Evaluation for what?

Lee Harvey

Abstract

External quality monitoring of higher education is explored and analysed. The different types of external bodies and their modus operandi are reviewed. The reasons for undertaking the evaluation are also examined in detail. It is argued that external evaluation legitimates the *status quo* and fails to ask substantial questions about the nature of learning. External quality monitoring is preoccupied with method and has almost entirely ignored a quarter of a century of research into learning theory, the nature and styles of learning and classroom innovations. It is suggested that higher education monitoring agencies need to adopt an explicit 'transformation' approach to quality and to address the implications of that for student learning. Only then will they shift from accountability- and compliance-oriented agencies to ones that raise substantial questions about improving student learning.

Introduction

Higher education is increasingly subject to evaluation by external quality monitoring bodies. This paper, taking its cue from Robert Lynd's (1939) *Knowledge for What?,* asks who are the quality monitors, what, how and why do they evaluate? Lynd argued that social science in the inter-war years asked insubstantial questions, which failed to match the tenor of the cataclysmic times. He argued, among other things, that social science was characterised by technicians who failed to address contemporary issues because of an over-concern with developing method.

The monitoring and evaluation of higher education at the turn of the 21st century evokes similar feelings. External evaluation, in legitimating the *status quo*, fails to ask significant questions about the reality of the learning experience for students at a momentous historical juncture for post-compulsory education. Evaluators appear to be preoccupied with the method of evaluation rather than the substance.

The wide scope of evaluation is explored and some analytic frameworks are suggested. In conclusion, the paper reinforces the need for quality monitoring to engage more directly and explicitly with innovations and improvements in learning and teaching.

Who are the quality monitors?

Evaluation is both *external* and *internal* to the higher education institution. Most countries have developed some form of external evaluation of higher education. However, in addition, very few institutions have the luxury of operating without any internal evaluation mechanism.

The range of internal processes is briefly outlined below but the paper will primarily focus on external monitoring. However, it is important to bear in mind the symbiotic relationship between internal and external processes.

Internal

Internal monitoring operates at three-levels (the institution, the course and the module) and includes:
- institutional units, such as audit and assessment units, institutional research units, management information units (such as statistics section of Registry);
- sub-committees of academic boards/senate;
- standing institutional audit and review bodies;
- specially convened review boards;
- faculty-based units;
- sub-committees of faculty boards;
- programme boards (or sub-committees, programme directors);
- individual teachers and researchers;
- student organisations;
- formal or *ad hoc* groups of students at programme or module level;
- external examiners: these are usually appointed by the institution to provide an external view of the academic process, standards of programmes and, where appropriate, professional development. They tend to be independent academic appraisals although the outcomes are usually presented in confidence to the institution (Silver, 1993; Warren Piper, 1994);
- invited consultants: Jiménez, Escalante and Aguirre-Vasquez (1997, p. 265), for example, discuss the role played in 1989 by independent external experts in helping the educational community at the Autonomous University of Southern Baja California learn to adapt to changing conditions. They conclude that it is important, when using external experts, not to become dependent on them but rather to learn from them.

Internal processes are vital to ensure the continued standard of qualification and the ongoing improvement of the education and research provided by the institution. Martin Trow (1995, p. 22) argues that 'internal reviews and assessments are more accurate and fruitful than those done by outsiders' a view reinforced by Bente Kristensen in her view of the Danish system. She notes that, while there can be a fruitful synergy between external and internal processes, external monitoring can never stand alone and 'will never be able to replace valuable internal quality monitoring' (Kristensen, 1997, p. 91).

What, then, constitutes a 'valuable' process? Arguably, it is one undertaken in a spirit of responsiveness to stakeholders — new collegialism rather than cloisterism (Harvey, 1995a). There is growing concern about the articulation between formal, managerialist structures, and less formal collegiate processes. A recurring theme is the dissolution of trust between academics and managers (Trow, 1994, 1996; Dearlove, 1995; Lindsay, 1994; McInnis *et al.* 1995; Gibbs, 1998). External monitoring tends to

be perceived, by teaching staff, as reinforcing managerialism and as compromising mutual trust.

External

It is not the intention of this paper to explore the meanings of the different terms associated with external monitoring: 'evaluation', 'assurance', 'accreditation', 'validation', 'assessment', or 'audit' (Harvey, 2002). The paper is more concerned with what the quality monitors actually do and why they do it.

Auspices

A crude classification of auspices would suggest three significant dimensions (Figure 1):
- whether the monitoring agency is established or empowered by legislation (law, decree, statute). Statutory agencies are usually government departments, agencies that are ultimately responsible to a government department (education, science, employment, treasury) or bodies with delegated regulatory powers;
- whether the genesis of the monitoring agency is within or outside the higher education sector. Sector-initiated agencies are often established by the Committee of Rectors or a similar organisation. Externally, initiated agencies tend to be government or state agencies or professional/employment linked evaluation agencies;
- the degree of independence of the monitoring agency, which might be measured by the extent to which agencies are expected to show allegiance to initiators, are put under pressure by funders, or are constrained in the methods and processes of operation.

Statutory
For example, the New Zealand Qualifications Authority (NZQA) is a crown agency empowered by the 1990 *Education Amendment Act* to require providers to develop and implement mechanisms that ensure they offer quality education to their clients. The government elected in November 1999 has re-affirmed the role of NZQA as the 'overarching' body for quality assurance in the post-compulsory sector. Although able to develop its own methods of working it has been heavily influenced by government policies.

In Sweden, the National Agency of Higher Education (Högskoleverket), which is a government agency funded by government money and established by law in 1995, has responsibility among other things for monitoring quality. However, it has a fair amount of autonomy to make its own enquiries and evaluations.

In the United Kingdom, there are regulatory bodies empowered by statute to control education in specific fields such as medicine (The General Medical Council) architecture (Architects Registration Council of the UK) and social work (Central Council for the Education and Training in Social Work). Their work, although affected by changes in legislation is almost entirely independent of government interference.

Figure 1: Key dimensions of auspices ABOUT HERE

Non-statutory

The Hong Kong Council for Academic Accreditation, established in 1990 by ordinance, was an independent, self-financed organisation that reviews, validates and accredits programmes mainly on the basis of the standard of inputs.

The New Zealand Universities Academic Audit Unit, unlike NZQA, is a non-statutory agency with a governing board appointed by New Zealand Vice-Chancellor's Committee and is paid for, in part, by fees from the universities it audits. However, although established by the NZVCC, the Academic Audit Unit in New Zealand operates independently of the sector. Nonetheless, it cannot avoid being mindful of the government's 'value for money' expectations of the public sector.

Institutional accreditation in the United States is a voluntary process undertaken by six regional bodies. The government recognises accreditation agencies as providing a framework for evaluating quality but they are not statutory bodies.

In South Africa, the Quality Promotion Unit of the South African Universities' Vice-Chancellors' Association (QPU), established in February 1996, was owned and paid for by the universities. It acted relatively independently of both government and the sector but its lack of statutory status resulted in it being wound up very rapidly when South Africa embarked on a review of its external monitoring processes.

Another voluntary, non-statutory evaluation process is that established by the European Rectors' Conference (CRE), which provides a Europe-wide, sanction-free, auditing procedure available, for a fee, to institutions on request. The use of international teams of peer reviewers and the lack of pressure from state governments affords a good deal of independence to the process.

In the United Kingdom, professional bodies validate and accredit programmes of study. Some have statutory powers such as the Law Society who effectively control the flow of graduates into the legal profession. Most professional bodies have no statutory powers, recognition by the Institute of Financial Accountants is not a necessary prerequisite for a career in accountancy (Harvey & Mason, 1995). A similar, well-established, process of professional accreditation also takes place in the United States. For example, the Accreditation Board for Engineering and Technology undertakes visible and accountable inspections and graduates can only become licensed engineers if they graduate from accredited schools (Adelman & Silver, 1990). While, these agencies in the UK and US act independently of government, they act very much in the interest of the professional area.

The plethora of industry-originated, consultant-enabled, self-evaluation models provides another example of external, non-legislated evaluation. Having moved on from TQM the latest vogue in Europe is the Business Excellence model. These models appear to be independent, however, the consultants involved have a vested interest in 'selling' them.

Independence

Whether, in theory, the agency is independent the reality, in practice, will be mediated by the organisational ethos, culture, and working practices. Hence any mapping of ostensive status has to take account of the sociology of the organisation.

For example, Teaching Quality Assessment in England used to be the responsibility of the Quality Assessment Division of the Higher Education Funding Council for England (HEFCE). It was established in response to legislation that said the Council must ensure quality monitoring took place. However, the assessment process was not as closely controlled by government as this might imply. The Minister that set this process in train really wanted an inspection system but ended up with peer review.

Although government initiated, HEFCE is not a government department but a 'quango' — a semi-independent appointed body, government funded, but autonomous in operation, delegated with a specific task which otherwise might be a civil service function. Furthermore, the quality function was the responsibility of the Quality Assessment Division (QAD), a division within a council primarily concerned with funding. So, as in any diverse organisation, QAD developed its own culture. In addition, the QAD was 'overseen' by an appointed steering committee made up of diverse people from higher education, industry and government. Finally, academics were used to do the assessments. All of this meant that despite its 'provenance', Teaching Quality Assessment, in practice, was a long way from being a government-controlled process. Indeed, the November 1997 budget letter to HEFCE stated 'the secretary of state expects the council to consider further ways of linking funding for teaching with assessment of quality' (Tysom, 1998, p. 48).

Thus, as a recent EC report suggests, the character of the process tends to be a different issue from, and independent of, the matter of formal ownership' (EC, 1998, p. 7).

Although, independence is influenced by operational culture, it is also affected by both the responsibilities to stakeholders and the boundary constraints of its work. The funding, terms of reference, degree of permanency, sanctions available and political agendas all impact on the responsiveness of the agency to the pressures placed upon it.

For example, in the UK, the new Quality Assurance Agency (QAA) is, supposedly, independent. Yet it is clear from talking to civil servants in the Department for Education and Employment (DfEE) that they 'network closely' with the agency, indeed the head of the Higher Education Quality and Employability section of DfEE is on the QAA Board. There is a clear 'expectation' that QAA will be responsive to government priorities, a point no doubt reinforced in the occasional meetings between the Director of QAA and the Minister for Higher Education.

The formal auspices, the culture and ways of working of an organisation and its degree of independence offer considerable opportunities for detailed comparative research, based on the sociology of organisations, of the way agencies mediate their brief in practice.

In summary, quality-monitoring agencies in most countries tend to have some statutory basis especially where they have accreditation responsibilities. In many cases accreditation of institutions was formerly a government activity that has been delegated to agencies and this is also often the case with programme accreditation, especially in professional fields. The British model of professional and regulatory bodies responsible for accreditation and review is somewhat unusual. The American system of voluntary accreditation, that derives from the market system in the US, is unusual and not easily imported into non-market systems or into countries such as Eastern Europe and South America that have witnessed a rapid burgeoning of relatively unregulated private provision. In these situations, a central, government-endorsed body has usually been set up to ensure private provision meets basic minimum requirements. The more the central agency works with the higher education sector the more genuine improvement appears to take place, especially if the agency is perceived to be relatively independent of both the universities and the government, as was the case in audit period in Sweden. Where the agency attempts to impose or cajole the sector, resistance builds quickly and confidence in the process breaks down, which has happened in the UK over the last three years. In general, as agencies mature they tend to place more emphasis on improvement activities rather than

accountability while simultaneously becoming more influenced by politicians, keen to ensure that the agency retains or develops 'teeth'.

What do they evaluate?

Evaluation, undertaken by quality monitoring agencies, includes most aspects of higher education. Learning and teaching is a primary area and tends to include, curriculum content and delivery; teacher performance, assessment/grading of students, standards of academic attainment or competence and, more recently, employability of graduates (HEFCE, 2000; Harvey, 2000).

Resources have been a conventional concern of external agencies, and provide an easy way of 'measuring' inputs: the se include floor space per student, number of volumes in the library, number of computer terminals per student and, in some cases, the qualifications of staff and the entry qualifications of students are viewed as input resources. Management is sometimes evaluated but usually only indirectly through its budgeting ability or via the management information systems and quality systems that are in place. The quality of management *per se* is rarely evaluated.

Research is another major area of evaluation: outputs and costs are most often evaluated but the nature of the research community is also taken into consideration in some monitoring of the quality of research. Finally, evaluation examines external links with the community and with other stakeholders, notably employers.

This is a quite staggering degree of examination, which has grown up rapidly over the last decade and reflects the intrusion of the evaluative state into all areas of public life (Neave, 1998).

Why evaluate?

Most external or internal agencies undertake evaluation for one or more of the following purposes:
- accredit or license institutions;
- audit institutional procedures;
- accredit programmes of study;
- assess the quality of teaching;
- assess research projects/proposals;
- assess research outputs;
- set or define academic standards or standards of competence;
- check levels of academic standards or standards of professional competence;
- ensure conformance with regulations or procedures;
- assess participant ('customer' or 'client') satisfaction with service provision (at institutional, programme or module level);
- gauge effectiveness (of teaching and learning, management systems, quality systems);
- ensure value for money (of teaching and learning, research, resources, management);
- evaluate the quality improvement process — identify improvement projects and evaluate their effectiveness;
- develop a system of credit transfer;

- disseminate good practice.

The array of purposes is impressive and suggests that evaluation is addressing itself at every facet of the higher education process. The reality, though, as will be shown, is that real evaluation of the nature of higher education is subsumed by political imperatives and despite the comprehensiveness of the purposes, significant questions are side-stepped.

What is the focus of the evaluation?

Although evaluations have a range of purposes it is not always clear whether the focus is on the quality of the process or the standard of the outcome. Is it the process of learning or the standard of what is learned? Is it the quality of the learning environment and its management or is it the fulfilment of specified service standards?

Quality

Quality is not uni-dimensional but has several dimensions, for example, quality as exceptional, as perfection (or consistency), as fitness for purpose, as value for money and as transformation (Harvey & Green, 1993) (Figure 2).

It is arguable whether theses dimensions are discrete or overlapping constructs. Similarly, it has been suggested that transformation is a meta-quality concept and that other concepts such as perfection, high standards, fitness for purpose and value for money are *possible* (although not very good) operationalisations of the transformative process rather than ends in themselves (Harvey, 1994, p. 51; Harvey & Knight, 1996, pp. 14–15).

Standards

Standards are measures of what is achieved, usually outcome standards. (Although, of course, value-added evaluations depend on the measurement of some form of input standards against which the output standard can be compared). Standards relate to the academic attainment of students and research attainment of staff, levels of competence of students (on professional courses), service standards (such as teaching standards and the standard of learning resources) and organisational standards. (Figure 2)

<u>Figure 2 Definitions of quality and standards ABOUT HERE</u>

Relating standards to quality generates a grid with 20 cells, the content of which indicate the focus for evaluation (Figure 3).

<u>Figure 3 Relationship between quality and standards ABOUT HERE</u>

Many agencies conflate quality with standards and most fail to disentangle different notions of quality and outcome standards. Often, fitness for purpose is used as a convenient definitional shortcut for quality. However, this is rarely invoked in a pure form, in practice, as most agencies are drawn to providing explicit or implicit gradings or ranking of institutions against some more-or-less fixed criteria. In all

events, the purpose is rarely if ever defined by the 'customer' but based on governmental/agency criteria or the stated mission of the institution: it is a producer-determined fitness.

Worse, the rhetoric of 'quality standards' provides a convenient means to side-step any real evaluation of learning theory or innovation in learning and teaching. Only two of the cells (in Figure 3) — transformation by academic standards and by standards of competence — emphasise the learning process and afford opportunities for evaluation to explore innovation in learning and an engagement with learning theory. In practice, few agencies explicitly address quality as transformation. They prefer, instead, a simpler model, such as 'fitness for purpose', that allows them to prioritise method rather than the substance of the learning process: learning theory and its implementation. Maureen Tam (2001, p. 53), for example, has recently reasserted the transformative element of higher education quality monitoring:

> Any measurement of quality and performance evaluation in higher education that falls short of the centrality of student's experience is bound to be peripheral and fail to provide information about how students find the experience and how much they are learning and progressing both intellectually and emotionally throughout their years in university.

How do the quality monitors evaluate?

Despite the varied objects of evaluation and the array of different types of agencies, there is a surprising conformity in the methods that are adopted. Approaches to evaluation in higher education, as has frequently been pointed out are heavily dependent on three basic elements (Frazer, 1995; Harvey & Knight, 1996; Massaro, 1996; EC, 1998).
- self-assessment (or submission);
- peer evaluation;
- statistical or performance indicators.

The results are prepared as a report that usually becomes a public document, albeit that a more detailed version may remain confidential.

Process

Typically, the procedure is for the institution or programme of study (or subject area) to produce a self-evaluation report or some other form of submission for assessment, such as a research profile. The qualitative self-evaluation is often complemented by statistical data.

The report (and the appropriate statistical data) are scrutinised by an external body. Sometimes more information is requested either by the co-ordinating body or the team of 'respected' peers who will subsequently visit. This additional material may be received in advance or be available during the visit.

The peer-review panel visits the institution. Usually such a visit lasts between one and four days. They attempt to relate the self-assessment document to what they see or, in practice, hear. In some cases the peers may observe facilities or even the teaching and learning process itself, although the latter is rare.

If the intention is to evaluate evaluators it is necessary to address the taken-for-granteds that are buried within the dominant methodology. The following brief

discussion suggests that reliance on these methods render evaluative processes far less effective than they might otherwise be.

Critique of process

Statistical indicators

Statistical indicators, as evaluative tools, have their limitations as measures of quality performance. In the early 1990s there was much research on 'performance indicators', most of which suggested that statistical indicators, whether reliable or not, are rarely *valid* operationalisations of quality (Klein & Carter, 1988; Cave & Kogan, 1990; Goedegebuure *et al.*, 1990; Head, 1990; HMI, 1990; Johnes & Taylor, 1990; Pollitt, C, 1990; Cave *et al.*, 1991, Gallagher, 1991; Yorke, 1991; Murphy, 1994). Furthermore, despite being 'indicators' it is unclear, exactly, of what performance they are indicative.

> What, for example, does an increase in percentage of 'good' degree classifications tell us about quality? Does it indicate that the student learning performance has improved? Does this mean that the teaching staff have performed better, or are the students learning more despite the teachers? Or does it mean that academic standards have fallen? Similarly, what does the employment rate of graduates within the first six months after graduation tell us about the performance of the institution? Perhaps it says more about the vagaries of the recruitment process and the differential in take-up rates between different subject specialisms than provide any indication of the performance of the institution. In short, so-called performance indicators are invariably simplistic, convenience measures that bear no relation to any notion of quality. (Harvey, 1998, p. 243)

Yorke (1998) suggested that the benefit that might accrue from improving statistical measures to make them into really meaningful performance indicators is outweighed by the cost that would accrue.

Although there is a good deal of scepticism about the value of quantitative indicators of higher education quality, politicians still like to use them. In some countries, such as the United Kingdom, statistical indicators play a minor role in quality evaluations. However, there is a growing desire for performance indicators, even in the UK, which suggests a revival of a 'quick-fix' policy agenda. There are, of course, other countries, such as Brazil, where fairly crude indicators have been part of institutional evaluations. Conversely, in the United States, where quantitative indicators are a way of life, there has, purportedly, been a gradual shift towards placing more credence on qualitative assessments based on peer reviews (Banta, 1995).

Self evaluation

Whereas statistical indicators invite creative accounting, self-evaluation, in the right context, is useful for encouraging fundamental reviews of objectives, practices and outcomes. Account after account has indicated how important self-evaluation is

(Karlsen & Stensaker, 1995; Saarinen, 1995; Rasmussen, 1995; Bazargan, 1999). There remain questions about the appropriate frequency and depth of self-assessment and the relevance of different models of self-assessment. What is clear is that the less threatening the evaluation process the more open, honestly reflective and useful is the self-evaluation process.

Peer review

Although self-assessment is often taken seriously only if peer review follows, the peer reviews themselves are not particularly an effective or efficient means of unravelling what is really going on. In the main, peer-review teams make judgements based on what they are told and tend to look for discrepancies in the story. They attempt to relate what they hear (and sometimes see) to the self-assessment document. However, in practice, there tends to be a significant gap in the perceptions of peers and the authors of self-assessment documents. Peer groups see relatively little as they spend most time closeted in a room having discussions with group after group of 'selected' discussants. It is also unusual for peer reviewers to have detailed documentation or, if they do, the time to read and evaluate it thoroughly. Even if the peer team has documentation, which allows some form of cross-checking, and they observe facilities and practices first-hand, they tend to see and assimilate only a tiny fragment of the entire institutional operation.

Peer reviewers are encouraged to ask questions but they are not trained as investigators. Sometimes they are not trained at all. There is very little attempt to challenge the preconceptions and prejudices of peers—after all their views are to be 'respected'. The little 'training' or 'briefing' given to peer groups is usually about what areas need to be examined and the sorts of things on which to focus. Peers are rarely trained *how* to identify and interpret what they see. A study in Chile, for example, suggested that, even in the newly developing private university sector, peer reports, in 90 per cent of cases were simply confirming what the institutions already knew and, furthermore, the prior experience of peer reviewers tends to influence the outcome of reports (Silva, Reich & Gallegos, 1997, p. 31).

Setting

In the appropriate setting, self-evaluation and peer review can be a significant spur to fundamental self-reflection. If the institution wants to explore its purpose, its areas of effectiveness, its weaknesses and future opportunities then self-evaluation, followed by a peer-review process that involves open dialogue and helpful feedback, can be an invaluable tool. It can help develop a future strategy for continuous improvement. However, the long-term effectiveness is entirely dependent on the establishment of internal procedures and development of a culture of continuous improvement. For example, the European-wide, CRE-Audits, undertaken on a voluntary basis, have been useful in helping most universities that have taken part to develop strategic plans. Whether, in the long term, they will result in a process of continuous quality improvement depends on how well the outcomes are communicated and linked in with the day-to-day activities of the teaching and research staff.

If the monitoring is compulsory it is much more likely that the staff in the institution see the self-evaluation as part of a judgmental process, especially if it is linked to

status rankings or to funding. In which case, there will be a disinclination to be open and frank about weaknesses and a tendency to overstate strengths. On the contrary:

> There is clearly gamesmanship. Academics are clever people, they can find a way round. The longer it has been going on, the more people work out how to play it. There is a whole infrastructure of training and information to help people get the best out of the system. (Brown, quoted in Baty, 1999, p. 6)

If people are inhibited about being open then the self-evaluative process becomes a defensive account rather than an opportunity to explore future development and change. In such circumstances, self-evaluation followed by an inquisitorial peer review reveals little and does not provide the basis for effective improvement. Instead of open responsiveness there tends to be a reversion to the defensiveness of the cloister (Harvey, 1995a; 1998).

Performance

Mark Barrow (1999) characterises the game-playing as a 'performance' designed to provide the illusion of 'objective' evaluation. He elaborates this in the specific case of NZQA accreditation of programmes. A central feature of the accreditation process is a requirement that 'there is a collective demonstration that the components of the quality-management system are understood and will be applied in the field being examined' (Barrow, 1999, p. 33). The process adopted is the familiar one of the accreditation panel interviewing a range of staff from the institution.

> These institutional staff are encouraged to demonstrate their competence by revealing an understanding of the processes in each of the eight elements of the quality system. The staff are not expected to elaborate on the product that they will develop or deliver. Indeed, during the accreditation exercise, the accreditation panellists are discouraged, by the panel's chairpersons trained by NZQA or NZPPC, from looking at the product, rather they are encouraged to look for a 'systems approach' and to examine the staff's understanding of the various sub-systems. (Barrow, 1999, p. 33–4)

Barrow (1999, p. 34) claims that this process encourages a 'dramaturgical compliance' in which staff do or say things that accord with a rule. The staff are not exhibiting a shared agreement on substantive matters, nor, he claims, are they exhibiting 'a meaningful understanding of the purpose of the system — the achievement of quality definitions'. Rather, the aim of the performance by staff is to present themselves in a positive light' and to 'play out the role assigned to them in the situation'. In short, in performing, staff 'withhold any sign of another role they may play which, if visible, would or could threaten the impression being given at the moment'. In Goffman's (1971, p. 139) terms, staff in accreditation situations in New Zealand practise 'audience segregation'.

This is not a situation unique to New Zealand. It is unlikely that there is a single country in which these performances are not played out. In the UK, for example, staff indicate that they have been told that they should not raise issues with external assessors that might indicate problems. This has meant, for example, that even when these staff are aware of how resourcing cuts or the imposition of modularisation and semesterisation are undermining the quality of the student experience, they have to

pretend that they endorse and support initiatives and that everything is satisfactory. Even in cases where lecturers have broken ranks and been blunt with assessors, the shared 'performance' is suitably 'convincing' — that is, it complies with expectations and requirements of a system approach — that the malcontent is regarded as a maverick or troublemaker and the honest comment is effectively disregarded.

In the UK, millions of pounds are spent every year to discover that, on the basis of the teaching quality assessments, fewer than one per cent of courses are failing.[1] This, Haakstad, (2001, p. 81) suggests is 'a heavy price to pay (for the government and for the institutions) for the tracing of a tiny number of sub-standard programmes, whose shortcomings might possibly be found out through simpler control procedures'.

Similarly, in many UK higher education institutions the Quality Audit process (formally undertaken by HEQC) is heavily orchestrated. Typically, auditors 'hold court' in the University Senate Room and see a stream of visitors, usually in small groups. These groups are summoned early by the university senior managers, briefed before they go in to see the auditors and de-briefed when they come out. The auditors hear a story that reflects the formal organisational process. However, formal structures may not match the reality of the living and dynamic organisation that is the university. If the audit process wants to know what really goes on, then a different approach is necessary: one that involves 'grubbing about in departments', which might be achieved by an audit with 'drill down' as currently being proposed for the UK or the extended audit procedure to be launched in Australia (INQAAHE, 2001).

Conclusion

Despite the vast array of agency types, purposes and focuses of quality evaluation of higher education there is remarkably consistency of approach. In the main, inspection is eschewed in favour of 'peer review'. The quality monitoring procedure is preoccupied with method and not student learning. As Bjorn Stensaker (1999, p. 365) noted, in his review of Swedish Audits, 'on the whole, the visits appeared to be more geared to the needs of the auditing teams than those of the institution' .

External quality monitoring is primarily to ensure accountability and conformity. Improvement is an 'add-on' that is presumed to result from compliance with the method. However, improvement in student learning requires a real engagement with learning processes, with the fundamentals of learning. Playing around with documentation and peer reviews, be it via audits, assessments of teaching or benchmarking, is an inefficient way to encourage and support the development of student-oriented learning facilitation. David Gosling and Vaneeta D'Andrea (2001) argue, for example, that the enormous growth in quality assurance processes over the last ten years in higher education in the UK has required significant extra effort by academic staff but this effort to conform to agency requirements is misplaced and does not necessarily produce enhancement of the student experience.

Most agencies, it seems, have not even considered how their method-led, bureaucratic procedures fail to engage student learning. They have tended to ignore a quarter of a

---

[1] Ironically, as the system has become more 'sophisticated' its ability to identify 'failures' has decreased — the old system identified 1.2% 'unsatisfactory' provisions compared to 0.6% by 1999 (Baty, 1999b). This could, of course be that the sector is improving or is better at 'performing' or that the modified more sophisticated approach is actually less effective.

century of research into learning theory, the nature and styles of learning and classroom innovations. Agency reports fail to act as a conduit for good practice. The only thing most external quality monitoring is useful for is the 'peer reviewers' who learn a lot from the experience (Niklasson, 1997). Unfortunately, there is little evidence that this invaluable personal experience is either shared with the reviewers' colleagues or reflected in the experience of the reviewed.

Until higher education monitoring agencies take an explicit 'transformation' approach to quality (Harvey and Knight, 1996, Astin, 1985) and seriously think about what that involves for student learning they will continue to be 'technicians' failing to ask substantial questions.

References

ADELMAN, C. & SILVER, H. (1990) *Accreditation: The American experience.* (London, CNAA).

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

BANTA, T. (1995) An assessment of some performance indicators used in funding: performance funding in Tennessee at age sixteen, paper presented at the 17th Annual EAIR Forum, *Dynamics in Higher Education: Traditions Challenged by New Paradigms*, Zurich, Switzerland, 27–30 August 1995.

BARROW, M. (1999) Quality management systems and dramaturgical compliance, *Quality in Higher Education*, 5(1), pp. 27–36.

BATY, P. (1999) Encounters of an unfair kind, *Times Higher Education Supplement, 19* March 1999, pp. 6–7.

BAZARGAN, A. (1999) Introduction to assessing quality in higher medical education in Iran: challenges and perspectives, *Quality in Higher Education*, 5(1), pp. 61–68.

CAVE, M. AND KOGAN, M. (1990) Some concluding observations, in M. Cave, M. Kogan, & R. Smith (eds.) (1990) *Output and Performance Measurement in Government: The state of the art,* pp. 179–87 (London, Jessica Kingsley).

CAVE, M., HANNEY, S., KOGAN, M. & TREVETT, G. (1991) *The Use of Performance Indicators in Higher Education: A critical analysis of developing practice.* (London, Jessica Kingsley).

DEARLOVE, J. (1995) Collegiality, managerialism and leadership in English universities, *Tertiary Education and Management*, 2 (2), pp. 161–169.

EUROPEAN COMMISSION (EC) (1998) *Evaluation of European Higher Education: A status report*. (Brussels, EC).

FRAZER, M. (1995) Ten papers on national developments, in International Network of Quality Assurance Agencies in Higher Education (INQAAHE) (1995) *Third Meeting of the INQAAHE, 21–23 May 1995, Utrecht, the Netherlands: Proceedings*. Netherlands, VSNU/IHO/INQAAHE, pp. 55–71.

GALLAGHER, A. (1991) Comparative value added as a performance indicator, *Higher Education Review*, 23 (3), pp. 19–29.

GIBBS, P. (1998) Competence or trust: the academic offering, *Quality in Higher Education*, 4(1), pp. 7–16.

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* (Culemborg, Lemma).

GOFFMAN, E. (1971) *The Presentation of Self in Everyday Life.* Harmondsworth, Penguin.

GOSLING, D. & D'ANDREA, V., 2001 Quality development: a new concept for higher education, *Quality in Higher Education*, 7(1), pp. 7–18.

HAAKSTAD, J. (2001) Accreditation: the new quality assurance formula? Some reflections as Norway is about to reform its quality assurance system *Quality in Higher Education*, 7(1), forthcoming.

HARVEY, L. (1994) Continuous quality improvement: A system-wide view of quality in higher education, in P.T. Knight (ed.) (1994. *University-Wide Change, Staff and Curriculum Development,* Staff and Educational Development Association, SEDA Paper, 83) May 1994) pp. 47–70.

HARVEY, L. AND GREEN, D. (1993) Defining quality, *Assessment and Evaluation in Higher Education: An International Journal*, 18 (1), pp. 9–34.

HARVEY, L. AND KNIGHT, P. (1996) *Transforming Higher Education* (Buckingham, Open University Press and Society for Research into Higher Education).

HARVEY, L. (1995) The new collegialism: improvement with accountability, *Tertiary Education and Management*, 2 (2), pp. 153–60.

HARVEY, L. (1995b) Editorial, *Quality in Higher Education*, 1(1), pp. 1–12.

HARVEY, L. (1998) An assessment of past and current approaches to quality in higher education, *Australian Journal of Education*, 42(3), pp. 237–255.

HARVEY, L., (2002), 'External quality evaluation: some international trends', keynote at the Higher Education Conference, Oslo, 22 January 2002.

HEAD, P. (1990) *Performance Indicators and Quality Assurance.* Information Services Discussion paper, 4 June, 1990. (London:, CNAA.)

HIGHER EDUCATION FUNDING COUNCIL FOR ENGLAND (HEFCE), (2000) *Consultation on employment indicator.* Letter from Brian Fender, Chief Executive and Annexes A and B. (Bristol, HEFCE).

INTERNATIONAL NETWORK OF QUALITY ASSURANCE AGENCIES IN HIGHER EDUCATION (INQAAHE), 2001, 'Australian Universities Quality Agency (AUQA)', *QA*, 21, November, pp. 7–8.

JIMÉNEZ, J., ESCALANTE, J.C, & AGUIRRE-VASQUEZ, J. (1997) The role of external experts intervention in higher education for radical change, *Tertiary Education and Management*, 3 (3), pp. 249–58.

JOHNES, J. AND TAYLOR, J. (1990) *Performance Indicators in Higher Education.* (Buckingham, Society for Research into Higher Education (SRHE)/Open University Press).

KARLSEN, R. AND STENSAKER, B. (1995), Between governmental demands and institutional needs: peer discretion in external evaluations—what is it used for?, paper presented at the 17th Annual EAIR Forum, *Dynamics in Higher Education: Traditions Challenged by New Paradigms*, Zurich, Switzerland, 27–30 August 1995.

KLEIN, R. AND CARTER, N. (1988) Performance measurement: a review of concepts and issues, in D. Beeton (Ed.), *Performance Measurement: Getting the concepts right*, (Public Finance Foundation).

KRISTENSEN, B. (1997) The impact of quality monitoring on institutions: a Danish experience at the Copenhagen Business School, *Quality in Higher Education*, 3(1), pp. 87–94.

LINDSAY, A. (1994) The impact of the Australian government's quality assurance program on institutional roles and management processes, paper presented at the AIR Forum, New Orleans.

LYND, R. (1939) *Knowledge for What?* (Princeton, Princeton University Press)

Massaro, V. (1996) Learning from audit? A survey of OECD countries, paper presented at the Swedish National Agency for Higher Education Conference: quality assurance as a support for processes of innovation, Castle Lejondal, September.

McInnis, C., Powles, M. & Anwyl, J. (1995) Australian academics perspectives on quality and accountability, *Tertiary Education and Management*, 2 (2), pp. 131–139.

Murphy, P. (1994) Research quality, peer review and performance indicators, *The Australian Universities Review*, 37 (1), pp. 14–18.

Neave, G. (1998) The evaluative state reconsidered, *European Journal of Higher Education*, 33(3).

NIKLASSON, L., (1997) *Vad tycker de inblandade i första omgången bedömningar av universitetens och högskolornas kvalitestarbete?* (Stockholm, National Agency for Higher Education, 1997: IR)

POLLITT, C. (1990) Measuring university performance: never mind the quality, never mind the width, *Higher Education Quarterly*, 44 (1):, pp. 60–81.

RASMUSSEN, P. (1995) A Danish approach to quality in education: the case of Aalborg University paper, with additional comments, presented at the Organisation for Economic Co-operation and Development (OECD), Programme on Institutional Management in Higher Education (IMHE) Seminar, at OECD, Paris, 4–6 December (1995.

SAARINEN, T. (1995) Systematic higher education assessment and departmental impacts: translating the effort to meet the need, *Quality in Higher Education*, 1 (3), pp. 223–234.

SILVA, M., REICH, R. AND GALLEGOS, G. (1997) Effects of external quality evaluation in Chile: a preliminary study, *Quality in Higher Education*, 3 (1), pp. 27–36.

SILVER, H. (1993) *External Examiners: Changing Roles?* (London, CNAA).

STENSAKER, B. (1999) External quality auditing in Sweden: are departments affected?, *Higher Education Quarterly*, 53(4), p. 353–368.

TAM, M., (2001) Measuring quality and performance in higher education, *Quality in Higher Education*, 7(1), pp. 47–54, forthcoming.

TROW, M. (1994) *Managerialism and the Academic Profession: Quality and Control.* (London, Open University Press).

TROW, M. (1995) Aspects of quality in higher education in Trow, M., (1995), *Two Essays on Quality in Higher Education*. Stockholm, Kanslersämbetets Skriftserie 1995:2, pp. 7–27.

TROW, M. (1996) Trust, markets and accountability in higher education: a comparative perspective, *Higher Education Policy*, 9(4), pp. 309–324.

TYSOM, T. (1998) The quality street to cash, *Times Higher Education Supplement*, 16 January, 1998, p. 48.

WARREN PIPER, D. J. (1994) *Are Professors Professional? The organisation of university examinations*. (London, Jessica Kingsley).

YORKE, M. (1991) *Performance Indicators: Observations on their use in the assurance of course quality*, Council for National Academic Awards Project Report 30, January (London, CNAA).

YORKE, M. (1998) Performance indicators relating to student development: can they be trusted? *Quality in Higher Education*, 4(1), pp. 45–61.
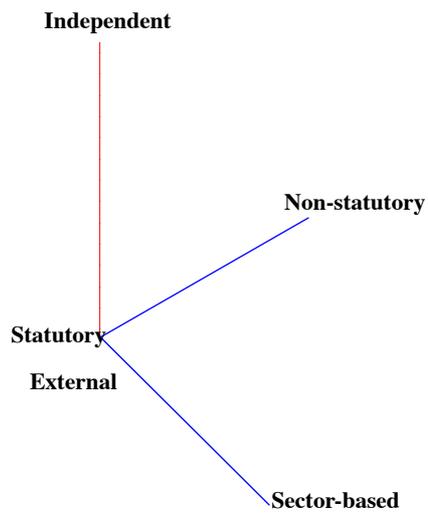
Figure 1: Key dimensions of auspices

Independent

Non-statutory

Statutory

External

Sector-based

Figure 2 Definitions of quality and standards.

Quality	Definition
Exceptional	A traditional concept linked to the idea of 'excellence', usually operationalised as exceptionally high standards of academic achievement. Quality is achieved if the standards are surpassed.
Perfection or consistency	Focuses on process and sets specifications that it aims to meet. Quality in this sense is summed up by the interrelated ideas of zero defects and getting things right first time.
Fitness for purpose	Judges quality in terms of the extent to which a product or service meets its stated purpose. The purpose may be customer-defined to meet requirements or (in education) institution-defined to reflect institutional mission (or course objectives).
Value for money	Assesses quality in terms of return on investment or expenditure. At the heart of the value-for-money approach in education is the notion of accountability. Public services, including education, are expected to be accountable to the funders. Increasingly, students are also considering their own investment in higher education in value-for-money terms.
Transformation	Sees quality as a process of change, which in higher education adds value to students through their learning experience. Education is not a service for a customer but an ongoing process of transformation of the participant. This leads to two notions of transformative quality in education: enhancing the consumer and empowering the consumer.
Standards	
Academic standards	The demonstrated ability to meet specified level of academic attainment. For pedagogy, the ability of students to be able to do those things designated as appropriate at a given level of education. Usually, the measured competence of an individual in attaining specified (or implied) course aims and objectives, operationalised via performance on assessed pieces of work. For research, the ability to undertake effective scholarship or produce new knowledge, which is assessed via peer recognition.
Standards of competence	Demonstration that a specified level of ability on a range of competencies has been achieved. Competencies may include general transferable skills required by employers; academic ('higher level') skills implicit or explicit in the attainment of degree status or in a post-graduation academic apprenticeship; particular abilities congruent with induction into a profession.
Service standards	Are measures devised to assess identified elements of the service provided against specified benchmarks. Elements assessed include activities of service providers and facilities within which the service takes place. Benchmarks specified in 'contracts' such as student charters tend to be quantified and restricted to measurable items. *Post hoc* measurement of customer opinions (satisfaction) are used as indicators of service provision. Thus, service standards in higher education parallel consumer standards.
Organisational standards	Attainment of formal recognition of systems to ensure effective management of organisational processes and clear dissemination of organisational practices.

Source: adapted from Harvey, 1995b

Figure 3 Relationship between quality and standards.

Standards / Quality	Academic standards	Standards of competence	Service standards	Organisational standards
Exceptional	Emphasis on summative assessment of knowledge and, implicitly, some 'higher-level' skills. Implicit normative gold standard. Comparative evaluation of research output. Élitism: the presupposition of a need to maintain pockets of high quality and standards in a mass education system.	Linked to professional competence; emphasis mainly on traditional demarcation between knowledge and (professional) skills.	Input-driven assumptions of resource-linked service/facilities. Good facilities, well-qualified staff, etc. 'guarantee' service standards. Reluctance to expose professional (teaching) competence to scrutiny.	Clear role hierarchy reflecting academic status and experience. Often a heavy emphasis on 'traditional values'. Strong emphasis on autonomy and academic freedom. Aversion to transparency.
Perfection or consistency	Meaningless, except for an idealistic notion that peer scrutiny of standards or quality will be undertaken in a consistent manner.	Expectation of a minimum prescribed level of professional competence. Problem in assessing for 'zero defects'.	Primary relevance in ensuring service-standard based quality — mainly in relation to administrative processes (accuracy and reliability of record keeping, timetables, coursework arrangements, etc.)	Right first time. Document procedures, regulations and good practice. Obtain ISO9000 certification.
Fitness for purpose	Theoretically, standards should relate to the defined objectives that relate to the purpose of the course (or institution). Summative assessment should be criteria referenced, although as purposes often include a comparative element (e.g., in mission statement) these are mediated by norm-referenced criteria.	Explicit specification of skills and abilities related to objectives. Evidence required to at least identify threshold standards. Professional competence primarily assessed in terms of threshold minimums against professional body requirements for practice.	The purpose involves the provision of a service. Thus, process is assessed in terms of (minimum) standards for the purpose — usually in terms of teaching competence, the link between teaching and research, student support (academic and non-academic) and so on.	Ensure appropriate mechanisms in place to assess whether practices and procedures fit the stated mission-based purposes.
Value for money	Maintenance or improvement of academic outcomes (graduate standards and research output) for the same (or declining) unit of resource. That is, ensure greater efficiency. Similarly, improve the process-experience of students. Concern that efficiency gains work in the opposite direction to quality improvement. Provide students with an academic experience (qualification, training, personal development) to warrant the investment.	Maintain or improve the output of generally 'employable' graduates for the same unit of resource. Similarly, ensure a continual or increasing supply of recruits to post-graduation professional bodies. Provide students with an educational experience that increases competence, in relation to career advancement, which ensures a return on investment.	Customer satisfaction analyses (student, employers, funding bodies) to assess process and outcomes. Students and other stakeholders are seen as 'paying customers'. Customer charters specify minimum levels of service (and facilities) that students (parents, employers) can expect.	Relies heavily on periodic or *ad hoc* reviews of whether organisational structure is effective and efficient, often informed by management information (especially basic output statistics).
Transfor-mation	Evaluation of the learning environment and processes. Assessment of students learning: their acquisition of transformative knowledge and skills (analysis, critique, synthesis, innovation) against explicit objectives. Focus on adding value rather than gold standards. As transformation involves empowerment, formative as well as summative assessment is required. Transformative research standards are assessed in terms of *impact* in relation to objectives.	Provide students with enhanced skills and abilities that empower them to continue learning and to engage effectively with the complexities of the 'outside' world. Encouraging learning how to learn. Assessment of students' acquisition of transformative skills (analysis, critique, synthesis, innovation) and the transformative impact they have post-graduation.	Emphasis on specification and assessment of standards of service and facilities that enable the process of student learning *and* the acquisition of transformative abilities.	Emphasis on organisational structure that encourages dialogue, teamworking and, ultimately, empowerment of the learner. Delegated responsibility for quality and standards. Innovation, responsiveness and 'trust' are prominent.

Source: adapted from Harvey, 1995b