

**UCE**

University  
*of*  
Central England  
*in*  
Birmingham

# **Evaluating the evaluators**

**Lee Harvey**

**Opening keynote of the  
Fifth Biennial Conference of the  
International Network of Quality Assurance Agencies in  
Higher Education**

**Santiago, Chile**

**May, 1999**

**Centre  
*for*  
Research  
*into*  
Quality**

## Introduction

The theme of this Conference is 'evaluating the evaluators' and it is therefore appropriate to ask who the evaluators are, what, how and why they evaluate before exploring who evaluates the evaluators.

The first part of the presentation, thus, poses some analytic questions about who are the evaluators and what do they do? The second explores under what circumstances evaluators are evaluated and the purpose of such evaluations. The second part of the paper raises rather more critical questions about the 'politics of quality'.

The conclusion addresses methodological issues, in particular, how to incorporate a political critique in evaluations of the impact of evaluation.

The aim in the first part of the paper is to explore the wide scope of evaluation and to suggest some analytic frameworks. However, it does not attempt to be exhaustive.

## Who are the evaluators?

Evaluation is both *external* and *internal* to the higher education institution. More and more countries are developing 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, we should not forget 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).

The crucial feature is what constitutes a ‘valuable’ process. I would suggest it needs to be one undertaken in a spirit of responsiveness to stakeholders — new collegialism rather than cloisterism (Harvey, 1995). 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’. The paper is more concerned with what evaluators actually do and why they do it, as this is a necessary prerequisite for evaluating evaluators.

## *Auspices*

A crude classification of auspices would suggest three significant dimensions (Figure 1):

- whether the 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 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 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*

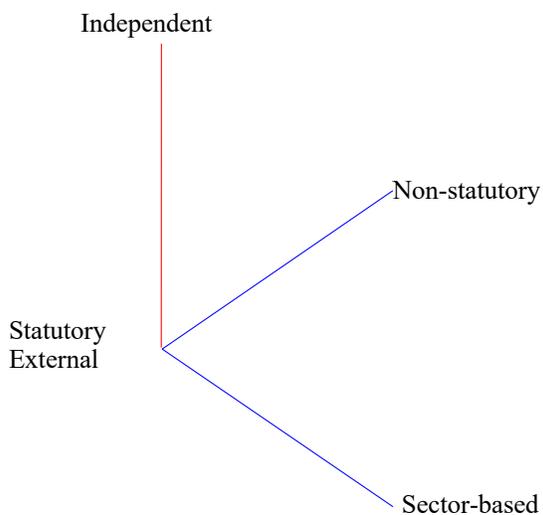
For example, the New Zealand Qualifications Authority (NZQA) is a crown agency empowered by the 1990 *Education Amendment Act* to require providers to develop and

implement mechanisms that ensure they offer quality education to their clients. Although able to develop its own methods of working it has been heavily influenced by government mandates.

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**



In Chile, The Consejo Superior de Educacion (CSE) was established by law in 1990 and is partly government funded and partly funded by fees paid by institutions undergoing evaluation. It is a public agency, chaired by the Ministry of Education with appointed council members. However, it operates as an autonomous private organisation within the public system — the government does not control it, and the institutions do not own it.

#### *Non-statutory*

On the other hand, the Finnish Higher Education Evaluation Council has no legislative status although it is funded by the Ministry of Education and plays an advisory role as part of the Government Development Plan for Education and University Research for 1995–2000 and implements evaluations as part of this plan.

The Hong Kong Council for Academic Accreditation, established in 1990 by ordinance, is 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, is owned and paid for by the universities. It acts relatively independently of both government and the sector but it is soon to be absorbed into a new system of evaluation.

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). As yet, this has not happened.

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.

## **What do they evaluate?**

Evaluation covers a wide array of aspects of higher education including the following.

Learning and teaching:

- curriculum;
- curriculum delivery;
- teacher performance;
- assessment (grading) of students;
- standards of academic attainment;
- standards of competence;
- employability of graduates.

Resources:

- computing;
- library (number of volumes);
- space (floor space per student).

Research:

- research outputs;
- research costs;
- research community.

Management:

- management information systems;
  - quality systems.
- (Rarely the quality of management *per se*)

External links:

- community impact;
- links with 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).

## How do they evaluate?

Despite the very varied objects of evaluation and the array of different types of agencies, there is a surprising conformance 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. Often, they see relatively little as they spend most time closeted in a room having discussions with group after group of 'selected' discussants. 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-granted 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.

### Statistical indicators

Statistical indicators, as evaluative tools, are of dubious worth. 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, to whose performance they relate.

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)

A recent study 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 (Yorke, 1998).

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 but even here there is a growing desire for performance indicators, which suggests a revival of a 'quick-fix' policy agenda. There are, of course, other countries, such as Brazil, where fairly crude indicators are part of institutional evaluations. Even 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.

### **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 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 institutional staff 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, 1999a, p. 4)

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, 1995; 1998a).

## 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, 0.6% of courses are failing.<sup>1</sup> Similarly, in many UK

---

<sup>1</sup> Ironically, as the system has become more ‘sophisticated’ its ability to identify ‘failures’ has decreased — the old system identified 1.2% ‘unsatisfactory’ provisions (Baty, 1999b). This could, of course be that the

institutions the institutional Quality Audit process (formally undertaken by HEQC) is entirely 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. Formal structures, though, are significantly removed from the reality of the living and dynamic organisation that is the university. If the audit process wants to know what really goes on, then an entirely different approach is necessary: one that involves grubbing about in departments.

### *A critical case study: Thames Valley University*

Thames Valley University (TVU) in the UK is a critical case study of the effectiveness of external quality monitoring. TVU moved from the status of a technical college to a university in two years. It became a university in 1992, prior to that it had been the Polytechnic of West London, founded in 1991 as a result of a merger of four technical colleges.

In the same way as every other university in the United Kingdom it has been subject to teaching quality assessments of its subject provision. So, for example, in November 1995 the Modern Language provision (French, German and Spanish) was assessed, covering 20 programmes of study. The overall score was a satisfactory 18 out of 24 and the report listed a number of strengths the last of which was:

the generally effective systems for quality assurance and enhancement, which are responsive and supportive of the new-style provision. (HEFCE, 1995b, p. 12)

As in any report there were a number of areas for improvement, but even given the British flair for understatement there were no real problems.

In March 1996, sociology was assessed and also approved with a very high score of 22 out of 24. The only concerns expressed were trivial. In short, the review suggested excellent provision. The final positive comments was:

systems for the assurance and enhancement of quality and for the regular and rigorous review and monitoring of performance are comprehensive and have a direct impact on the quality of the student experience. (HEFCE, 1996a, p. 10)

In May 1996, linguistics was also approved, again with a score of 22 out of 24. Again seven positive statements were countered by only two minor reservations. Again the final positive comment was:

a rigorous quality assurance system which has successfully devolved ownership of quality to the subject level where student feedback is taken into account; the system can be seen to impact both on curricular development and on teaching and learning. (HEFCE, 1996b, p. 13)

---

sector is improving or is better at 'performing' or that the modified more sophisticated approach is actually less effective.

In March 1997, American Studies was also approved by a visiting panel, although the score of 15 out of 24 was poor but not fatal. Nonetheless, the subject area was commended for the high quality of teaching, the learning materials provided to students, the supportive area, the wide participation, the transparency of processes and the:

detailed arrangements for monitoring and evaluating delivery of individual modules by the Subject Quality Group and the Subject Assessment Board. (HEFCE, 1997, p. 10)

Then there was a rapid change in the fortunes of the university. There were rumours that Thames Valley University was ‘officially’ dumbing down standards. At first, it was said there was an administrative error that had allowed failed-grade students to pass. The press got involved and although formally cleared of ‘dumbing down’, the damage was done. The press persisted and eventually the governors of the university invited Quality Assessment Agency (QAA) ‘inspectors’ to examine the university.

The QAA inspectors produced a damning report that, among other things, stated that the university ‘is in a position where its academic standards and the quality of its students’ experience were — and are — under threat’ and that standards can only be maintained by special measures including the appointment of an external review team to prepare an action plan ‘to rescue the University’ (Baty, 1998).

Although brought to a head by QAA, the whole issue had blown up *not* as a result of external quality monitoring. On the contrary, subject evaluations had drawn a veil over the real problems at TVU, which dated back several years. The real issue was poor staff-management relations. This was an issue raised repeatedly by the Teaching Union (NATFHE), but ignored by governors and external agencies. The crisis came once the press got hold of the story.

The reason for the initial sequence of events that had led to the ‘dumbing down’ accusations was that staff felt the students had been disadvantaged by the industrial action that had taken place at the university. The strikes were a result of a failure of dialogue between teachers and the Vice-Chancellor over the change to a new student-centred learning environment. The Vice-Chancellor was regarded by many teachers as ‘too prescriptive’ in the implementation of the plan, failing to ‘engage in dialogue’ and not ‘consulting with those at the ‘coal-face’.

Having been invited in and had the background to the dumbing down issue explained to them, QAA then produced a report that said:

We did not find there to be a self-critical academic community, but instead found a culture in which blame is cast on others..... It was evident to us that those disaffected with management and its style could not be dismissed simply as a small hard core of troublemakers. (Baty & Thomson, 1998, p. 6)

The report went on:

Furthermore, the institution had “major problems” with almost all elements of provision. External examiners were ‘disgruntled and disbelieving’ and their reports were serious indictments of the university as a degree awarding body.

Yet there was not a word of this, as far as I can see, in the subject assessments. Why not? The special enquiry discovered it easily enough. In short, the subsequent institutional enquiry, given a clear lead as to the fundamental problem, was able to go on and discover the state of the institution — a state that had not come to light in innumerable previous subject visits. Why? Because the commissioners were ‘free’ to look at what they liked and people spoke to them openly. Staff at TVU were no longer *performing*.

In conclusion, the QAA report on TVU found that the university’s:

‘structures and procedures for safeguarding quality and standards were inadequate’.

This was despite four consecutive subject reviews that had been unequivocal in declaring that TVU had:

- generally effective systems for quality assurance and enhancement;
- detailed arrangements for monitoring and evaluating delivery;
- comprehensive systems for the assurance and enhancement of quality;
- a rigorous quality assurance system.

The case study does cast a heavy shadow over the effectiveness of external quality monitoring. It suggests that the normal processes are imbued with amateurism, game-playing and performance.

## 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 purposes is impressive and suggests that evaluation is addressing itself at every facet of the higher education process.

# 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 a number of dimensions (Harvey & Green, 1993) (Figure 2):

- exceptional;
- perfection (or consistency);
- fitness for purpose;
- value for money;
- transformation.

It is arguable whether they 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 (Figure 2):

- the academic attainment of students and research attainment of staff;
- levels of competence of students from professional courses;
- service standards (such as teaching standards and the standard of learning resources);
- organisational standards.

### Figure 2 Definitions of quality and standards.

<i>Quality</i>	<i>Definition</i>
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.
<i>Standards</i>	
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. <i>Post hoc</i> 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

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

**Figure 3 Relationship between quality and standards.**

<i>Standards</i>	Academic standards	Standards of competence	Service standards	Organisational standards
<b>Quality</b>				
<b>Exceptional</b>	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.
<b>Perfection or consistency</b>	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.
<b>Fitness for purpose</b>	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.
<b>Value for money</b>	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 <i>ad hoc</i> reviews of whether organisational structure is effective and efficient, often informed by management information (especially basic output statistics).
<b>Transformation</b>	Assessment of students in terms of the standard of 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 <i>impact</i> 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. Assessment of students in terms of the 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 <i>and</i> 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.

## Who evaluates the evaluators?

Evaluation of evaluators appears to be a rather *ad hoc* process. Unlike the published reports of the evaluators, research reports that evaluate evaluators are much less easy to locate.

There are three types of evaluation of evaluation systems.

1. Opinionated or ‘theoretical’ analyses that tend to ‘predict’ the likely affects of the introduction of, or change in, evaluation systems (Wilson, 1996).
2. Analyses that are based on, often limited available evidence, much of it anecdotal: for example, the value attributed to self-evaluation is based on accumulations of such anecdotal evidence (Harvey, 1998a).
3. Analyses that are based on systematic data collection.

The paper will focus on the third of these. These ‘type-three’ analyses tend to be of three sorts.

1. Self-evaluation;
2. External evaluations initiated by agency or parent body — sometimes using consultants as ‘friendly’ advisors;
3. Independent evaluations initiated and undertaken by an individual, research centre or organisation, as one-off studies or as part of a research programme in higher education.

## For what purpose?

These ‘systematic’ analyses tend to serve a range of purposes.

1. Feasibility studies or evaluations of pilots;
2. Modifications to an existing process;
3. Evaluations of effectiveness or ability to deliver the ‘underlying rationale’;
4. Fundamental review of impact on the sector.

A two-dimensional analysis of evaluators by purpose would suggest, on this simplified framework, twelve types of evaluation (Figure 4). Selecting some examples from the literature suggests that feasibility and modification studies tend to rely heavily on agency self-evaluation or agency-initiated review. Effectiveness and impact research is more likely to be done by independent researchers. Not surprisingly, feasibility and modification studies tend to precede impact and effectiveness studies.

**Figure 4: Type and purpose of systematic evaluations (with examples)**

	<i>Self-evaluation</i>	<i>Self-initiated external</i>	<i>Independent</i>
<i>Feasibility</i>	Barblan, (1995) TQA, in UK		
<i>Modifications</i>	<i>van Hartingsveld (1999)</i> <i>Scheele (1999)</i> <i>Breslin &amp; López (1999)</i>	Barnett (1993), CHES (1994),	<i>Hu &amp; Chen (1999)</i>
<i>Effectiveness</i>	<i>Bergseth et al. (1999)</i>	<i>Meade &amp; Woodhouse (1999)</i> <i>Nilsson &amp; Wahlén (1999)</i>	Westerheijden (1996) <i>Newton (1999)</i>
<i>Impact reviews</i>	<i>Chené &amp; Freedman (1999)</i>	McInnis & Marginson (1994)	Warde (1996), Lee & Harley (1998) Baur & Kogan (1995), Baur & Henkel (1997). <i>Bégin-Heick (1999)</i> <i>Moon (1999)</i> Horsburgh, (1998, 1999)

*Those in italics are presentations at the INQAAHE Conference, 1999.*

## **Examples of feasibility studies**

Andris Barblan's report on the initial review, by CRE, of its audit process is an example of a self-evaluation feasibility study. Similarly, the feasibility and appropriateness of the proposed TQA methodology in the UK was based on a process of proposal, consultation and response to consultation.

These may not seem very exciting examples. However, they are determining moments in the introduction of processes. What is noticeable about these self-evaluation, feasibility studies is that, in most cases, the purpose and underlying rationale are rarely topics for consultation, let alone fundamental review. The emphasis tends to be on the detail of methodology, the grading systems, the type of reporting and whether there should be any link to funding. The latter may seem fundamental but is usually determined, in any event, by politicians as a matter of expediency irrespective of any 'rational' analysis.

## **Examples of modification studies**

Similarly, self-initiated external reviews designed to help modify processes, such as the report for the Higher Education Funding Councils for England and Wales by the Centre for Higher Education Studies (CHES) in April 1994, are not designed as fundamental reviews. The brief for the CHES study did not include the possibility of proposing scrapping the whole procedure. Indeed the tender contract made this clear and the 'Terms of Reference' required answers to such questions as:

- To what extent are the purposes of quality assessment being met?
- How are the self-assessments used to shape the conduct of the assessment visit?
- How did the assessment teams perform their work?
- Did the training adequately prepare assessors for their work?

The Report concluded:

We believe that it may be possible to conduct the QA system so that it meets even more effectively both the political imperatives for public accountability and the wish of all parties to deploy the QA system so as to improve the quality of provision. Our central proposals (i) visiting all institutions (ii) moving to a threshold system of judgement ... are intended collectively to attend this dual end. (CHES, 1994, p. 45)

This has been the dominant trend of initial 'official' evaluations in the UK and it is evident, for example, in the early review of the Scottish system (Barnett, 1993) and the evaluation of the Higher Education Quality Council by a firm of management consultants, whose main recommendation was that the guarded language of the reports might be made more forthright.

Apart from that, HEQC has produced overviews, or 'meta-evaluations' about what has been learned from its audits (HEQC, 1996) as have the funding councils (for example, HEFCE, 1995a) but these are more a comment on the sector, and the sector's ability to address the evaluation, than a fundamental self-reflection of the effectiveness or impact of

the evaluation process itself. This is also the case in other countries, such as Spain (CDU 1997).

In the Netherlands, the system is that the Inspectorate revisits institutions, two years after the VSNU evaluation, to explore what action has been taken. In 1992, for example, the Inspectorate produced a 'meta-evaluation' of the process that the institutions pay attention to the quality of education in a more systematic and structural way than they did before a systematic process of EQM was established. However, in general, institutions still have problems with the formulation and realisation of consistent, well-planned and managed responses to the reports of visitation committees: improvements are scattered and actions have a *short-term* character (IHO, 1992). (See also Scheele (1999) and van Hartingsveld (1999) on the Netherlands and Hu & Chen, (1999) on Taiwan).

### **Examples of evaluations of effectiveness**

Evaluations of effectiveness usually consider how effective monitoring has been from the point of view of a single institution, or attempt an evaluation across a national system.

For example, a self-study at the Hogeschool Holland, revealed that external monitoring has helped to clarify the purpose and focus of internal quality assessment. However, it has been far more effective in encouraging the development of systems of quality assurance and in improving self-evaluation than it has in enabling effective, continuous improvement of the student learning experience (van Schaik & Köllen, 1995).

In Chile, a review of initial quality monitoring procedures in a two institutions suggests that the existence of external quality monitoring has led to the establishment of permanent quality control or accrediting processes within institutions. Furthermore, it has resulted in some significant curriculum content reforms, improvement of instruments for assessing student learning and the implementation of pedagogical upgrading programmes. What is less clear is that the process is leading to a change in culture towards one of internally-driven quality improvement (Silva, Reich & Gallegos, 1997).

The Swedish National Agency commissioned a peer-review-style evaluation of its own operations, which included an evaluation of quality evaluation. The ensuing report noted that the agency's approach appears to have had a beneficial effect on the operations of the institutions of higher education. The agency is encouraged to continue scrutiny of the institutions' quality improvement efforts but to place more emphasis on assessment and follow-up. (See also, Nilsson and Wahlén (1999) and Bergseth *et al.*, 1999 on the Swedish National Agency, and Meade & Woodhouse (1999) on the effectiveness of New Zealand Universities AAU).

The Centre for Higher Education Policy Studies (CHEPS) provides good examples of independent research exploring the effectiveness, at a system level, of external quality monitoring. A study in 1993, for example, showed that, although quality was clearly on the agenda of institutions in the Netherlands, there was no clear linear relation between recommendations made by the visitation committees and measures taken by the institutes (Frederiks, Westerheijden & Weusthof, 1993).

Don Westerheijden's (1996) study of the VSNU research evaluation in Dutch universities similarly showed that the evaluations had rapidly become part of the 'culture' of Dutch universities. The VSNU system was credible and provided a basis for internal differentiation between research areas. However, the research suggested that the credibility was mainly because the VSNU reviews were compatible with existing internal judgements and external reputations among Dutch universities. He also asked whether the process

improved the quality of research? Nonetheless, the VSNU reviews are successful in fulfilling the underlying function of identifying active researchers, that is, those who publish. The public evaluation process meant that ‘‘weak’, non-publishing researchers cannot remain hidden’ and will be ‘‘weeded-out’ in the long run’. However, the question remains as to ‘whether more publishing researchers, hence more publications also means *better science?*’ (Westerhijden, 1996, p. 12). The methodology and available resources permitted no definitive analysis of this question.

This raises questions about the limits of effectiveness studies. Is it appropriate to evaluate only how well the quality system achieves its specified purposes? Or is it legitimate to address the extent to which it is effective in establishing the underlying rationale of the monitoring process. Or, further, should effectiveness studies explore intended or unintended consequences? Is Westerheijden’s question about the improvement of science appropriate in an effectiveness study?

### **Examples of evaluation of impact**

Evaluations of impact are also institution-specific or sector-wide. For example, Gay Baldwin’s (1997) evaluation of the impact of the short-lived Australian evaluations on Monash university suggests that it focused attention on teaching and learning. The result has been an intensification of debate about effective learning and the use of student perspectives in the shaping of programmes of study. She suggests that, coincident to the development of external quality monitoring, there has been an improvement in the quality of attention given to teaching and learning at Monash. However, this may be as much to do with the impact of new technology as to external quality monitoring. Furthermore, some of her colleagues, she reports, are far from convinced that external quality monitoring represents an overall gain rather than loss as the costs of the process include excessive bureaucratisation, greatly increased administrative workload, a formalism that can stifle creativity and individuality and an implicit lack of trust in academic staff.

Berit Askling (1997) has explored the impact of the Swedish National Agency on Linköping University. She suggests that the Agency acts as a catalyst, rather than having a direct impact. That ‘in a soft and supportive way it has... encouraged the institutions not to just establish routines for a management of quality but for elaborating infrastructures for quality enhancement’ (Askling, 1997, p. 25). However, similar to Baldwin, she points out that although external quality monitoring has been a catalyst, ‘it is itself a response to changes that are exerting great direct and indirect impact on institutions’ (Askling, 1997, p. 25). (See also Chené and Freedman (1999) for a view from Quebec).

Craig McInnis and Simon Marginson (1994) were commissioned by the Department of Employment, Education and Training (DEET) to undertake a study of the impact of the 1987 Pearce Committee evaluation of Australian law schools. Their impact study clearly revealed that the Pearce Report, which followed the evaluation, had a considerable impact especially on those institutions where it had highlighted weaknesses. Its emphasis on library standards, for example, influenced law library developments and its advocacy of small-group teaching led to shifts in teaching practice (at least until deteriorating staff:student ratios undermined it). However, its main impact was the encouragement of a climate of debate, discussion, critical thinking, self-evaluation and continuous improvement. On the other hand, the report was almost counterproductive in some areas, not least its attempt to restrict Masters courses, its suggestion that one school might close, and its failure to grasp the significance of diversity. McInnis and Marginson note, however,

that the impact of the Committee was no greater than concurrent policy changes in Australian higher education, which raises an issue, discussed below, about identifying impacts.

Alan Warde (1996), on behalf of the British Sociological Association, undertook a small-scale study of the impact of the 1992 Research Assessment Exercise in the UK. Warde, was far more willing (on less evidence than Don Westerheijden in the Netherlands) to highlight apparent impacts of research evaluation that called into question the whole process. Warde's results showed that there was a greater emphasis on research in universities, leading, in some cases, to the appointment of additional 'research professors' and a redistribution of teaching away from research-active to non-research active staff. 'Most respondents noted greater encouragement, or pressure, to publish' although 'several doubted whether the increased emphasis on publication had done anything for the quality of the content of sociological journals' (Warde, 1996, p.1). The most remarkable impact appeared to be the 'sense of declining morale, loss of job satisfaction and a decline of collegiality'. The greater pressure meant that staff had less time to talk to colleagues, levels of personal interaction had reduced and there was greater competitiveness. Whether this could all be laid at the door of the RAE is, of course, debatable and begs methodological questions about isolating effects. However, for Warde, the RAE is symptomatic of a move towards increased managerialism: 'collegiality and a sense of democratic control had been eroded... because of changed, more centralised and more directive management practices'.<sup>2</sup>

Highly significant, for Warde, was that no-one in the small sample reported any positive effects of the RAE.

Most thought it detrimental to quality, of both teaching and research. This stands in stark, if predictable contrast, to the self-assurance of the people responsible for the exercise who, without making explicit the grounds for their beliefs, proclaim its unquestionable success.... [For example] Brian Fender, chief executive of HEFCE... was quoted by the THES (29/9/95) as saying "there is no doubt that the assessment exercise has enhanced quality". (Warde, 1996, p. 2)

A rather more rigorous and systematic analysis of research evaluations was provided by Frederic Lee and Sandra Harley (1998) who studied the impact of the British RAE on the discipline of economics. They conclude that the RAE has led to the reinforcement of the primacy of 'mainstream' economics to the detriment of alternative approaches. This has happened because the RAE process has been appropriated by the 'mainstream' economics establishment, exemplified by the Royal Economic Society (RES). The RES's power came from its control of the 'reputational' system that is central to the organisation of academic work. Although the RAE could have undermined the control of RES, it had, on the contrary, reinforced it.

By virtue of their already dominant position, the leadership of the RES was able to capture the process by which assessors were appointed to the economics panel. In particular, the Society actively supported the establishment of the Conference of Heads of University Departments of Economics (CHUDE) in 1987, whose

---

<sup>2</sup> The effects of this increase in bureaucratic or commercial practice by managers were seen 'as unjust, sometimes as highly provocative, sometimes as a form of bullying, and sometimes as threatening the health and well-being of lecturers'.

most important activity was the selection of RES-acceptable candidates for the economics panels.

Consequently, the assessors appointed to the 1989, 1992 and 1996 panels were overwhelmingly mainstream economists holding significant positions within RES, or on the editorial board of *The Economic Journal* or a member of the CHUDE Standing Committee. The impact was that this ‘rigged jury’, whose ‘paradigm-bound view’ was that ‘the quality of non-mainstream research is largely inferior to mainstream research’ rewarded departments who did mainstream research and published in core mainstream journals (Lee & Harley, 1998, p. 198).

The result of this financial intervention has ‘driven British economic departments to discriminate against non-mainstream research and the hiring of non-mainstream economists as well as to restrict if not eliminate the teaching of non-mainstream economics to students’ (Lee & Harley, 1998, pp. 198-99). Thus, for example, Marxist economists on retirement are not replaced and courses of Marxist political economy are dropped.<sup>3</sup> For Lee and Harley, the message is clear, improve ratings and hence funding by hiring mainstream economists. The impact is the loss of non-mainstream research and teaching. A kind of ‘academic cleansing’. More studies that address the micro-politics of quality evaluation are needed.

These impact analyses, thus, move us clearly into the realms of the ‘politics of quality’. It is not enough, I suggest, to undertake effectiveness or impact research that attempts to transcend the political dimension. This is a view endorsed by the ongoing British, Swedish and Norwegian *International Study of Higher Education Reforms* research programme, which has as its overall aim to:

examine how higher education reforms have been received and implemented or counteracted by academic faculty and academic leaders; and to analyse how reform policies and structural changes have affected working lives, values, identities and productivity of academics in higher education institutions. (Baur & Henkel, 1997, p. 1)

## The politics of quality

The ‘politics of quality’ refers to the macro and micro agendas that accompany the introduction of quality monitoring procedures. It is naive to think that quality monitoring is not linked to political agendas. Indeed, Alvesson & Willmott, (1996, p. 11), suggest that the achievement of quality in higher education ‘is essentially political in origin’. The politics, though, are concealed behind a facade that suggests ‘that “achieving quality” is amenable to technical and bureaucratic solutions’.

Thus, any evaluation of evaluation systems needs to unravel the politics of quality. Equally, there is also a need, as in any social science, to explore the values and political agendas of researchers as well as those who commissioned the research.

---

<sup>3</sup> The 19 departments that increased their ratings overwhelmingly hired mainstream economists (82 out of 89= 92%). Those who did not improve their lower-end ratings (1–3) hired relatively fewer mainstream economists (37 out of 54= 69%).

At the risk of oversimplification, evaluation models can be seen to have three underlying rationales, each of which have political implications. Evaluation is used to ensure or encourage:

- accountability;
- conformance;
- improvement.

I will not linger long over these issues, much has been written about them and they have been major topics of previous INQAAHE conferences (INQAAHE, 1995).

## **Accountability**

With pressure on budgets and a growing ethos of evaluation, higher education has had to become more accountable for the money it receives. In effect, this means that higher education must not only be explicit about where it spends the money but also must endeavour to provide good value for the money it receives.

A central feature of accountability is:

that of 'rendering an account' of what one is doing in relation to goals that have been set or legitimate expectations that others may have of one's products, services or processes, in terms that can be understood by those who have a need or right to understand the 'account'. For this reason, accountability is usually, if not always, linked to public information and to judgements about the fitness, soundness, or level of satisfaction achieved (Middlehurst & Woodhouse, 1995, p. 260).

Accountability is not, of course, a political-neutral underlying rationale for quality monitoring. Indeed, quality monitoring can be used to legitimate changes in the structure or resourcing of higher education, while simultaneously providing reassurance to external stakeholders about the 'standard' or 'quality' or 'international comparability' of higher education at a time of rapid change.

The 'politics of quality' might also include the role that quality monitoring has in introducing value-for-money practices, or redistributing limited resources on the basis of an apparent value-for-money exercise, such as a research assessment exercise where money is concentrated in institutions that provide 'excellent' research output,<sup>4</sup> or, as Lee and Harley suggest, concentrated in institutions that adhere to a particular paradigm.

## **Compliance**

Evaluation also encourages compliance to (emerging or existing) government policy preferences or the preferences or policy of stakeholders such as professional bodies and

---

<sup>4</sup> Such exercises, of course, rarely measure the value of the output against the cost of the research, but assume, implicitly, that well-rated research, in terms of peer review, is 'good value'. Such practices also have another political dimension, to ensure that substantial research money is concentrated rather than spread too thinly and that it is awarded to the 'correct' institutions, not least to ensure the status quo is retained.

employer lobbies. Government is usually the most important and powerful as far as higher education goes because it supplies so much of the money and in many cases controls the licensing of institutions.

Governments around the world are looking for higher education to be more responsive, including:

- making higher education more relevant to social and economic needs;
- widening access to higher education;
- expanding numbers, usually in the face of decreasing unit cost;
- ensuring comparability of provision and procedures, within and between institutions, including international comparisons;
- responding to value for money imperatives.

‘Quality’ has been used as a tool to ensure some degree of *compliance* to these overt political agendas as well as to less overt ones, such as attempts to reduce the autonomy of academia and questioning the extent to which mass higher education is producing ‘work-ready’ graduates.

At the simplest level, quality monitoring has encouraged, or even forced, compliance in the production of information, be it statistical data, prospectuses, or course documents. Such compliance means that taken-for-granted practices and procedures have had to be confronted and clearly documented. ‘It represents the minimum required shift from an entirely producer-oriented approach to higher education to one that acknowledges the rights of other stakeholders to minimum information and a degree of ‘service’.’ (Harvey, 1998b)

Quality monitoring is also a tool that can be used to ensure compliance at the local level, that is, within an institution. For example, quality assurance can be a tool to unify disparate institutions behind a common goal or framework. In the new polytechnic sector in Finland, Rektors are using the Finnish Higher Education Evaluation Council’s (FINHEEC) pilot quality audits as a way to focus the attention of very diverse component institutions onto the new polytechnic mission and procedures.

Quality monitoring can also be used as a smoke screen to ensure compliance to resource restrictions by covering the issues that arise when student numbers increase rapidly without a commensurate increase in staffing and resources: the very element of the ‘performance’ played out in peer visits that most demotivates academics.

Indeed, a quality-management system can be viewed ‘from a Foucauldian perspective as a method of controlling’. Control of academia is not via direct intervention, rather the state and the institutional management ‘maintain a degree of surveillance from a distance that ensures that the requirements of the system are met’ (Barrow, 1999, p. 35).

## **Improvement**

Evaluation is usually seen as leading to improvements in the quality of processes or the standards of outcomes. On the surface, improvement appears to be an uncontentious goal. However, once we ask what is to be improved? in what ways? and for whose benefit?, then the political dimension is immediately evident. Value-for-money improvements may, for example, not be congruous with transformative improvements.

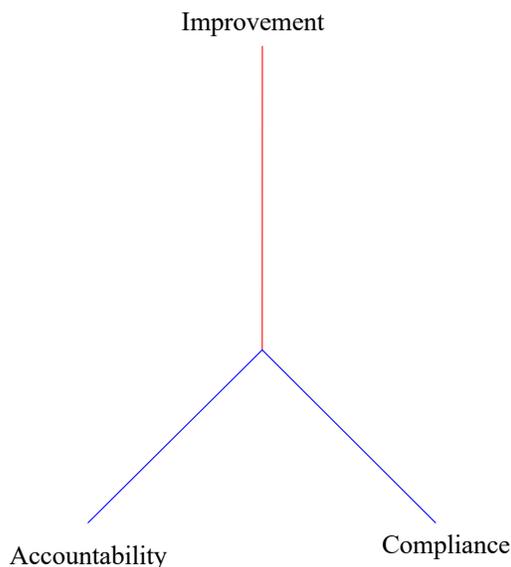
The introduction of quality improvement systems may not gel with the working practices and the culture of the institution. At the institutional level, for example, the politics of

quality can extend to leveraging a more open approach to teaching and learning via feedback from students and action based on a culture of improvement. However, this may result in a defensive, hostile reaction from teaching staff, who feel threatened.

Of course, there is a long-standing and well-rehearsed debate about the accountability and improvement axis (Vroeijenstijn, 1995; Middlehurst & Woodhouse, 1995; Massy, 1997). The crux of the debate is whether it is possible to be both an ‘accountability’ *and* ‘improvement’ organisation. Do the strictures of accountability inhibit the improvement role? Conversely, is continuous improvement its own accountability? Can an agency be both an accountability agency and an improvement agency?

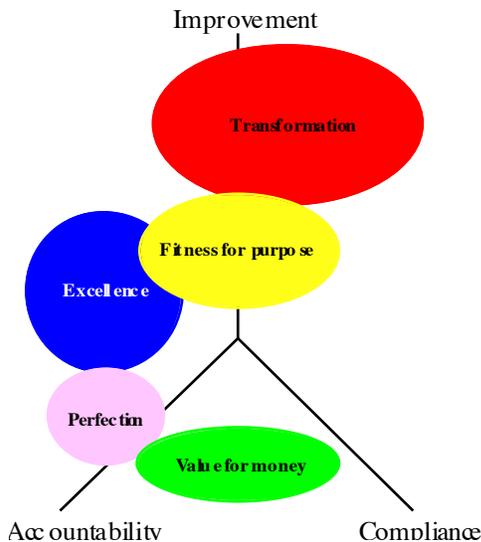
Perhaps the accountability versus improvement approach is not legitimate. Rather than see accountability and improvement as ends of a single axis it is more fruitful to conceive of them as two intersecting axes. Furthermore, these axes can be set against a third axis: compliance (Figure 5).

**Figure 5: Underlying dimensions**



Conceptually one might also link notions of quality to the three-dimensional space defined in relation to improvement, accountability and responsiveness (Figure 6).

**Figure 6: Underlying dimensions and quality**



### **Negative view**

It is the political agendas of quality that have led to steadily growing negative views of quality amongst academics (McInnis *et al.*, 1995). The structuring of quality procedures entraps academics into endorsing the ‘quality’ of a system that they know to be providing a worsening student experience. ‘Quality’ is disengaged from their own primary concerns: the enhancement of students, the development of their research, the financial management of the institution. Furthermore, ‘quality’ is structured as an exercise or game in which most academics only fleetingly take part. Academics consider that the imposition of a top-down model of accountability as a burdensome but pointless process. They really want an exploration of how quality is really improved or how improvement is impeded at the operational level. Instead, as individuals, they are encouraged to become increasingly passive and to allow their roles to be ‘shaped by the technical, instrumental rationality of systems rather than draw upon the lifeworld; their accumulated individual experiences and understandings’. As this lifeworld is ‘colonised’ there is a process of ‘cultural impoverishment’ (Alvesson & Willmott, 1996, p. 82).

Furthermore, colluding with political agendas via ‘quality performances’ that cast the programme or institution in a golden light is not only demoralising, it is dishonest. There are more TVUs to be uncovered.

Funding rewards for quality can also lead to a ‘compliance culture’, skewing the system to ‘following the money’. However, as Robin Middlehurst and David Woodhouse point out:

getting compliance is exactly what is wanted (providing the goal is clearly and unequivocally what is stated), and the criticism then turns on philosophical considerations, such as whether it is intrinsically better to have compliance by free

will than by compulsion or incentives. (Middlehurst & Woodhouse, 1995, p. 262).

Compliance by free will is, of course, congruent with academic autonomy. Middlehurst and Woodhouse argue, furthermore, that improvement only follows from intrinsic motivation. If the motivation for change is solely extrinsic, then 'compliance is the best that can be hoped for'. Indeed, 'compliance may pass for improvement in the short term but as soon as the need to display 'improvement' has passed, old habits are likely to re-emerge' (Middlehurst & Woodhouse, 1995, p. 263). This is analogous to the launching of a satellite, external evaluation provides the initial thrust that is hopefully sufficient to send the quality-bearing rocket into space and enable the launch of the satellite. Unless there are real internal procedures with local control of a process of continuous improvement the orbit will rapidly decay (Harvey, 1995). Baldwin (1997, p. 61) argues, similarly, that elaborate quality assurance mechanisms may have been a necessary catalyst for change but that if their principles are not internalised they will not be effective.

## **Methodological issues**

A fundamental issue is how compliance requirements are obscured by a thinly-veiled discourse about improvement, information requirements, stakeholder rights and so on. In any social research, especially of a socio-political setting, it is necessary to engage and critique ideology. Yet there has been a reluctance within 'higher education', as a discipline, to critique the ideology of quality evaluation.

The foregoing discussion is incompatible with a view that suggests evaluation is a neutral measuring process. Quality evaluation is a process imbued with politics.

### **Dialectical analysis not causal modelling**

Furthermore, evaluation has impacts, both intended and unintended. Thus, evaluating the evaluators raises fundamental methodological issues.

There is the enduring issue of identifying and isolating 'impacts'. However, this is not just a matter of causal attribution and the building of causal models. The relationship between quality monitoring on one hand and changes in the quality of provision on the other are not linear but *dialectical*.

The process is not reducible to systems analysis. It is important to distinguish between the quality system and the methodology used to evaluate them. Systems approaches are prevalent in quality work in higher education and, as Barrow (1999) suggests, they tend to operationalise the work of academics so that it is amenable to the development of auditable systems. However, this prevalence of quality systems has nothing to do with the methodology for evaluating them. One does not need a systems approach to evaluate a system. On the contrary, a systems approach will not get beyond the evaluation system to explore the wider context in which it is situated.

The 'politics of quality' mean that the dialectical relationship between quality monitoring and the quality of what is monitored must be located within a holistic context that accounts for structure, history and ideology. To model a system of quality disengaged

from such processes is fundamentally misconceived and colludes with the dominant political agendas rather than regards them as subject matter for investigation.

### ***Critical social research***

Thus, the suggestion is that studies of impact or effectiveness need to develop a *critical social research* approach (Harvey, 1990; Harvey & MacDonald, 1993; Harvey, MacDonald & Hill, 1999).

Evaluating the evaluators is not just a matter of specifying a remit, identifying evaluation criteria and adopting the same methods of investigation as the evaluators adopt in their evaluations. This is frequently the case in agency-initiated evaluations because, in the main, their concerns are instrumental. Rather than simply reproduce the dominant methodology, any analysis of effectiveness or, especially, impact, needs to develop a critical methodology.

That is, a methodology for evaluating evaluators must engage the politics of quality, go beyond any limiting specification, critique ideology, dig beneath the surface, dialectically deconstruct the prevailing perception, question the rationale and reconstruct an alternative understanding of the process.

In his detailed case study of the implementation and impact of quality in a Welsh institution, Jethro Newton (1999) has addressed these kinds of issues. Based on qualitative data from semi-structured interviews with both 'front-line' staff and academic managers, he explored the divergence between the views of 'managers' and 'managed'. He identified a 'policy implementation gap' and argued that situational factors and context are crucial in quality development. The 'success' of a system may be dependent less on rigour of application, than on its contingent use by actors and interest groups, and on how the system is viewed and interpreted by them. Hopefully, the INQAAHE Conference will lead to more contextualised research of this type.

### ***Student experience***

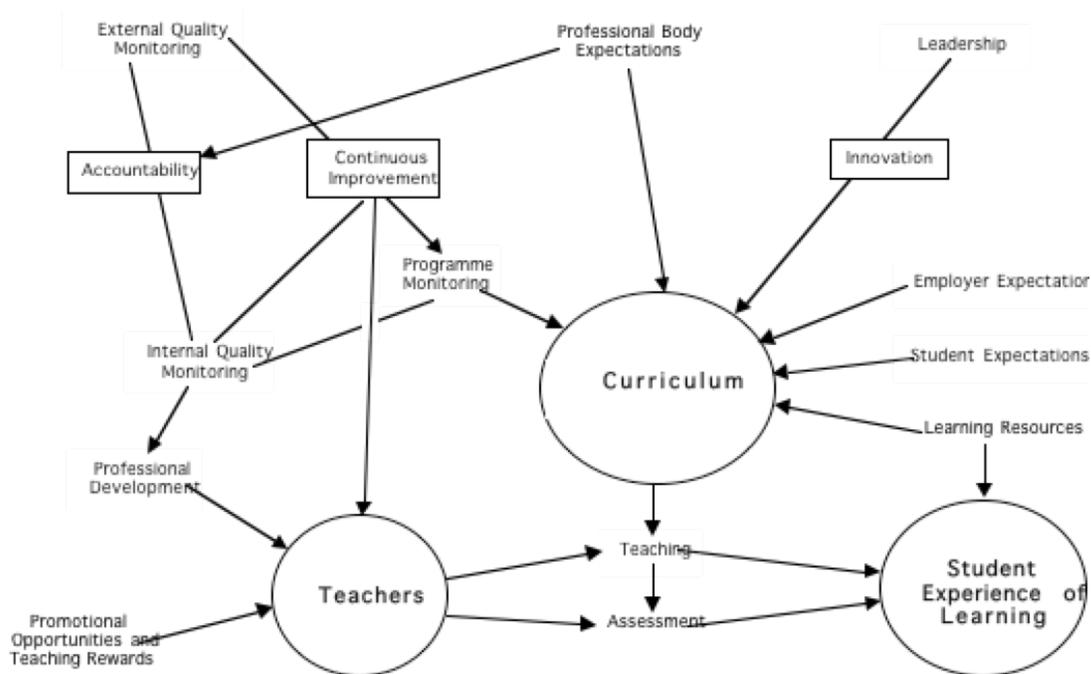
Most effectiveness and impact studies have focused on the affect external quality monitoring has on staff, internal procedures or management structures in institutions. It is far less clear what impact external and internal quality monitoring is having on the student experience. There appears, for example, to be little articulation between quality monitoring and innovation in learning and teaching. Indeed, there are few studies that attempt to address the impact on the student experience. One exception is Margaret Horsburgh's (1998) longitudinal study of the role and importance of external processes in changes to the transformative learning experiences of students. Rather than start from the external processes and explore whether changes recommended have been put in place, or attempt to construct a decontextualised model of the effects of a quality system, Horsburgh approaches from the other direction. She starts by identifying the elements needed for transformation and constructs a framework that structures her observations, interviews and wide-ranging document reviews. On the basis of this, eight key factors are shown to have a significant impact on change management on two degree programmes. These are:

- the curriculum intent and philosophy, along with the programme-development team who determine what this will be;
- leadership;

- members of external approval panels, external monitors of programmes and a professional body;
- the teaching staff and how they teach and assess students;
- the teaching staff and the environment in which they teach;
- programme-specific internal quality monitoring, both formal and informal;
- resource issues;
- student expectations.

Horsburgh evolved a map of the dialectical interrelationship between the factors that impact on degree programmes and the student experience of learning (Figure 7). This is not in any way intended as a causal model but as a map of the complexity of the processes and as indicative of the wider context and state of flux in higher education.

**Figure 7: Factors impacting on the student experience of learning**



Source: Horsburgh, 1999, p. 22

She suggests that there are far more important factors impacting on innovation in learning than external quality monitoring. Horsburgh (1999, p. 23) concludes that, overall, ‘the greatest impact on student learning was the curriculum, factors that influence the curriculum, and the teachers. The most direct impact on student learning was from teacher practices, how they help students learn and the assessment practices they employed’. On the other hand, in an institution that has comprehensive quality systems in place, which have been influenced by external requirements, ‘there was minimal impact of quality monitoring activities on actual learning and teaching practices’.

The factors impacting on the student experience of learning confirmed what many higher education commentators report, that for quality monitoring to have an impact on the effectiveness of higher education, the emphasis must be on the curriculum, learning and teaching. Quality monitoring must focus on more than systems, inputs and outputs, if effectiveness is to be enhanced. (Horsburgh, 1999, p. 23).

## Conclusion

There is a need for more research into evaluation of evaluation. There is scope for research that:

- explores how agencies mediate their brief in practice;
- addresses the micro-politics of evaluation (such as that by Lee & Harley, 1998);
- undertakes rich detailed analyses of implementation and impact (such as that by Newton, 1999);
- starts from the transformative learning of the student and critically deconstructs the politics and ideology of the learning situation (such as that by Horsburgh, 1998).

In particular, there is a shortage of research into the impact on the student experience (Moon, 1999). There is a lack of research that starts from the transformative learning of the student and attempts to unravel an array of interrelated 'factors'. However, such research should not become preoccupied with building and validating static causal models or systems. Rather, what is needed is research that critically deconstructs the politics and ideology of the learning situation. It would be timely to develop an international collaborative and comparative study, based on the approach developed so effectively in Margaret Horsburgh's work.

## References

- Adelman, C. & Silver, H., 1990, *Accreditation: The American experience*. London, CNA.
- Alvesson, M. & Willmott, H., 1996, *Making Sense of Management. A critical introduction*. London, Sage.
- Askling, B., 1997, 'Quality monitoring as an institutional enterprise', *Quality in Higher Education*, 3(1), pp. 17–26.
- Baldwin, G., 1997, 'Quality assurance in Australian higher education: the case of Monash University', *Quality in Higher Education*, 3(1), pp. 51–61.
- 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.
- Barblan, A., 1995, 'Management for quality: the CRE programme of institutional evaluation: issues encountered in the pilot phase — 1994–1995' paper submitted to 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.
- Barrow, 1999a, 'Quality management systems and dramaturgical compliance', *Quality in Higher Education*, 5(1), pp. 27–36.
- Baty, P. & Thomson, A., 1998, 'Brought low by high ambitions', *Times Higher*
- Baty, P., 1998, 'Thames Valley put on report', *Times Higher Education Supplement*, 13 November 1998, p. 1.
- Baty, P., 1999, 'Encounters of an unfair kind', *Times Higher Education Supplement*, 19 March 1999, pp. 6–7.
- Baty, P., 1999b, 'A quality game where cheating is allowed', *Times Higher Education Supplement*, 12 March 1999, pp. 4–5.
- Bauer, M. & Henkel, M., 1997, 'Responses of academe to quality reforms in higher education: a comparative study of England and Sweden', *Tertiary Education and Management*, 3 (3), pp. 211–28.
- Bauer, M. and Kogan, M., 1995, 'Evaluation systems in the UK and Sweden: Successes and difficulties', paper for the the Conference on 'Evaluating Universities', AF-Forum, Rome, 26–27 September, 1995.
- 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.
- Bégin-Heick, N., 1999, 'Impact of program reviews on faculty, and faculty perception of them', paper at the Fifth Biennial Conference of the International Network of Quality Assurance

- Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Bergseth, B., Håggström, I., Karlsson, S., 1999, 'Quality: an issue of both assurance and development', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Breslin, M.B & López, C., 1999, 'Evaluating the accreditation processes: findings and consequences of the study conducted by the North Central Association of Colleges and Schools, Commission on Institutions of Higher Education', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Cave, M. and Kogan, M. (1990) 'Some concluding observations' in Cave, M., Kogan, M. & Smith, R., (Eds.) (1990) *Output and Performance Measurement in Government: The state of the art*, London: Jessica Kingsley, pp. 179–87.
- 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.
- Centre for Higher Education Studies (CHES), 1994, *Assessment of the Quality of Higher Education: A review and an evaluation*. Bristol, Higher Education Funding Council for England.
- Chené, L. & Freedman, B., 1999, 'Evaluating evaluation : a case-study', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Consejo De Universidades Secretaría General (CDU) 1997, *Report on the Assessment of the Quality of Universities in Spain*. Madrid, Ministerio de Educación y Cultura.
- Dearlove, J., 1995, 'Collegiality, managerialism and leadership in English universities', *Tertiary Education and Management*, 2 (2): 161–69.
- European Community (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.
- Frederiks, M.M.H., Westerheijden, D.F. & Weusthof, P.J.M., 1993, 'Self-evaluations and visiting committees: effects on quality assessment in Dutch higher education', paper presented to the 15th EAIR Forum, University of Turku, 15–18 August, 1993.
- Gallagher, A., 1991, 'Comparative value added as a performance indicator', *Higher Education Review*, 23 (3): 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.
- H.M. Inspectorate (HMI) (1990) *Performance Indicators in Higher Education*. A Report by HMI, Reference 14/91/NS. Jan.–April, 1990. London: DES.
- Harvey, L. (1994) 'Continuous quality improvement: A system-wide view of quality in higher education', in Knight, P.T., (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, no. 1.
- Harvey, L. and Knight, P., 1996, *Transforming Higher Education*. Buckingham, Open University Press and Society for Research into Higher Education.
- Harvey, L. and MacDonald, M., 1993, *Doing Sociology: A Practical Introduction*. London, MacMillan.
- Harvey, L. and Mason, S., 1995, *The Role of Professional Bodies in Higher education Quality Monitoring*. Birmingham: QHE.
- Harvey, L., 1990, *Critical Social Research*. London, Unwin Hyman.
- 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. 237–55
- Harvey, L., 1998b, 'Getting 'quality' right: the relationship between quality and learning' introductory keynote presentation at the International Seminar of the Consejo Superior De Educación, Santiago, Chile, 23–25 September, 1998.
- Harvey, L., MacDonald, M., and Hill, J., 1999, *Theory and Method*. London, Hodder and Stoughton.
- Head, P., 1990, *Performance Indicators and Quality Assurance*. Information Services Discussion paper, 4, June 1990. London: CNA.
- Higher Education Funding Council for England (HEFCE), 1995a, *Report on Quality Assessment: 1992–1995*. Bristol, HEFCE.
- Higher Education Funding Council for England (HEFCE), 1995b, *Quality Assessment Report by the HEFCE for Thames Valley University: Modern Languages*, November 1995. Bristol, HEFCE.
- Higher Education Funding Council for England (HEFCE), 1996a, *Quality Assessment Report by the HEFCE for Thames Valley University: Sociology*, March 1996. Bristol, HEFCE.
- Higher Education Funding Council for England (HEFCE), 1996a, *Quality Assessment Report by the HEFCE for Thames Valley University: SLinguistics*, May 1996. Bristol, HEFCE.
- Higher Education Funding Council for England (HEFCE), 1997, *Quality Assessment Report by the HEFCE for Thames Valley University: American Studies*, March 1997. Bristol, HEFCE.
- Higher Education Quality Council (HEQC), 1996, *Learning from Audit 2*. London, HEQC.
- Horsburgh, M., 1998, *Quality Monitoring in Higher Education: A Case Study of the Impact on Student Learning*. Unpublished doctoral thesis, Charles Sturt University.
- Horsburgh, M., 1999, 'Quality monitoring in higher education: the impact on student learning', *Quality in Higher Education*, 5(1), pp. 9–26.
- Hu, Yuehluen & Chen, Han-Chiang, 1999, 'The meta-evaluation of 1997 higher education evaluation in Taiwan', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Inspectie Hoger Onderwijs (IHO), 1992, *De bestuulijk hantering van de resultaten van der externe kwaliteitszorg 1989 in het wetenschappelijk onderwijs*, rapport 1992–8, Zoetermeer, Ministerie van Onderwijs en Wetenschappen.
- 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.
- 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 Beeton, D. (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.
- Lee, F.S and Harley, S., 1998, 'Economics divided: the limitations of peer review in a paradigm-bound social science', in Jary, D. and Parker, M., (Eds.), 1998, *The New Higher Education: Issues and directions for the post-Dearing university*, Stoke-on-Trent, Staffordshire University Press, pp. 185–204.
- 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.
- 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 Lejonald, September.
- Massy, W.F., 1997, 'Teaching and Learning Quality-Programme Review: the Hong
- McInnis, C & Marginson, S. with Morris, A., 1994, *Australian Law Schools after the Pearce Report*. Canberra, Australian Government Publishing Service.
- McInnis, C., Powles, M. & Anwyl, J., 1995, 'Australian academics' perspectives on quality and accountability', *Tertiary Education and Management*, 2 (2): 131–39.
- Meade, P. & Woodhouse, D., 1999, Evaluating the effectiveness of the New Zealand Academic Audit Unit: Review and Outcomes', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Middlehurst, R. & Woodhouse, D., 1995, 'Coherent systems for external quality assurance', *Quality in Higher Education*, 1(3), pp. 257–68.
- Moon, S., 1999, 'The impact of quality monitoring on stakeholders: The case of the University of Central England in the UK', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Murphy, P., 1994, 'Research quality, peer review and performance indicators', *The Australian Universities' Review*, 37 (1): 14–18.
- Neave, G., 1998, 'The evaluative state reconsidered', *European Journal of Higher Education*, 33(3), pp.
- Newton, J., 1999, 'Academics' responses to quality assurance and evaluation: exploring the tension between the views of 'managers' and 'managed'', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Nilsson, K-A & Wahlén, S., 1999, 'Institutional response to the Swedish model of quality assurance', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- Pollitt, C., 1990, 'Measuring university performance: never mind the quality, never mind the width', *Higher Education Quarterly*, 44 (1): 60–81, Winter.
- 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): 223–234.
- Scheele, J.P., 1999, 'Evaluating evaluation in the Netherlands' paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- 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: CNA.
- 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.
- Van Hartingsveld, L.M., 1999, 'Learning by action: concepts and instruments for improving the assessment and diagnosis of educational quality', paper at the Fifth Biennial Conference of the International Network of Quality Assurance Agencies in Higher Education (INQAAHE), Santiago, Chile, May.
- van Schaik, M. and Köllen, E., 1995, *Quality Management At The Hogeschool Holland: towards a policy of systematic quality assessment*. AG Diemen, Hogeschool Holland.
- Vroeijenstijn, A.I., 1995, *Improvement and Accountability, Navigating Between Scylla and Charybdis: Guide for external quality assessment in higher education*. London, Jessica Kingsley.
- Warde, A., 1996, 'The effects of the 1992 Research Assessment Exercise', *Network: Newsletter of the British Sociological Association*, 64, January, pp. 1–2.
- Warren Piper, D. J., 1994, *Are Professors Professional? The organisation of university examinations*. London: Jessica Kingsley.
- Westerheijden, D., 1996. 'A solid base for decisions: effects of the VSNU research evaluations on Dutch Universities' paper to the 18th EAIR Forum, Budapest, 25–28 August, 1996.
- Wilson, B., 1996, 'Quality in universities, Sir Robert Menzies Oration, 4 October 1995', *Journal of Higher Education Policy and Management*, 18(2), pp. 149–58.
- Yorke, M., 1991), *Performance Indicators: Observations on Their Use in the Assurance of Course Quality*, Council for National Academic Awards Project Report 30. London: CNA, January.
- Yorke, M., 1998, Performance indicators relating to student development: can they be trusted?' *Quality in Higher Education*, 4(1), pp. 45-61.

