

# Never Mind the Quality, Measure the Length

## Issues for Lifelong Learning

Paul Armstrong, Birkbeck College, University of London, UK

### Introduction

In outlining the position of quality assurance in lifelong learning, this paper will restrict itself to addressing two issues:

- What is quality?
- How can we measure quality?

Developments in quality assurance in lifelong learning in Britain and elsewhere will be mentioned in passing, but detailed description of the new models of quality assurance impacting on lifelong learning will not be provided here. This is documented and accessible, and the references will be given as they occur in this critical examination of ways in which quality may be measured.

At the outset I should state my own personal position. I am in favour of quality for lifelong learners. Indeed, it would sound rather odd to make the opposite assertion. However, it is worth making this clear from the outset because there has been considerable resistance in Britain to the idea of quality assurance, particularly in the further and higher education sectors.

In brief, this resistance has come about because models of quality assurance have been imposed on the sectors rather than being encouraged to emerge from within. These models have not necessarily been the most appropriate for learning 'industries' - and it has often been an industrial model of quality assurance that has been promoted, with its cult of measurement and a culture of accountability. Moreover, the resistance has been exacerbated by the linking of quality to funding. Having said that it should be made clear that there has been a differential experience here. Funding arrangements have been such that those providers of 'quality' education and training have been rewarded by being given a larger share of the distribution of resources. This means, of course, that those deemed not to be providing 'quality' have been penalised by being given fewer resources. There is an inevitable and increasing divide between those who provide 'quality' provision and those who do not. Those who benefit from this arrangement are often content with the systems in place. Those who suffer from the inequity of the system are those seen to complain or be cynical about quality.

And there are many reasons to be cynical - the bureaucracy, the imprecision and lack of validity of measurement tools, and above all the diversion of resources that could contribute to quality improvement being channelled into the demonstration that quality exists. That this takes up a disproportionate amount of time and

resources, and diverts attention away from the provision of learning leads to the paradox that in meeting the requirements of quality assurance procedures, the quality of learning and the experience of learners are diminished.

But, as this paper seeks to show, this is due to the strategies and systems of monitoring, inspection and accountability, rather than commitment to ensuring that learners receive the best quality service that institutions are able to provide, within the constraints of funding and distribution of scarce resources.

### What is quality?

The application of quality assurance to learning is a relatively recent development. Inspection, monitoring and evaluation and assessment, however, have been around for some time. These are all constituent elements of any system of quality assurance. Increased emphasis has been given to the importance of inspection in recent years as the major means of evaluating formal provision for learning. In the United Kingdom the demise of Her Majesty's Inspectorate (HMI) has been witnessed, replaced under a Conservative Government by the Office of Standards in Education (Ofsted), which has shifted the emphasis from a professional advisory service to a commissioned and privatised inspection service, but still contained within the government department responsible for education. This has been a controversial development that has more clearly delineated the role of inspection and its links with funding and political decision-making about education and training. At the moment this is primarily responsible for schooling, but as will be pointed out later, there is already an influence on lifelong learning, which is likely to be extended in the near future.

The accountability aspect of quality assurance is well represented by the nature of inspection, although it also has an awareness raising function for government policymakers and politicians. There is an argument that would suggest that the reporting mechanism used in school inspections is also vital for providing advice on improvement in terms of both provision and the quality of learning.

But it is the links between funding and inspection that have been of most concern. Education providers deemed to be failing may be taken over by the government and placed in the hands of private companies who, at least in the short-term, would be responsible for ensuring the improvement of both efficiency and effectiveness of the provision of learning. There is very little educational literature about the nature of inspection in learning. According to Wilcox (1992) this is because of a lack of 'a broadly agreed conceptual and methodological base'.

Typically, the definitions of quality give weight in varying proportions to the dimension of efficiency and effectiveness. In the 1980s and throughout most of the 1990s, there was a definite focus on economic efficiency. Inevitably what is required is a trade off between efficiency, economy and effectiveness. Economy is the most simplistic, asking merely 'what does it cost?', which out of context can

be quite meaningless. Reducing the costs of inputs was given primacy in the equation even if it meant a reduction in effectiveness. However, an efficient provision of learning, which required investment in auditing and monitoring and the development of management information systems so that the efficiency could be measured, did not appear to contribute to an upturn in economic fortunes either of a nation nor individuals, nor to the enhancement of the life chances of a significant proportion of the population. If anything it may well have reduced the 'social capital' and life satisfaction of a vast proportion of the population, whose experience of learning was restricted to instrumental skills-based training or learning leading to qualification outcomes, but not necessarily employment.

In considering the balance between efficiency, economy and effectiveness dimensions of quality, it has to be said that this is too simplistic. We can add at least another three 'Es' to the equation (Armstrong, 1999):

- equality or equity
- ethics; and
- environment

It has to be acknowledged that not all learning is intended to lead to economic development. Learning, including training, can be for other intended purposes. There is growing recognition that the assessment of the quality of a learning programme needs to be understood in terms of its specified purpose, as well as in terms of the reasons why learners undertake a course of study. A legacy of the mistaken strategy of defining quality in terms of the acquisition of one quality kitemark or another is the notion of 'fitness for purpose'. In other words, a product or a service may be considered quality if it is effective in meeting its specified purpose. Aims and objectives of learning programmes then become pivotal to the discourse of quality assurance. Did the learning serve its purpose? If so, then it may be considered to be of quality.

However, there are at least two further considerations that reflect the over-simplistic nature of this assumption. The first is that there is often a difference between the specified aims and objectives as seen by the learning providers, and those of the learners who undertake a course of study. In practice this means that a learner may feel satisfied with a learning programme because it meets their needs, whereas measured by the specified aims and objectives, the programme may be considered to be failing. Conversely, a learning programme may be able to demonstrate it has achieved its aims, but leave learners feeling dissatisfied. In other words, official outcomes never tell the whole story. Yet this leads to assumptions being made around 'performance' or 'quality' indicators that may prove to be unwarranted.

The second consideration revolves around whether quality assurance expects providers to satisfy all learners all the time, most of the learners most of the

time, or some of the learners some of the time. The issue here is to do with diversity. It is well recognised that adult learners participating in a learning programme do so for a diverse range of purposes and reasons. The standardisation of the definition of quality then becomes problematic. Issues of equality or equity need to be introduced into the increasingly complex equation.

I have discussed the distinction between equality and equity elsewhere (Armstrong, 1990). There is no space here to rehearse the arguments, but in practice, the distinction has implications that impact on quality. In terms of equal opportunities, the assumption is that we should be able to widen participation to provide equal access for all. However, access in itself is not sufficient, as the current vogue for widening participation is making apparent. If there is to be equality of opportunities for achievement, more needs to be done in terms of providing learning support over and beyond the curriculum. And, importantly, this may well mean providing additional resources for some individuals or groups rather than others. Those with positive prior learning experiences tend to accumulate more. The resources need to be directed at those who previously have had negative experiences of learning. This has implications for determining what constitutes quality. If learning is efficient, but only effective for a minority, then its quality is problematic on the grounds of failing to contribute to equality. There can be no quality without equality.

Quality is in the eye of the beholder. The idea that there can be global agreement on definitions of quality is mistaken. All definitions are invariably situated in a context, and a reflection of the interactions between a range of agencies, including the individual learner whose needs and expectations form part of the equation. The definitions are a cultural product and are underpinned by cultural values. In short, there is always an ideological as well as an ethical basis to definitions of quality. The commitment to learning for a specific purpose reflects ideology, ethics and values. Quality cannot be understood nor defined outside this local frame of reference.

This is not an entirely individualistic interpretation of quality. For even in these post-modern times there are still shared meanings and the fragments of a collective consciousness, communities of beliefs and communities of practice. One example of this is the contemporary awareness of the need for environmental sustainability and the importance of learning in the process. I have observed elsewhere (Armstrong, 1999) that there is case for including environmental sustainability in definitions of quality. If efficiency is at the expense of environmental sustainability, then can it be considered quality?

### **How should we measure quality?**

Unless there is some broad agreement about what constitutes quality in lifelong learning, it is difficult to bring about any degree of standardisation in its

measurement. There are two extreme positions. At one end of the scale, anything goes. Lifelong learning providers can claim that they are quality providers and if necessary provide evidence of what they claim to be quality. At the other end, government intervention may involve the explicit statement of what counts as quality. The argument is for the need for standardisation and comparators between providers, so that we can not only measure their quality but determine which is the best - often succinctly summarised in the form of misleading 'league tables'.

The table below summarises different strategies that have been, and are being, used to measure quality.

**Table 1. Strategies for Assuring Quality**

Focus	How?	Who?	Why?
Fitness for purpose	Specification of <ul style="list-style-type: none"> <li>• Mission</li> <li>• Aims</li> <li>• Objectives</li> <li>• Learning outcomes</li> </ul>	Organisation (self assessment)	To establish market position To derive evidence of success and achievement To ensure customers are satisfied
Industry standards	Specification of minimum standards	(Inter) national industry experts	To ensure that provision meets national or global standards. To achieve a kitemark
Best practice	Benchmarking	Researchers Organisation (self assessment)	Comparisons with those known to provide best practice
Accountability ('Best value')	Specification of performance indicators for efficiency and effectiveness	Central government Funding bodies Auditors Inspectors	To demonstrate value for money To demonstrate effectiveness

This paper will focus on the four different kinds of measurements that have been introduced. A case study from higher education will be used to illustrate the issues that arise from these different strategies.

### **Fitness for purpose**

In accounting for quality in lifelong learning, we have traditionally fallen back on assessment and accreditation of learning. Qualification outcomes have been used as the major critical success factor. However, there has always been something of a contradiction here, for two reasons. The first is whether we can assume that if learners succeed, it is necessarily a reflection of the quality of the teaching or

training providers. A recent experience in training with the introduction of outcome funding or 'payment by results' has been some evidence of a conspiracy to lower standards. It works this way. The government sets National Education and Training Targets, and it is important for them that these targets are seen to be achieved. The funding agencies are given outcome targets to achieve, and they pass these on to the training providers, who need to maximise their income, and do everything they can to ensure their learners achieve the required qualification outcomes. The accreditation bodies, now competitors as well as the upholders of standards, are under pressure to increase their income, which they can do by ensuring their qualifications are seen to be easier to achieve than those of other accreditation bodies. The second contradiction has been experienced in recent years in schools, as they have sought to respond to a whole range of 'raising achievement' initiatives, and they have successfully managed to get more young people achieving higher grades on national examinations. Each year as the pass rate goes up, the more criticism that standards are falling is heard.

Apart from the contradictions inherent in using qualification outcomes as a measure of quality, there is the strong view that accreditation is not the purpose of all learning, and therefore an inappropriate way of measuring quality. This is particularly the case in lifelong learning, where the 'liberal tradition' has not been entirely eradicated by the economic instrumentalism of the radical right in the 1980s. Extra-mural provision had to move to outcome funding, but those outcomes did not necessarily have to be expressed in terms of gaining qualifications. However, there is a requirement on providers to specify what their aims are, how they intend to achieve those aims, and whether the learning has taken place, through the assessment of learning outcomes.

In the UK, in higher education, the models of quality assurance have been changing over recent years, but in the past decade, there has been a definite focus on programme aims and objectives, and the qualification outcomes. More recently, there has been a shift to a focus on learning outcomes. An important development for lifelong learning is that the learners themselves might wish to negotiate the learning outcomes, as well as to claim credit for unintended learning that might well occur.

The change has been reflected in the broadening of quality assurance. Whilst there continues to be a significant focus on the assessment of learning, and the reliability of the assessment procedures, there is also a close check on the validation of programmes, and teaching quality, so that what is learned, and how it is learned can be deemed to be appropriate, if not excellent. The major concern in higher education, however, is the maintenance of standards. Before 1992, the polytechnics and other higher education institutions awarding degrees did so through the Council for National Academic Awards, an accreditation body. The older universities operated their own systems of examination that utilised peers as external examiners to assure the quality of the qualification outcomes. Whilst this

has yet to be seriously challenged, this aspect of quality assurance has been placed within a broader quality framework by the various agencies responsible, now concentrated in the hands of the Quality Assurance Agency (QAA)<sup>1</sup>.

### Standards

The advantages of the 'fitness for purpose' model is that providers of learning can potentially define their own quality criteria: state what your purpose is, how you will achieve that purpose, and by when, and then provide evidence that you have. The phrase 'fitness for purpose' was coined by the British Standards Institute for its BS5750 (now ISO9000) kitemark. The limitations were apparent in the kitemark - the quality inspectors checked that organisations had detailed their procedures for producing a product or providing a service, and then checked that those procedures were being complied with. Whether those procedures were appropriate or whether the product or service was needed, whether they were providing value for money was not a concern. If customers were continuing to purchase the product or service, then the quality must be self-evident (although the influence of the kitemark as a quality insignia, a denotation to the consumer that this is a quality product or service should not be underestimated).

What was missing was the notion of national or international standards that those who make the product or deliver the service should aspire to. In Britain, with the emphasis on competence-based training, national standards have been derived for a vast range of occupations to inform training. These are represented in National Vocational Qualifications (NVQs), supposedly set by industry themselves as minimum requirements for successful implementation of a set of job functions. There is an extensive literature criticising NVQs<sup>2</sup> and the idea of standards that would take us beyond the main focus of this paper. In higher education, there is much rhetoric about maintaining standards, but there have been few attempts to specify what those standards are, to make them explicit and transparent. It is one of the mysteries of professionalisation - the acquisition of the idea of the standards through engagement with higher education. Throughout the world, efforts are being made to specify qualification standards, if only to be able to be more precise about equivalence when students go to study in another country, and want their prior educational achievements recognised. For example, the university qualifications in Australia have developed from their origins in the degrees of older universities that in the nineteenth century were modelled on British universities:

---

<sup>1</sup> This paper is focusing on lifelong learning rather than the total quality of a higher education institution. If it did, this paper would need to also consider the research aspects of higher education. In Britain this is carried out through the quinquennial (approximately) Research Assessment Exercise, undertaken by the Higher Education Funding Councils, together with their new performance indicators for research outputs mentioned briefly later in this paper.

<sup>2</sup> For example, Hyland (1994)

Today, Australian universities, wherever their location and whatever their selected profile, must enable their graduates to operate anywhere, and in any sphere, at a level of 'professionalism' consistent with best international practice, and in ways that embody the highest ethical standards. (Australian Qualifications Framework Advisory Board, 1998, v)

In Britain, the QAA has only recently published its first draft of the National Qualifications Framework<sup>3</sup>, although there have been attempts outside the higher education sector to develop this for nearly 15 years. It was a task for the National Council for Vocational Qualifications (now the Qualifications and Curriculum Agency) when it was first established in the mid-1980s. It was not until the Dearing Report on Higher Education (NCIHE, 1997) recommended that universities too address this issue that any real progress could be made. This requires, at last, higher education to make its standards transparent.

### **Best practice through benchmarking**

Standards are useful in determining the minimum acceptable degree of quality, but the idea of 'meeting standards', as with 'competence', has a connotation of mediocrity. 'Satisfactory' is not always good enough. Learners are rarely satisfied with 'satisfactory', for they are seeking excellence. In the model of Total Quality Management (TQM), there is an assumption that organisations should aspire to excellence. This is not to say that they will necessarily achieve it, but with the philosophy of continuous improvement, there needs to be this aspiration for excellence. Business excellence became the new drive for quality, and it was no longer enough to *satisfy* customers; they had to be *delighted*. With this in mind, the question was how to recognise best practice, and how can it be measured. The proposed solution is through the process of benchmarking.

Benchmarking is not about comparing standards. Benchmarking is sometimes seen as a way of avoiding

reinventing the wheel. To make the most of other people's expertise. It has been described as 'the art of stealing shamelessly'<sup>4</sup> (Owen, 1999, 2)

Owen goes on to define benchmarking as a

---

<sup>3</sup> The draft framework can be found at <http://www.qaa.ac.uk>. At the same site, there are two relevant codes of practice that can be accessed: one on assessment of learning, and the other on external examination arrangements.

<sup>4</sup> As an aside, there is an interesting paradox here. The idea of 'stealing shamelessly' is not approved of in most formal assessment of learning situations. It is labelled as 'plagiarism' and is treated as one of the most serious misdemeanours a learner can engage in, and the quality assurance system needs to be able to identify when this happens. Yet, in many other domains, it is an appropriate means of learning - copying how someone else does something (imitation).

systematic method of improvement that utilises others' good practice to improve your own processes. It can be used as a way of improving any process from ordering paperclips to recruiting staff.

Benchmarking is a method of identifying what must be improved in an organisation, finding ways of making those improvements and then implementing the improvements. It requires an organisation to fully understand its processes and its customers' and stakeholders' needs. From that point it is possible to identify gaps between needs and performance. Once an organisation knows what to improve it can use the knowledge and experience of the organisations it is benchmarking with ... to identify better ways of working. (Owen, 1999, 2)

Described in this way, benchmarking is a research activity, although as Owen says, it is not just taking part in a survey. The identification, description, analysis and evaluation of good practice requires an extensive range of research activities. Moreover, it is supposed to be a typical action research project insofar as it seeks not only to have beneficial outcomes (quality improvements) for the researching organisation, but is of mutual benefit to the organisation being researched. It is also a learning process - learning from the experience of others, and for the clear purpose of quality improvement (and not just an opportunity for educational or industrial tourism).

What are used as benchmarks will vary. Lifelong learning in higher education is likely to be seen in disciplinary or subject terms. Like research, quality in higher education has focused on the teaching and qualification outcomes of subjects. Indeed, in the most recent round of teaching quality assessment, the processes had been renamed as subject review. This contrasts with other learning providers, including schools and colleges, whose quality reviews are based on whole institutional provision. Recent revisions<sup>5</sup> to the QAA approach to quality assurance have been in some ways quite radical, and although it will now be called academic review, the subject aspects have been far from lost. What have been introduced are subject benchmarks. These have come from groups of 'subject experts', invited from institutions that scored highly in those subjects in earlier rounds of teaching quality assessment. Piloted in chemistry, history and law, these have been extended to a further 19 subjects or subject clusters that will be reviewed over the next three to four years, with the remaining 20 subject areas being benchmarked in readiness for the two to three years' time. Their use will be largely through self-assessment. Each institutional provider of the subject will be subject to internal review where they, among other quality assurance activities, measure themselves against these subject benchmarks. At the moment, the subject benchmarks are limited to what undergraduates should be able to know, understand or do having studied the subject for three years' full-time. This has

---

<sup>5</sup> These are outlined in *Higher Education 6* and *7*, available from the QAA. Details of the 22 subject benchmarks are available in the same place. See footnote 3.

some major implications for lifelong learners, who may study at a lower level, even in higher education, part-time, through a range of modes of study (including distance learning), and across a range of multi- or inter-disciplinary subjects. The subject benchmarks may prove to be invalid for use with the diversity of lifelong learning.

There are a number of aspects of this new model which will impinge on lifelong learning, but it is really too early to say too much about them at this time. One point is worth making about the future of lifelong learning in further education. Here benchmarks are being used already, and Owen's manual (1999) is an attempt to support the use of institutional rather than subject benchmarks across the college. This is a positive response by the further education sector, but at a time when the future of the inspection of its quality is a cause for concern. Within the next year there are to be radical re-organisations around the funding and provision of further education and training in England and Wales, with the development of new Learning and Skills Councils, which will be regional bodies, replacing the existing network of Training and Enterprise Councils, who have also had a quality assurance role in monitoring and inspecting training programmes. The new Councils will also be responsible for funding and quality assurance. Among the current proposals is one that there should be a common inspection framework in institutions where adults and 16-19 year olds learn together - an Adult Learning Inspectorate. But there is also a proposal that this should be given to the existing, and controversial, agency responsible for inspecting primary and second schools, Ofsted (Office for Standards in Education), which may well be totally inappropriate for lifelong learning (NIACE, 1999).

### **Performance Indicators**

The idea of a performance indicator has come out of accounting. It is derived from a performance measure, which is a fairly precise tool, used when efficiency, economy and effectiveness can be measured with great accuracy. Performance measures are particularly well developed in manufacturing industry; for example, the number of person-hours it takes to build a car. But the precision of these measures makes them less appropriate for service or people-processing industries, including education. When it is not possible to obtain a precise and unambiguous measure, the term performance indicator is used. Performance indicators are intended to be quantifiable measures of performance in a particular area. In short, they are assumed to tell us something about the efficiency, economy and effectiveness of educational institutions.

Within lifelong learning, the early development of performance indicators in the UK came out of the Audit Commission. Its report *Obtaining Better Value from Further Education* (1985) examined the ways in which 165 of 550 polytechnics (now the 'new' or 'post-1992' universities) and colleges of further education had used their resources since December 1983. The context for the report was

- more than £2 billion (gross) spent on further education by local education authorities (this was before schemes of incorporation that took colleges of further education out of the control of local education authorities, with the development of the national Further Education Funding Council)
- net expenditure after income and grants exceeds £1.5 billion;
- 90,000 staff employed to teach;
- 400,000 full-time and part-time students on courses of advanced further education and 1.7 million on non-advanced courses; in total 900,000 FTE students;
- uncertain economic outlook and collapse of traditional apprenticeships;
- increasing demand, but difficult to forecast;
- increasing proportion of places funded through what was then the Manpower Services Commission (later Training and Enterprise Councils), such as through the Youth Training Scheme (the forerunner of National Traineeships and the New Deal) , and part of the Further Education element in rate support grant to LEAs used to fund Manpower Services Commission courses;
- predicted demographic changes of falling numbers in 16-19 year age group up to 1994.

Against this perceived background, the report made suggestions about ways of improving marketing, tailoring teaching resources more closely to demand better cost recovery, and tighter control over non-teaching costs. In the process, the Audit Commission introduced the notion of performance indicators, meaning such things as links between FE and local schools and employers; attendance and retention rates of students; progression or destination figures; staff-student ratios; grading of staff.

The Audit Commission report focused on little more than simple indicators of efficiency, and led to the publication of the Joint Efficiency Study (JES) report, *Managing Colleges Efficiently* (1987) which represented an attempt to progress measures of effectiveness to counterbalance the emphasis on efficiency found in the Audit Commission report. The JES indicators of effectiveness included:

- Staff-student ratios based on FTE enrolled student and academic staff numbers;
- Non-teacher costs per enrolled FTE student;
- Costs per FTE student enrolled on a course;
- Completion rates for students enrolled on courses, and the cost per FTE student completing a course;
- Rates of target qualifications gained for students enrolled on a course, and the cost per qualified FTE student;
- Rates of employment or progression to further and higher education for students completing appropriate courses.

Indicators of efficiency suggested were

- Staff-student ratios;
- Unit costs linked to appropriate output indicators;
- Non-teacher cost per enrolled FTE;
- Systems for monitoring space utilisation.

Performance indicators relate to both inputs and outcomes.

Subsequent work by the Further Education Unit (FEU) and Unit for the Development of Adult and Continuing Education (UDACE) (later to merge together as Further Education Development Agency) was to give emphasis to processes, as well as inputs and outcomes. The FEU worked with a college in Lancashire to develop an Educational Audit Reference Inventory as a working document to be used to assist in the creation of reliable audit instruments. This was published in *Towards an Educational Audits* (FEU, no date), and added responsiveness and standing to the more obvious measures of efficiency already identified by the Audit Commission and the JES.

In *Performance Indicators and the Education of Adults* (UDACE, 1989), it was suggested that while some of the measures advanced in *Managing Colleges Efficiently* were of direct relevance to any educational provision in the post-compulsory sector, there was still a need to develop specific performance indicators that fitted the purposes of adult continuing education. The UDACE Performance Indicators Project sought to work towards the goal of identifying indicators that reflected the diversity of adult continuing education and the special characteristics of the adult learner. It was not possible at the outset to identify a set of performance indicators that could be used across the whole range of adult continuing education. The project tested different approaches to performance monitoring and determined what was most likely to be applicable and transferable across a variety of settings. They gave overriding importance to effectiveness, and the focus became the assessment of institutional effectiveness; that is, the service's or institutions' ability to achieve the policy goals which have been developed in response to the needs of a specified population. It gave emphasis to achievement and benefits, user satisfaction, economy and efficiency, equity and excellence. Partners in the UDACE Project chose to investigate and develop particular parts of this model. For example, part of the UDACE performance indicators project focused on the extent to which adult learners gain qualifications and accredited success, including credits gained via 'non-traditional' accreditation. It also wished to develop the notion of and ways of measuring 'added value'.

In 1991, UDACE published a document on *Measuring Performance in the Education of Adults* (Powell, 1991). At the same time the annual National Institute of Adult Continuing Education (NIACE) conference in Leicester concerned itself with issues

relating to performance and quality assurance. By this time, there was almost uncritical acceptance of the idea of measurement and performance indicators, particularly as a driving force for change. Three reasons were given for the need for measurement:

- to enable learners (individuals and groups) to understand their existing and newly acquired capabilities and have those accredited, both for personal satisfaction, and to enable them to progress in education, paid or unpaid work;
- to enable providers of education and training to measure their effectiveness in terms which are comprehensible and public;
- to enable those who are paying for education or training (whether learners, central or local government or employers) to make wise decisions about the investment of time and money.

Performance indicators, the UDACE document argued (Powell, 1991, 1), can benefit everyone - users, managers, and staff - because they can help in

- evaluating policy implementation
- assessing whether overall objectives have been achieved
- monitoring plans
- ensuring accountability
- guiding resource allocation
- improving performance
- increasing public support.

The use of performance indicators for comparative purposes, over time or across organisations or services, was recommended by UDACE, including the production of 'league tables' at the macro-level, if identical indicators were to be applied. UDACE was well aware of how performance indicators might be used in misleading ways. For example, the need to set operational standards or targets, if any performance indicator was to be used to compare actual performance against desired performance was recognised; it also recognised that performance indicators have to satisfy the methodological requirements of validity, reliability and standardisation. Very different uses would be made of performance indicators, and it is made clear that they are not merely data for senior managers and policy-makers. We were also reminded that performance indicators were precisely that - indicative not explanatory:

It is only when an indicator suggests that performance is not within the range identified as acceptable that further action or investigation is required. If performance is below the standard agreed managers will need to identify the cause and make appropriate changes. If, on the other hand, performance is above the expected standard, it may be worth investigating how this came about, with a view to replicating good practice elsewhere.

In this sense, performance indicators do not provide answers, but raise questions as to why the set standard has not been met. (Powell, 1991, 5)

The UDACE report quotes an earlier FEU (1986) document approvingly:

performance indicators should not be regarded as precise, sophisticated measures or ends in themselves, but as the basis of subsequent investigation. (Powell, 1991, 6)

Although well aware of the technical limitations and difficulties of the use of performance indicators, the UDACE documents did not contribute to the development of a critique of the idea and use of performance indicators. In the earlier document, they demonstrated their concern that sets of performance indicators should not be imposed, but they built their work on the assumption that some range of measures are required that might constitute a 'basket of tools for the assessment of institutional [or service] effectiveness'. It recognised that 'performance indicators are still regarded with some suspicion by many of those working in adult continuing education', which they saw demonstrated a need for 'cultural change before all adult educators (managers as well as teachers) recognised the value, for themselves and their students, of the use of performance indicators'. This view was shared by others. For example, Peter Mortimore, writing in *The Times Educational Supplement* (25 November 1988) concluded his assessment of the value and validity of performance indicators by saying that 'valid performance indicators, if they can be developed, will benefit the whole education service'. Providing the performance indicators are negotiated or agreed rather than received, and providing the data is valid and reliable, that the system is simple, that the indicators are standardised, and it is remembered that they are only indicators, then the critics would appear to be satisfied.

What was missing was a systematic critique of the notion of performance and its measurement. The language and metaphors were by now so deeply embedded in our policies and practices that we found it difficult to develop and sustain a critique in broader terms. The broader framework needed to consider the political and economic context in which the drive to measure performance was located. It was a climate that stressed value for money, for economic efficiency. But behind this, major and rejuvenated forms of organisational management and control were being encouraged. This has been characterised as the 'rise of managerialism', emerging from Fordism, Taylorism and scientific management. According to Stephen Ball,

the language of management deploys rationality and efficiency to promote control; it is a regime of 'jurisdiction' and 'veridiction'. As a discourse, a system of possibility for knowledge, it eschews or marginalises the problems, concerns, difficulties, and fears of 'the subject' - the managed. (Ball, 1990, 157)

The critique that the shift towards measurement and accountability was part of a 'radical right' thrust to gain more precise and close control over the processes of schooling and education, and that the discourse of management played an essential role in achieving this shift, was very persuasive.

Whilst the critique did not deter the increasing pervasiveness of performance measures and indicators, the higher education sector, for reasons beyond the scope of this paper, managed to resist performance measures in assessing its quality for another decade. It was not until the late 1990s that the idea of performance indicators in higher education began to take hold. In late 1999, the Higher Education Funding Council published performance indicators for 170 universities and colleges. They were the first measures of their kind to be published for higher education. Earlier attempts to produce performance indicators and league tables outside measuring teaching quality and research assessment had been resisted because they were not measuring like with like. It has proved impossible to produce a single league table that consolidates the performance of each university and college. This comes back to the point made earlier about indicators of quality being a trade off.

### Case Study: University of East London

This can be illustrated very neatly by looking at the experience of just one higher education institution - the University of East London (UEL). The four performance indicators used by the Higher Education Funding Council were widening participation, drop out rates, projected learning outcomes, and research outputs and inputs<sup>6</sup>. For brevity, this case study does not intend to look at research inputs, nor the wider aspects of institutional quality, including Research Assessment Exercise and subject review, which a fuller case study analysis would need to encompass.

#### 1. Widening participation

Measured as percentage of young students (full-time and part-time undergraduates) from state schools, or from social classes IIM, IV and V, or from low-participation neighbourhoods, or mature students as a percentage of all students, at each institution in 1997-98.

**Table 2. UEL's performance on widening participation**

<i>Performance indicators</i>	Young people 38% of all FT entrants 8% of all PT entrants		Mature students 62% of all FT entrants 92% of all PT entrants	
<b>Full-time</b>	%	Benchmark +/-	%	Benchmark +/-
From state schools	94	89% +5%		

<sup>6</sup> The data in this section is taken from 'Performance Indicators' in *The Times Higher Educational Supplement*, 3 December 1999; pp i-xii

From lower social classes (IIIM, IV, V)	40	32% +8%		
From low participation neighbourhoods	15	15%. 0%	17	16% +1%
<b>Part-time</b>				
From low participation neighbourhoods	33%	16% +17%	6%	9% -3%

## 2. Drop out rates

Measured by percentage of young, full-time first degree and mature entrants in 1996-97, who had dropped out within the first year.

**Table 3. UEL's performance on Drop Out Rates**

<i>Performance indicators</i>	Young people N= 963		Mature students N=1556	
	%	Benchmark +/-	%	Benchmark +/-
<b>Full-time</b>				
Drop outs one year after entry	15	11% +4%	21	17% +4%
With previous HE qualifications			16	18% ...-2%
With no previous HE qualifications			22	16% +6%

- **Projected learning outcomes and efficiencies**

Measured by numbers projected to complete degree, another award, or transfer to another course or institution, and no award nor transfer, for full-time students starting first degree courses in 1997-97.

**Table 4. UEL's performance on projected learning outcomes**

Number of starters	2376		
% starters who are mature	60		
Degree as projected outcome	55%	Benchmark 68%	-13%
Award as projected outcome	1%		
Transfer as projected outcome	8%		
Neither award nor transfer as outcome	36%	Benchmark 23%	+13%
Projected average time to transfer or complete (years)	Actual 4.2	Efficient 2.9	
Efficiency	Projected 68%	Actual 79%	

So what does all this tell us about the quality of UEL's provision? These figures indicate that the measurement of performance has moved beyond simple rank order on a single index.

The benchmarks should make it possible for UEL to assess its own performance without explicit comparison to external standards or any other institution. A benchmark is initially established sector wide, and then adjusted for each institution to make allowance for local factors, such as entry qualifications, subjects studied. For widening participation, it is the value that the whole UK sector would have if it had the same subject and entry qualification profile as the institution. For projected learning outcomes, the distinction is made between younger and mature learners because those institutions that have more than 40

percent of mature students tend to have benchmarks that are too stringent, and for those with less than 15 percent of mature students the benchmarks may be over-tolerant.

We may still need to know something about the performance of other higher education institutions as measured by the same indicators for comparative purposes. On widening participation, using the percentage of young full-time undergraduates from state schools, we know that Queens University in Belfast did better than expected - 97% came from state schools against their benchmark of 77%. Ulster had a similarly high figure - 95% compared with their 86% benchmark. In Scotland, Glasgow Caledonian achieved 94% against their benchmark of 86%. In England, the best performers were Lancaster who achieved 87% against 79%, and Keele who managed 88% against 80%. Whilst UEL achieved 95%, higher than Glasgow Caledonian and the same as Ulster, its benchmark was also higher at 89%. In other words, UEL's differential was +6%, where as Queens was +20%, Ulster +19%, Lancaster +18%, Keele +8% and Glasgow Caledonian also +8%. In this 'league table' UEL came in the top 26 HE institutions, including six institutions all exceeding their benchmark by 6%. In England alone, UEL came in at eighth equal with five others. In London, UEL was not only the top institution on this indicator, but was one of the very few to exceed their benchmark, with many London institutions under-performing, including the University of London achieving -17% and Imperial College at -18%.

On the other three indicators of widening participation, UEL has a difference of +8% between its benchmark of 32% and its actual percentage (40%) of students from the lower social classes. Across the UK, this is eighth equal with five other institutions. The highest positive differential in participation is Bolton Institute and University of Wolverhampton both with +13%. In terms of representation from low participation neighbourhoods, UEL only broke even on this, achieving their benchmark. Five of the top six exceeding their benchmark were Scottish universities, with the top English representative being Sunderland, which was second overall with +12%, although looking at this indicator for part-time younger students, UEL came out fifth equal with +16%. Sunderland were top of the list for those whose percentage of mature students from less affluent residential areas without a prior higher educational qualification, with a differential of +16%. Again, with +1% UEL was only halfway on the list of 170 institutions,

In terms of student drop out, generally there was a much closer correlation between the actual and benchmark percentages. UEL failed to meet its benchmark by 4% (15% compared with 11% benchmark for younger learners, and 21% compared with 17% benchmark). With most institutions reaching their benchmark, or slightly below it, this puts UEL down in the 'worst than expected' group, with not only one of the highest dropout rates (shared with South Bank and University of North London), but also a significantly large differential with the benchmark. The worst was Napier in Scotland with 18% compared with an 11% benchmark. On

a slightly more positive note, although UEL had a 14% drop out rate for younger full-time degree students in their first year from low participation areas, this exceeded the benchmark by just one percent. Twelve other institutions in England had higher rates than this.

Finally, in terms of projected outcomes, UEL is fourth in the list of 'worse than expected' with a negative differential of 11% (68% of first year undergraduates studying full-time who will eventually get a degree, compared with its benchmark of 79%). The University of North London was similarly placed.

We might consider comparing only like with like. UEL might only wish to compare itself with former polytechnics (the 'new universities') or those that serve similar constituencies. The point is that 'league tables' themselves can continue to be misleading, even with the introduction of benchmarks, which are an attempt to introduce a degree of standardisation or equity in the scoring system. This has happened elsewhere outside of lifelong learning. For example, in looking at schools performance, the league tables are now informed by an 'index of deprivation'. Schools in areas of high deprivation are not competing on equal terms with other schools, and therefore to 'blame' low examination results or poor attendance rates on the school and its staff, without taking into its social and economic context is misleading. The index of deprivation aggregates a number of indicators reflecting the quality of life, including health, housing, environment, crime rates, and unemployment. This index can then be used to weight the scores on performance indicators, that enables allowance to be made of extraneous factors that might impact on the simple percentage score that is supposed to represent an indicator of performance.

The purpose of publishing these performance indicators by the Funding Council was to give information to the public about the performance of universities, to ensure public accountability, and to provide the institution's managers with management information (Bekhradnia, 1999, 16). The Funding Council expects to utilise the information to target improvements, in widening participation, reducing drop out and raising achievement:

But, above all, these indicators ... provide a basis for driving forward the government's agenda for universities. This raises the question of whether these are the right indicators by which to judge universities and that in turn raises the question of what universities are for. The education department [Department for Education and Employment] whose indicators these in effect are, seems to see universities mainly as agents of social inclusion,. Hence the emphasis on attracting and keeping students for under-represented groups and deprived neighbourhoods ... (THES, 1999, 14)

The use of a range of indicators makes the construction of simple league tables based on a single index problematic:

The outcome is a set of indicators so complicated that they will, as the devisers doubtless hope, make it almost impossible for anyone to create unitary league tables, the bane of the university system that is becoming increasingly diverse. (THES, 1999, 14)

This enables institutions like UEL to demonstrate their progress on widening participation, even if they are yet to improve the numbers eventually achieving a degree. The use of league tables that show the University of Oxford and some of the University of London colleges to be performing 'worse than expected' on widening participation may go some way to redressing the balance.

The benchmarking system means, for example, that high-prestige universities setting high entry grades are not damned for failing to take in as many students from under-privileged backgrounds as universities with more open entry. And those with more open entry are not expected to graduate so high a proportion. (THES, 1999, 14)

But the performance indicators and benchmarks do need to be treated with a degree of caution: 'experts are questioning the basis on which the indicators have been calculated':

*Maggie Woodrow, author of *From Elitism to Inclusion: Good Practice in Widening Access to Higher Education*, criticised the calculation of the benchmark figure, which does not condemn universities with high entry qualifications and certain subject mixes for recruiting fewer students from lower socio-economic groups. 'We should ask whether the funding councils should look at the benchmarks again: is an institution's subject range a legitimate excuse for keeping out the lower social classes? The sector benchmarks imply that it is okay for some universities to do more [on social exclusion] than others', she said. (Goddard, Thomson and Wojtas, 1999, 4)*

One suspects, however, that research will remain the basis for judging the quality of higher education, rather than the contribution to lifelong learning:

Universities are about advanced learning, scholarship and research. They are necessarily selective and meritocratic. The present set of indicators, largely ignores these broader academic purposes. They should be treated as what they are: a measure of how well universities are meeting government requirements. They should not be regarded as measures of universities' general performance as centres of academic excellence. (THES, 1999, 14)

## References

- Armstrong, P (1990) 'Extending Equality: Equity and Entitlement' in Pam Kleiber and Libby Tisdell (eds), *Proceedings of 31<sup>st</sup> Annual Adult Education Research Conference*, University of Georgia, Athens, May, pp 1-6
- Armstrong, P (1999) 'Sustaining commitment to social responsibility: renewing the ethical basis of the quality agenda in the education of adults'. *Proceedings of 40<sup>th</sup> Annual Adult Education Research Conference*, Northern Illinois University, DeKalb, 23-25 May; available at <http://www.educ.ubc.ca/edst/aerc>
- Audit Commission (1985) *Obtaining Better Value from Further Education*, London, HMSO
- Australian Qualifications Framework Advisory Board (1998) *Implementation Handbook*, 2<sup>nd</sup> ed. Melbourne, Australian Qualifications Framework Advisory Board. Available at <http://www.curriculum.edu.au/aqfab.htm>
- Ball, S (1990) 'Management as moral technology: a Luddite analysis' in S. Ball (ed), *Foucault and Education: Disciplines and Knowledge*, London, Routledge, pp 153-166
- Bekhradnia, B (1999) 'Tables for our times', *Times Higher Education Supplement*, 3 December, p 16
- FEU (Further Education Unit) (1986) *Educating the Auditors: Feasibility Study*, FEU/ACFHE
- FEU (Further Education Unit) (no date) *Towards an Educational Audit*, London, FEU
- Goddard, A, Thomson, A and Wojtas, O (1999) 'How did you perform?' *Times Higher Education Supplement*, 3 December, p 4
- Hyland, T (1994) *Competence, Education and NVQs: Dissenting Perspectives*, London, Cassell
- Joint Efficiency Study (1987) *Managing Colleges Efficiently*, London, Department of Education and Science
- Mortimore, P (1988) *The Times Educational Supplement* (25 November 1988)
- NCIHE (National Committee of Inquiry into Higher Education) (1977) *Higher Education for a Learning Society* (The Dearing Report), London, HMSO
- NIACE (National Institute of Adult Continuing Education) (1999) 'Learning and Skills: The Prospectus and the Bill - England. An early NIACE Briefing'. Available from <http://www.niace.org.uk>
- Owen, J (1999) *A College Guide to Benchmarking*, London, FEDA
- Powell, B (1991) *Measuring Performance in the Education of Adults*, London, UDACE
- THES (editorial) (1999) *Times Higher Education Supplement*, 3 December, p 14
- UDACE (Unit for Development of Adult and Continuing Education) (1989) *Performance Indicators and the Education of Adults*, London, UDACE
- Wilcox, B (1992) *Time-Constrained Evaluation: A Practical Approach for LEAs and Schools*, London, Routledge
- Paul Armstrong (July 2000) 'Never Mind the Quality, Measure the Length', Working Papers of the Global Colloquium on Supporting Lifelong Learning [online], Milton Keynes, UK: Open University. Available from <http://www.open.ac.uk/lifelong-learning> [Accessed: (insert date)]**