can be characterised in terms of a period of significant change and adjustment. Universities have been challenged by a series of dilemmas. It has been a period of significantly increased participation in the final year of secondary education leading to increased demand for higher education from school leavers. In parallel there has been the adoption of more flexible entry provisions to address issues of equity and access, leading to increased demand for access to higher education from groups other than school leavers. And the prevailing economic conditions have reduced the availability of alternatives to higher education. They have also changed patterns of demand for courses. The current downturn in demand for Education courses is but one example.

This increase in demand and the resulting increase in higher education places has occurred in an environment of general economic constraints, reducing resources and increasing competition for scarce resources. Partly as a result of these resource constraints, and partly also because of the perceived slowness of change in universities, there have been persistent and increasing calls for improved efficiency and public accountability in all aspects of higher education.

This has been reflected in pressure for the development of more objective and systematic procedures for the evaluation of universities, and for systematic monitoring of the performance of universities and the higher education system as a whole.

There has been accelerating change toward more formal, routine and quantitative approaches to evaluation of higher education. The focus has been on monitoring performance and productivity to assist institutions to improve their efficiency and accountability.

Some key developments

It is important to acknowledge some of the recent milestones which have influenced the way in which performance indicators are currently viewed.

1. The Review of Efficiency and Effectiveness in Higher Education (CTEC, 1986) marked a significant advance in government pressure for visible progress in establishing systematic mechanisms for performance appraisal of higher education institutions. To that time such performance appraisal had been largely at the discretion of the institution, intermittent, predominantly subjective, confined essentially to single institutions with inadequately defined central data systems giving little basis for comparability.

This Review encouraged the further development and application of performance indicators both within institutions and across the higher education system to assist in identifying strengths and weaknesses and in measuring progress toward specified institutional and systemic objectives.

2. The Discipline Review of Engineering (Williams 1988) which reported in 1988 systematically attempted to define and apply a range of input and performance indicators for higher education institutions, which to some degree covered the major functions of

higher education in professional discipline areas.

3. The Green Paper (Commonwealth of Australia 1987) marked a strengthening of government policy on institutional evaluation and performance appraisal. That paper proposed in specific terms that institutional performance in its various forms be quantitatively assessed and that this assessment should have some influence on institutional funding. The paper explicitly stated the Government's intention 'to fund on output and performance'. Importantly, there was an acceptance that performance indicators needed to be acceptable to both institutions and the Government.

The Government expects institutions, as part of their strategic planning, to give consideration to indicators that would help in measuring the achievement of their goals. In time, such indicators that are agreed to be useful would be built into the general funding system.

The need for performance indicators acceptable to both institutions and the Government points to the likelihood of further demands on the higher education statistical information base. (Commonwealth of Australia 1987, p.42)

4. The AVCC/ACDP Working Party on Performance Indicators was jointly established by AVCC and ACDP in response to these proposals in the Green Paper. This joint working party attempted to define what they saw as 'a proper context' in which performance indicators should be used.

The joint working party sounded some warnings to those tempted to enthusiastically embrace performance indicators:

Something resembling a Cargo Cult seems to have grown up around the notion of performance indicators, so that all manner of powers and virtues are ascribed to them and expectations are aroused that by collecting and using them great benefits will miraculously result. (AVCC/ACDP 1988, p.1)

In their report of December 1988, they proposed that 'a proper context' for the use of performance indicators be established by a policy agreed by institutions:

the best manner of assessing performance is by the judgement of knowledgeable and independent people, and that

an agreed set of indicators will form part of the material available to those conducting the reviews. (AVCC/ACDP 1988, p.4)

They went on to try to identify a range of possible indicators which would be acceptable to both institutions and the Government. They proposed that, in selecting performance indicators among the things that should be kept in mind were:

- indicators should be clearly related to the institution's prime functions and objectives
- indicators should cover as many of the institution's prime functions and objectives as possible
- indicators should form a coherent set
- indicators are of greater significance when considered in groups than singly
- as few indicators as possible should be used