

INTERNATIONAL PERSPECTIVES ON PERFORMANCE INDICATORS

In March 2004, CHET held a two-day seminar in Cape Town to discuss the proposals contained in Chapter 6 of this book (CHET 2004a). To set the context for discussions at the seminar, a number of the international participants were invited to offer accounts either of recent work they had undertaken on performance indicators or of developments that were taking place in their countries and regions. The presentations of Maurice Kogan and Frans Kaiser are reproduced in full below, and summaries are offered of the presentations of participants from three African countries.

MAURICE KOGAN: THE DEVELOPMENT AND USE OF PERFORMANCE INDICATORS AND PERFORMANCE MEASUREMENT IN THE UK

The search for appropriate performance indicators is part of a massive change in assumptions about ways of governing higher education, and indeed all of the other professionally led public services and institutions. Indulge me whilst I go back two or three decades in the history of higher education as perceived by a practising academic. In the 1970s, in the UK universities, we believed that we had a mandate to deliver education and perform research according to our own professional lights. The university was essentially a holding company for expert groups who, in doing their thing, were doing society's thing. The state had no doubt about the rightness of letting us get on with it, within broad resource limits, and we were allowed a corporatist bargain in which we delivered education and knowledge in return for freedom and its essential wherewithal.

By the mid-1980s, academics, however, were having to face the notions first put up by a few cranky vice-chancellors, accountancy professors, applied economists and Treasury officials, that our performance should be measured. This was reinforced by a

compliant committee – the Jarratt Committee – which wanted to turn us into corporate entities led by chief executives.

In the UK the performance indicator policy was seen as part of Margaret Thatcher's determination to bring the professions to heel. The performance indicator movement reached a kind of crescendo in the 1980s. There had been a prime concern with the potential of performance indicators to aid funding bodies or government in the allocation of resources in accordance with overarching goals, or to call institutions to account for the economic and efficient use of resources. But then they began to lose ground and were treated as indicative material which might be used by choice or, by the mid-1990s, to enable individual institutions to assess their own performance and act upon that assessment. Now performance indicators seem to have sprung back into institutional evaluation. Universities are grouped by the funding council into families based on their leading characteristics, e.g. universities with heavy engineering course commitments are grouped together, and then judged by such benchmarks as retention of students, participation and numbers from disadvantaged backgrounds.

Performance indicators, however, are perhaps a less important part of performance review than the assessment of teaching by a Quality Assessment Agency and the research assessment exercise. Both are critical for reputations and for research resources, but are not in essence performance indicator systems. Performance indicators can be referred to by assessors but they are simply data on which they can draw. Their use in institutional evaluation is not the sharp end of what has been in the UK an extremely sharp evaluation movement.

All of this takes us a long way from the cherished myth of British higher education that it was free of state interference and that its governance was in the hands of academics. I was long a sceptic about the utility of performance indicators. I'll articulate that scepticism again later. But the facts of higher education development were against the critics. Part of my grudging concession to them was the result of a visit by a new vice-chancellor to my department. A distinguished engineer, he had little time for much of what was going on. In his first year he visited no department, but let us stew. He then came in with a sheaf of papers and began. "You are a rather old lot, and therefore expensive." No one had said that to us before. But he had hit on the point, that we were rather stable – more of us ought to have moved on and made way for newcomers. He then said, "You publish a lot of books. Why so few articles?" Again he had penetrated the point that much of our stuff was in an historical rather



than a theoretical political science mode. So performance indicators had their uses domestically.

They obviously have an important place in the larger scale of things. Our system had expanded enormously. From 45, we now have 105 universities or 150 institutions of university status. What had been a loose assembly of free institutions was now to be treated as a system. Quality issues became sharp – were all of these places producing degree qualifications that measured up? And at the same time higher education became the target of social and economic policies, in which universities were to behave like profit-making corporations, disadvantaged members of society to be given full opportunities, and Britain made ready to compete with other contenders in the Knowledge Universe. This was not to be left to chance: an enormous apparatus of monitoring was established to make sure we were up to snuff, and to see that taxpayers' money followed what were thought to be the routes of most advantage.

Thus ended the period of trustfulness between universities and the state. If the UK needed monitoring systems, and evaluative data to go with them, it would follow that countries undergoing far more drastic reforms than we will need to be able to know where they are starting and where they are going. The problem for you (in South Africa), as it was for us, is to make sure that the monitoring devices do not bring unintended and undesirable consequences. There is a branch of engineering known as non-destructive testing. Have we got one?

If performance indicators or similar instruments are an unavoidable part of our collective machinery there are cautions that need to be observed. Some cautions have been sounded previously, but let me reiterate them.

I will follow three themes:

- ▶ The political, or, if you like, the ideological background to performance indicators, and related devices such as benchmarking and outcome measurement.
- ▶ The educational models favoured by their use.
- ▶ Some interrogatory critiques of performance indicators.

THE POLITICAL-IDEOLOGICAL BACKGROUND TO PERFORMANCE ASSESSMENT

In the UK the performance indicator policy was seen as part of Thatcher's determination to bring the professions to heel. In the UK, at least, there was a kind of

schizophrenia about the role of the state. The conservatives wanted to roll back the state, but in fact imposed degrees of centralisation and evaluative requirements unimaginable decades before. Universities had to use furniture vans to submit university research assessments to the funding councils. The bulk of senate papers at one time were photocopies of central government documents.

They were part, too, of a demand for performative systems which advanced the notion that all things and activities must have a use and results. A further thesis might have been that since higher education was not then a marketed good, performance indicators would provide information that would be the equivalent of prices in the market. They thus gave strength to customer pressures. Immediate utility had never been a classic higher education objective, from Newman onwards.

At the centre of the strategy was a commitment to neo-Taylorian concepts of management and the drive to ensure that clear objectives were set against which performance could be easily measured and resource allocation policies linked with that performance. Establishing performance indicators, reducing complexities of judgement to simple measures, was part of that.

However, at various points in the 1980s different governments were moving towards more complex concepts of the state, in which the market would assume a more central role and the structures and mechanisms of the public sector would be modelled on those of the private sector. Management continued to be seen as a key part of the strategy, but in the form of New Public Management, with its emphasis on decentralisation, organisational missions, contracting, income generation and the management of multiple objectives. Paradox and ambiguity, such as “centralised decentralisation”, became part of the language of management. Public as well as private organisations were persuaded that their survival depended on maximising not only economy and efficiency, but also quality. Evaluative institutions and managements themselves began to place more emphasis on self-evaluation and quality assurance.

The institution gained power as that of the academic base declined. Structures under pro-vice-chancellors, directors of quality and other actors in a kind of morality play, increasingly cut across the traditional dean-head of department-individual academic lines. Central administrators instructed departments on how to complete the evaluative exercises set by the national evaluative bodies. This gave them purchase on matters hitherto exclusively the domain of the academics.



Academics still retain a great deal of freedom, but the drive to performance meant that money for research was constrained by the pre-set demands of central government, often aided by research councils. This, incidentally, sometimes gives us the problem of “immature customers”. The curriculum can hardly fail to take notice of the benchmarks which are set up for each subject by national groups of academics. So far, benchmarking has not produced complaints of politicisation of the curriculum, but it must certainly constrain the creative and maverick teacher.

The ideology behind this is mixed. Part of it is the quest for quality. No doubt an expanded system, now accepting less well qualified students, has to change its teaching and learning ways and teach to the text rather than in the heat of individuals’ research interests. So mass markets mean mass production and mass monitoring. A second objective is marketisation, and this is signalled by the pervasive demand that both teaching and research demonstrate their market relevance.

The dominant model for the development of performance indicators is a production model linking input, process and output. Within this, emphasis may be laid primarily on input or output or the relationship between these two. It has a strong conceptual link with the objectives of economy (reduction or control of resources) and efficiency (the productive use of resources). It may encourage thinking of institutions as production systems, whose prime task is to convert inputs into desired outputs or add maximum value to specified inputs. It might therefore be consistent with promoting relative concepts of quality, but at the same time, used in the context of teaching, suggests the assignment of passive or inert roles to students in higher education institutions.

The third objective is a construction of managerial efficiency. Not only institutional review, but the two forms of academic review give ministers and institutional leaders data through which they can know a lot about institutions and their components. They know, and can therefore control, the inputs so as to correct faults and reward virtue, and can see that outcomes are improving. One recalls that in this, “words are fat and figures are thin”. System managers can play with averages, means, bottoms and tops. They are backed by an arithmetical epistemology. This visibility is a valuable managerial asset. One recalls deans in Lithuania who decline to let their universities, let alone the ministry, have key data – it is their property.

THE EDUCATIONAL MODELS FAVOURED BY PERFORMANCE MEASURES

Performance indicators have some interesting sisters, such as outcomes analysis (in terms of learning gains), targeting and benchmarking. Do they fit good models of curriculum development? A traditional model would start with individual teachers internalising their own subject expertise before expressing it for communication. It is then tested and changed through interaction with students and critical colleagues. An alternative proposed to the official UK model was to check on what outcomes have been achieved, but not be required to build on them. Or to talk to the students to find out what has been achieved.

The outcomes-driven alternative starts from externally proposed benchmarks and specifies outcomes. But some of the most important learning outcomes will emerge from the process of teaching and learning, and in a lively academic community will not remain as constants. Moreover, inputs and outcomes for individual students should be various. Thus, in historical studies, some students will achieve excellence, or competence, through their ability to interpret and assess the validity of great detail. Others will deal best in grand themes. So will their teachers. If this eclecticism is a desired product of higher education, what good is it to render the whole to pre-conceived outcomes and then attempt to demonstrate how they are achieved? Unless, that is, we are talking about minimal basic skills, which may not be the business of higher education anyway.

“Subject benchmarks are statements which represent general expectations about standards for the award of qualifications at a given level in a particular subject area.” Reviewers are to use benchmark statements to determine whether outcomes are appropriate. Is it objectionable to move towards national norms for each subject area? Some thought there could be no real objection to common benchmarking. There were elements of disciplinary formation with which all students should be acquainted. Not a big deal anyway.

There have been objections raised, however. It is difficult to avoid the lowest common denominator and to avoid being prescriptive. Outcomes and benchmarks are laid out in prescriptive form with thresholds and anticipated levels, and this does indeed encourage simply ticking them off as tiresome requirements rather than as developmental and quality assurance aids. “We had said to ourselves that we would produce a broadly-based social science education, but then found that we had to accord to benchmarks. We would have encouraged critical thinking and built blocks



from there. Now we have to think about compliance to something else.”

Another objection concerns the educational model that benchmarks imply. In the classic model of the teacher’s role in student learning, we absorb and internalise knowledge from reading, research and other academic activities which we then recreate for transmission. We have assumed that there will be outcomes beyond the interest of students in acquiring the knowledge and skills associated with the recreation of knowledge that we present them. Because students in a three-year sequence will probably take between nine and twelve courses led by different teachers, they will have been in contact with a large number of teacher perspectives as well as those from a wide range of written sources, and can choose the values, perspectives and skills from that repertoire on which to build their own intellectual identity. All of that is pre-empted by benchmarking and outcome measurement of learning gains based on it.

SOME INTERROGATORY CRITIQUES OF PERFORMANCE INDICATORS

Returning, finally, to performance indicators, it is obvious that no judgemental statement is innocent or purely technical. It entails a statement of underlying value, and a judgement based on it. The selection and use of performance indicators are significantly affected by the broader policy context, the nature of the state and the resource allocation mechanisms embedded within that. The larger contexts in which the various forms of interplay take place are not eternal verities, but are also in transition. Performance indicators have been forged in higher education systems in which resource allocation is either centralised or market-based. They thus by-passed the era of collegial or peer control. Now we have allocative systems in a changing political economy, characterised by multiple stakeholders, of which government is one.

The contexts help form their uses. Are they to be used summatively, to determine or inform decisions or judgements, or formatively, with a view to challenge, interrogation, dialogue, reflection or diagnosis? As monitoring instruments with a view to maintaining a system or assuring an agreed level of quality, or as instruments of presentation with a view to marketing or attracting investment or support or recruitment?

Because they are complex and admit of multivalence, we do well to interrogate them by a range of criteria, as follows:

- ▶ **Type of indicator:** Is the indicator a measure of input, output, productivity or final outcome? For example, numbers of graduate students in natural sciences can be an input – they do much of the work – or an output, adding to the stock of trained resources. Is it a simple quantity index, or is it adjusted for quality?
- ▶ **Relevance:** How accurately does the performance indicator measure true underlying performance, relative to the organisation’s given objectives? Put another way, can we say on looking at some of the performance indicators that they tell us what South African higher education is like?
- ▶ **Ambiguity:** Is it possible to identify a high or a low value of the indicator as unambiguously favourable or unfavourable? For instance, a high cost per student can be interpreted either as a sign of wasteful extravagance or of the giving of more individual attention.
- ▶ **Manipulability:** If a performance indicator can be manipulated by the individual or body which it is intended to assess, its value is reduced. For example, numbers of research students can be inflated if selection is indiscriminating. This also applies to degree classes.
- ▶ **Cost of collection:** Some performance indicators can be calculated readily from data already available to institutions or the higher education system. Others require costly and perhaps unobtainable data collection.
- ▶ **Level of aggregation:** Each performance indicator has its own natural level or level of aggregation – individual, department, discipline, institution or the higher education system as a whole. They may thus be correspondingly difficult to bring together or measure.
- ▶ **Relation to other indicators:** Often several indicators are used to measure the same or a similar aspect of performance. In such circumstances, the existence of multiple indicators is a useful consistency check, but these variables should not be regarded as independent in an overall evaluation. In other cases a single indicator – for example, number of research students – measures several dimensions of activity.

USING PERFORMANCE INDICATORS

The multivalence is brought into play if one asks: “How far are values shared between different levels and different stakeholders in the system? How and how easily are different values (for example, public accountability and quality improvement or efficiency and quality or accessibility and rigorous standards) reconciled?” And do



different modes of reconciliation emphasise the role of different levels or actors in the development and use of performance indicators?

Values are unlikely to be stable or coherent across higher educational systems. In practice the challenge will be the management of performance indicators when values are in conflict with one another. Higher education institutions are widely assumed to require a level of independence, self-determination and reflection, and they value criticism if they are to fulfil the demands placed on them.

Up to a point, the multiplicity of values required of higher education may be accommodated through a differentiated system, in which institutions can carve out their own distinctive missions and devise performance indicators to match them. But this is unlikely to provide more than part of the answer where higher education is conceived as fulfilling important functions on behalf of the state and is part of a general funding system. It is important to think about what degree of consistency between indicators it is essential to establish through the system and how this is to be achieved.

The importance of self-activation and motivation in higher education suggests that a premium needs to be placed on establishing structures for maximising consent and co-ordination in the selection and use of performance indicators. As yet, the evidence is that new measures incorporating more diverse criteria have not been institutionalised and that the dominant forms of evaluation are those of the state rather than those of the market. The main influence on the UK quality assessments is that of peer review. The role of performance indicators in the processes is either negligible, as in the case of teaching quality, or unclear, as in the case of research assessment. In each case the criteria adopted appear overwhelmingly to have been those of high academic excellence and reputation, despite policy commitments in the areas of research and teaching to more diverse concepts of quality.

Reaching consensus performance indicators for the quality of teaching has proved particularly difficult. The dual status of performance indicators as managerial tools designed for institutions' use and as public expressions of relative performance, readily convertible into league tables, makes the design of measures a highly political process in which compromises have to be reached. And those indicators, which embody consumer criteria or commitment to transformative or value-added concepts of quality, have been shelved on the grounds of the need for more research. State mechanisms and peer review seem to be holding their own against the grosser

intrusions of the market, consumerism and externally imposed quantitative measurement. At the same time, the difficulties of objective measurement and of fair allocative systems based on them have been exposed.

I return to my general theme. Performance indicators are an inevitable outcome of changes in systems towards mass recruitment, marketisation and the demands for managerially-led efficiency. They embody far more than techniques; they entail value preferences, although they cannot avoid ambiguity. So we do well to benefit from the questions they raise but need to use them cautiously – as “can openers, not dials” and alongside qualitative evaluation which can remedy some of their faults.

FRANS KAISER: THE ROLE OF INDICATORS AND BENCHMARKS IN THE EUROPEAN UNION

Globalisation, competitiveness and transition into the knowledge-based economy are some of the buzzwords that have stimulated national and international policy-makers to formulate great ambitions. Many policy-makers see the new dynamics as an opportunity to move forward and upward in the world. To achieve those ambitions, new steering mechanisms such as market-type mechanisms and contracts are explored, since traditional mechanisms like regulation and planning do not succeed any longer in co-ordinating and steering the behaviour of the actors in the field in accordance with the ambitions.

In this paper I shall focus on the situation in the European Union (EU) where, in 2000, a major ambition was formulated. What steering mechanism does the EU use, and how is the performance regarding the ambition assessed? The focus will be on indicators and benchmarks: how are they used, how are selections made? How to evaluate the selections made and how to improve them are leading questions I shall discuss in this paper.

THE LISBON AMBITION

Some four years ago, in Lisbon, the EU heads of state and government set a very ambitious aim for the EU to achieve by 2010. By that year the EU should be:

... the most competitive and dynamic knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion.



The role of education and training was considered to be crucial to the realisation of that ambition: without a high quality education and training system it is impossible to make the transition into a knowledge-based society and to further develop the knowledge-based economy.

Following the general Lisbon ambition, the Council of Ministers for Education in 2001 agreed to three strategic goals for European education and training systems:

- ▶ Improving the quality and effectiveness of education and training systems in the EU.
- ▶ Facilitating the access of all to education and training systems.
- ▶ Opening up education and training systems to the wider world.

These very broad strategic goals for European education and training systems were refined in 13 associated objectives. These objectives were presented in the “Detailed Work Programme for Monitoring the Progress Towards these Strategic Goals until 2010”, which the education ministers adopted in 2002.

The 13 objectives of the Detailed Work Programme are:

- ▶ Improving the quality of teachers and trainers.
- ▶ Developing skills for the knowledge society.
- ▶ Ensuring access to ICT for everyone.
- ▶ Increasing the recruitment to scientific and technical studies.
- ▶ Making the best use of resources.
- ▶ Creating an open learning environment.
- ▶ Making learning more attractive.
- ▶ Supporting active citizenship, equal opportunities and social cohesion.
- ▶ Strengthening links with working life and research and society at large.
- ▶ Developing the spirit of enterprise.
- ▶ Improving foreign language learning.
- ▶ Increasing mobility and exchanges.
- ▶ Strengthening European co-operation.

These objectives cover the whole education system, but the focus is mainly on primary and secondary education.

A NEW STEERING MECHANISM

The Lisbon ambition changed the relation between the EU and its member states. Up until 2000, the EU had a minor role in education policy. Its main instruments in (higher) education policy were the mobility programmes and the large research programmes. The Maastricht Treaty of 1993 stated firmly that education policy was a national responsibility and the EU should not interfere. However, since education was considered crucial to the realisation of the Lisbon ambition, there was a need for a clear role of the EU in the formulation or at least co-ordination of education policy. Since the Maastricht Treaty did not allow the EU to develop and use firm steering mechanisms, the EU Commission developed a new steering method that respected the subsidiary principle and at the same time could also lead to a convergence of national education policies along the lines the EU Commission had set. This has become known as the Open Method of Co-ordination.

This method involves:

- ◆ Fixing guidelines for the Union, combined with specific timetables for achieving the goals which they set in the short, medium and long terms.
- ◆ Establishing, where appropriate, quantitative and qualitative indicators and benchmarks against the best in the world and tailored to the needs of different Member States and sectors as a means of comparing best practice.
- ◆ Translating these European guidelines into national and regional policies by setting specific targets and adopting measures, taking into account national and regional differences.
- ◆ Periodic monitoring, evaluation and peer review organised as mutual learning processes.¹

The Open Method of Co-ordination is an instrument for identifying best policy practices, using the diversity of policy approaches in European countries as a grand reservoir of ideas for possible policy measures to achieve the agreed-upon objectives or outcomes. Crucial to this method is the use of indicators and benchmarks. This information has to bring national governments to change their national policies to achieve the common goals. The “peer pressure” will stimulate governments to look for best practices and to learn from those cases. This will eventually lead to a convergence in national education policies.

¹ *Conclusions of Lisbon European Council 23/24 March 2000 – paragraph 37.*



THE ROLE OF INDICATORS AND BENCHMARKS WITHIN THE OPEN METHOD OF CO-ORDINATION

The use of indicators for monitoring progress in the follow-up to the Lisbon conclusions is inherent to the process. Indicators and benchmarks are needed to make progress easily visible and to break down the overall ambition into achievable goals in different policy areas.

Every spring, the Commission reports on the progress made towards achieving the Lisbon ambition using a framework of structural indicators (involving 42 indicators in 2003).² The part education and training takes in this report is surprisingly small. Only four of these indicators are specifically relevant for education and training. These indicators cover spending on human resources, lifelong learning, science and technology graduates and early school leavers. This is partly due to the fact that the framework for monitoring progress in education and training systems – the Detailed Work Programme – was adopted in 2002 and the first progress report on that field was published in early 2004. However, this does not mean that indicators and benchmarks are not important in the education and training field. In the following section I shall elaborate on that role.

THE DETAILED WORK PROGRAMME

In the Detailed Work Programme (DWP), objectives were formulated (13 in total) for each of the strategic goals and for these objectives an “indicative list” of 33 indicators was proposed. The Commission also proposed a standard format to be used for measuring progress within the 13 objective areas (see Figure 1).

FIGURE 1: Monitoring performance and progress regarding education and training within the Open Method of Co-ordination

| INDICATOR | PRESENT LEVELS | | PROGRESS | | BENCHMARKS | | |
|-----------------------------|--------------------------|---|---------------|------------|------------|----------|----------|
| INDICATOR DEFINITION | Average performance (EU) | Average of 3 best performing countries (EU) | USA and Japan | Up to 2004 | Up to 2010 | For 2040 | For 2010 |

² See *Communication from the European Commission on “Structural Indicators” COM (2003)585 – 8 October 2003.*

On the basis of chosen indicators for each objective an interim report foreseen in 2004 and the final report foreseen in 2010 will include an evaluation of progress made. Where feasible, European-wide benchmarks could be set by the Council.

Clearly, indicators and benchmarks are considered to be important in monitoring progress, but what are indicators and benchmarks and how are they defined and used by the EU?

INDICATORS

There is no single authoritative definition of an indicator. Johnstone described it rather vaguely as “something that conveys a general impression with more or less exactness: something that gives a broad indication of the state of the situation investigated” (Johnstone, 1981:3). The discussion of definitions is further complicated by the introduction of “performance indicators”. Cave defined a performance indicator as:

... a measure – usually in a quantitative form – of an aspect of an activity of a higher education institution. The measure may be either ordinal or cardinal, absolute or comparative. It thus includes the mechanical applications of formulae and can inform, and be derived from, such informal and subjective procedures as peer evaluations or reputational rankings. (Cave, 1997:24)

What distinguishes a performance indicator from an “ordinary” indicator is that the former is used as an assessment of goal-attainment.

Following these definitions, the 29 indicators that the Commission now uses can be considered to be performance indicators: they are used to monitor progress toward the goals.

However, the Commission has stressed that indicators should not be considered only in their capacity for measuring progress. Indicators should function mainly as a basis for a constructive dialogue and exchange between Member States, as a tool to understand the reasons for differences in performance, so that other countries can learn from policy practices adopted by the most successful countries. Therefore, indicators can be used as an instrument for stimulating the exchange of good experience and new ways of thinking about policy approaches – which brings us to benchmarks.



BENCHMARKS

Benchmarking is an ongoing, systematic process for measuring and comparing the work processes of one organisation to those of another, by bringing an external focus to internal activities, functions or operations. Core questions in benchmarking are: How well are we doing compared to others? How good do we want to be? Who is doing it the best? How do they do it? How can we adapt what they do to our system? How can we be better than the best? The processes or organisation of another organisation or system are used as a point of reference (best practice).

In the context of the DWP, “benchmarks” function as reference points for where the EU wants to be in 2004 and in 2010. They point to areas where special policy efforts are necessary to improve education and training in Europe.

In May 2003, the Education Council selected five European benchmarks for the improvement of education and training systems in Europe up to 2010:

- ▶ By 2010, an EU average rate of no more than 10% early school leavers should be achieved.
- ▶ The total number of graduates in mathematics, science and technology in the EU should increase by at least 15% by 2010, while at the same time the level of gender imbalance should decrease.
- ▶ By 2010, at least 85% of 22-year-olds in the EU should have completed upper secondary education.
- ▶ By 2010, the percentage of low-achieving 15-year-olds in reading literacy in the EU should have decreased by at least 20% compared to the year 2000.
- ▶ By 2010, the EU average level of participation in Lifelong Learning should be at least 12.5% of the adult working age population (25–64 age group).

These European “benchmarks” are not concrete targets for individual countries to be reached by 2010. They are defined by the Council as “reference levels of European average performance”. However, in a number of member states (including Austria and the Netherlands) those European reference levels are set as targets for national education policy.

THE SELECTION OF INDICATORS AND BENCHMARKS

The selection of indicators and benchmarks is made within the context of the strategic goals and the 13 objectives. Within this political, normative frame, one may expect that the choice and definition of indicators is more or less objective and follows “technical” criteria rather than political ones. What are those “technical” criteria?

In the literature on (education) indicators there is a long list of criteria that indicators have to meet. These criteria not only refer to practical or technical aspects of indicators: indicators have to be used, which brings in political, or in more general terms, normative considerations regarding the choice and presentation of indicators.

- ▶ **Quantitative:** The phenomena to be indicated should be quantifiable. This does not mean that all indicators have to be measured on the ratio-level. As Cave stated, measurement can also be on interval or ordinal scales. The latter scales will be used frequently whenever “intangibles” (like cultural aspects) are to be quantified.
- ▶ **Content validity:** The indicator should measure the phenomenon it refers to. The content validity is a crucial requirement. The fact that the indicators are used in an international comparison complicates the quest for content validity even further: does this indicator mean the same in all the countries compared?
- ▶ **Face validity:** In addition to having a good content validity, the indicator should also be easy to understand for all users. The use of complicated formulae to calculate an indicator-value is not very helpful in getting a good face validity.
- ▶ **Reliability:** The data underlying the indicator scores should be trustworthy. If a measurement is repeated, the results should be identical.
- ▶ **Timeliness:** Indicators should produce information that is up-to-date. For many statistics, the time between the event and the availability of data on that event is relatively long (three to four years). Such time-lags may distort the judgement of policy-makers when deciding on action.
- ▶ **Feasibility:** Collecting data and calculating indicators that meet all requirements listed has to be feasible. This refers to costs of collecting data and calculating indicators, as well as capacity and time.
- ▶ **Set of indicators:** A crucial element in the discussion on interpretation is that indicators should not be used in isolation. Indicators point to different aspects of a certain phenomenon and the combined information may provide a true image of the phenomenon. In the choice of indicators, one has to make sure that all relevant

aspects are covered. To combine the information of indicators a map of the relations between the indicators is needed. Whether this mapping is a causal model or merely a presentation of the stocks and flows like the input-process-output model is not of prime interest, as long as relations are made explicit. Redundancy within the set of indicators should be avoided. If two indicators are strongly related to each other, the elimination of one of them should be considered.

The setting up of the selection process for the indicators was promising: eight working groups were set up in successive waves from the second half of 2001 in order to implement the common objectives. This first phase was necessary in order to identify the priority themes, make an inventory of existing experience, define a preliminary list of indicators for monitoring progress and to secure the consensus needed between all the interested parties. A Standing Group on Indicators and Benchmarks (SGIB) was established in 2002 to advise the Commission on the use of existing indicators and benchmarks and on the development of new ones. Its main concern has been to evaluate the pertinence of the suggested indicators in relation to the objectives and the availability and relevance of data within each indicator area.

The working groups produced long lists of indicators for the various areas. The lists comprised existing indicators as well as proposals for the development of new indicators. Based on these lists and the indicative list of 33 indicators, the SGIB advised the Commission to use the 29 indicators specified in the appendix. The list of 29 indicators was heavily criticised, especially by a number of members of the working groups. In their advice, the SGIB closely followed the indicative list of the Commission and took the advice of the working groups into account in only a limited number of cases.

The main problems are in the content validity of the indicators. For a number of indicators it is quite clear that the indicators selected do not provide an indication of the objective. Timeliness is also a problem that is often referred to by the advisory groups. For some indicators, most recent data refer back to the previous century, which is quite useless to monitor progress since 2000. Some indicators are very difficult to use or to understand. The validity of an indicator like the number of young people in the population to measure the quality of teachers and teacher education is, on face value, difficult to understand. Reliability is a relatively strong point of most indicators. A substantial effort has been put into this by Eurostat, the statistical office of the EU. The final criterion of the theoretical list calls for the use of a coherent set of indicators. In

the interpretation of the list of indicators and their scores, most analyses are partial, focusing on only one indicator. The lack of a coherent framework for the indicators prevents the user from making the fullest use of the information provided. On one criterion almost all indicators score highly, namely feasibility. This leads to the conclusion that feasibility has been the main criterion for the selection of indicators.

Why is it that feasibility is such a prominent criterion and content validity as a criterion is almost completely neglected? A concrete answer to these questions is difficult to give, but based on discussions with members of the SGIB and an analysis of the documents we can give some tentative answers.

Part of the answer probably lies in the political character of the process. The strategic goals were set at the political level, as, to a certain extent, were the 13 objectives. Congruence and coherence in the strategic goals and the associated objectives are not very likely to be important criteria in such political processes. In addition, heavily politicised multi-actor processes have the tendency to produce relatively broad (or should we say, vague) formulation of the objectives. This makes it difficult to define adequate indicators for which data are available.

It may also be partly related to the state of the art in educational indicator development. Although a lot of work has been done over the last 15 years, there are still only a limited number of indicators for which international comparative information is available.

Time pressure is a third possible explanation. For other policy areas, the array of indicators for which information is available is much larger and has already existed for quite some time. Previous progress reports have been written for those policy areas, while there were no full reports on education and training. The pressure to come up with such a report in the framework of the Lisbon ambition may have had an impact on the selection made.

RESULTS

For all 29 indicators on this list, statistical information was collected by Eurostat for the EU member states, the ten acceding countries and a number of other European countries. This process was finalised at the end of 2003.

In February 2004 the progress report, comprising the data on the 29 indicators, was adopted. However, results were published earlier, particularly on the five benchmarks the European Council had adopted. The Commission report on those benchmarks⁴

⁴ Communication from the Commission (2003), "Education & training 2010: The success of the Lisbon strategy hinges on urgent reforms", Brussels.



sent a critical message to the member states, urging them to increase their investment in human resources. A number of red lights are flashing which require immediate action.

THE WAY FORWARD

The Commission has acknowledged most of the problems described. It has stressed that the Detailed Work Programme was adopted only two years ago, which is a short period given the complexity of the task. The list of indicators will be further developed to create a more solid basis for the monitoring of progress and the Commission has identified a number of ways to move forward.

DEVELOPMENT OF NEW INDICATORS

The first way is to select new and additional indicators that may help to complete the insights into progress towards achieving the 13 objectives. Six indicators are now further developed.

- ▶ In a knowledge-based society, learning to learn skills has become crucial for success. Information on those skills is however sparsely available, which has led the Commission to set up a process to develop and implement a method to measure such skills.
- ▶ In a Union where over a dozen languages are spoken, foreign language skills are crucial for strengthening the Union. New methods are developed to measure those skills.
- ▶ Teachers are an important input into the education process for which no satisfactory information is available regarding their quality, especially the continuous professional development of teachers. A task force has been set up to develop such an indicator.
- ▶ There has been a strong call for the development of an efficiency indicator. The development of such an indicator has been discussed for quite some time and is still on the agenda. However, lack of consensus on the definition of the indicator (what outputs should be linked to what inputs?) frustrates this process.
- ▶ In the EU, mobility of students has been one of the main policy issues in European higher education policy for a long time, and high student mobility is still seen as an indispensable instrument to forge the EU into a real Union. The existing indicators are based on the exchange programmes only, which convey a limited picture of

student mobility. A task force has been set up to come up with a new indicator.

- ▶ A final indicator refers to the equity objective: steps are first taken to develop a framework for monitoring the socio-economic backgrounds of students in tertiary education.

There are some areas (equity and ICT) in which discussions are going on but no substantial plans for the development of concrete indicators exist yet.

It is clear that the development of these new indicators and methods to obtain data on them will improve the picture we have of the progress towards the 13 objectives. However, these benefits do not come without major costs. The development of the skills indicators will take a long time and considerable amounts of money, and the problems regarding the efficiency indicators will prove difficult to resolve. The development of the other indicators may happen faster, but the time needed to have time-series available on those indicators is still considerable.

THE USE OF COMPOSITE INDICATORS

Due to the large number of indicators necessary to cover the full range of policy fields involved in the follow-up to the Lisbon conclusions, efforts have been made by the Commission to develop specific composite indicators on “investment in the knowledge-based economy” and “performance in the transition to the knowledge-based economy”. In due time, such indicators can be applied to give an aggregated view of progress towards the Lisbon targets for the European knowledge economy. Composite indicators with a special focus on education and training are also considered.

This development may be a viable way to reduce the information overload that long lists of indicators produce. However, the methods and calculations used have to be transparent, and there has to be consensus regarding the relations between the components of the composite indicator and the relation between the composite indicator and the objective or goal. These conditions have not yet been met, which has led to some debate on the experimental composite indicators presented.

TOWARDS A COHERENT INDICATOR FRAMEWORK

A relatively new way forward is the development of a coherent framework for the indicators. Such a framework is seen as an important tool to improve the selection of new indicators and to evaluate the list of existing indicators. It may also be used to



check the validity of the interpretation of single indicator/objective scores and to improve the validity of composite indicators and develop a more transparent design of those indicators.

CHECK FOR COMPLETENESS

The goal set in Barcelona 2002 was “to make Europe’s education and training systems a world quality reference by 2010”. This put “quality” as a central concept on the agenda. Unfortunately, this is not very helpful for developing a set of indicators since there is no unambiguous definition of quality. In a recent study by Scheerens, et al, a number of perspectives on quality were described, based on a simple input-process-output model. From the productivity perspective, the focus is on the outputs of the education process. The instrumental effectiveness perspective focuses on the process: how the education process is organised in order to achieve the intended outputs with the inputs available. From the efficiency perspective, outputs are related to inputs; and from the fourth perspective, the focus is on equity. From this perspective the distribution of outputs, processes and inputs over various groups of participants is analysed.

Based on the objectives in the DWP, two types of outputs can be discerned: outputs in terms of participation and completion, and outputs in terms of skills. A large proportion of the indicators in the list of 29 refer to participation-outputs (14), whereas skills as an output are represented by only five indicators. Process indicators are completely missing in the list of indicators. Although there are some objectives that refer to the way the educational process is organised, no process indicator was selected.

From the efficiency view of quality, inputs are related to outputs. There are eight input indicators, but it proves to be extremely difficult to achieve any kind of consensus on what inputs should be related to what outputs. This stalemate position has led to the absence of efficiency indicators on the list.

Safeguarding or enhancing social cohesion is a key element in the Lisbon strategy. An important aspect of this is the way goods, services and opportunities are distributed across groups in society. Despite this central position of the equity perspective, there are no clear indicators that address distributional aspects.

This evaluation of the completeness of the list of indicators has highlighted white spots in a number of areas. The Commission has responded to this by starting the

development of new indicators. Among those new indicators there is no process indicator. This apparent lack of interest in how the education process is organised is partly due to the history of the relation between the member states and the European Union. In the Maastricht Treaty, education was deemed to be a national affair in which the EU had a very minor say. How the education process is organised is a matter that is seen as essential to national identity and too many process indicators would invade that national territory.

Equity has been put on the agenda by the Commission, but the development of equity indicators does not keep up with the growing need for that information.

MAPPING RELATIONS TO IMPROVE INTERPRETATION

A second type of evaluation of the list of indicators hinges on the relatedness of indicators and objectives and is important for the interpretation of the scores on the indicators. The setting up of strategic goals, objectives and indicators implies a hierarchical model as specified in Figure 2.

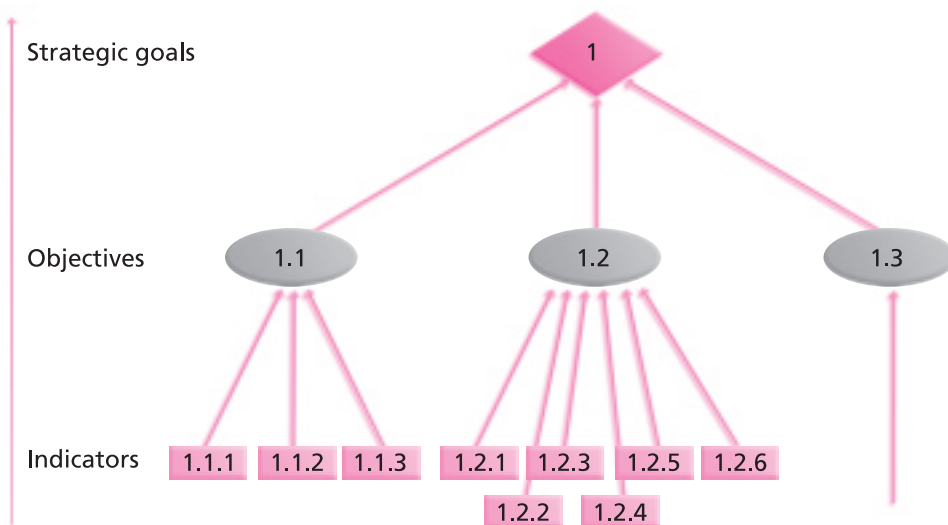


FIGURE 2: The hierarchical model

The basic assumptions in this “model” are that the indicators are independent of each other and that the objectives are independent of each other. However, this model does not apply to the reality of the list of 13 objectives and 29 indicators. What we found is that:

- ▶ The scope for a number of the objectives is much wider than the scope of the indicators; based on the indicators selected we can get only a very partial idea, if any at all, of progress towards the objective.
- ▶ Indicators within one objective are related to each other.
- ▶ Indicators across objectives are related to each other.
- ▶ Objectives are related to each other.

In order to give a valid interpretation of the score on a particular objective, it does not suffice to look at the score of a single indicator or even the combination of the indicators under one objective. It requires more in-depth analysis of which indicators may be related to a particular objective, why we think the indicators are related and how they are related.

As an example, we describe the mapping of the indicators related to objective 1.4: Increasing the recruitment to scientific and technical studies. In the DWP, the rationale for this objective is the following:

Scientific and technological development is fundamental for a competitive knowledge society. All citizens need a basic understanding of mathematics, science and technology. If Europe is to maintain, let alone to improve, its position in the world, and to meet the Lisbon targets, it must do more to encourage children and young people to take a greater interest in science and mathematics; and to ensure that those already in scientific and research careers find their careers, prospects and rewards sufficiently satisfactory to keep them there. In this context, gender balance must be encouraged.

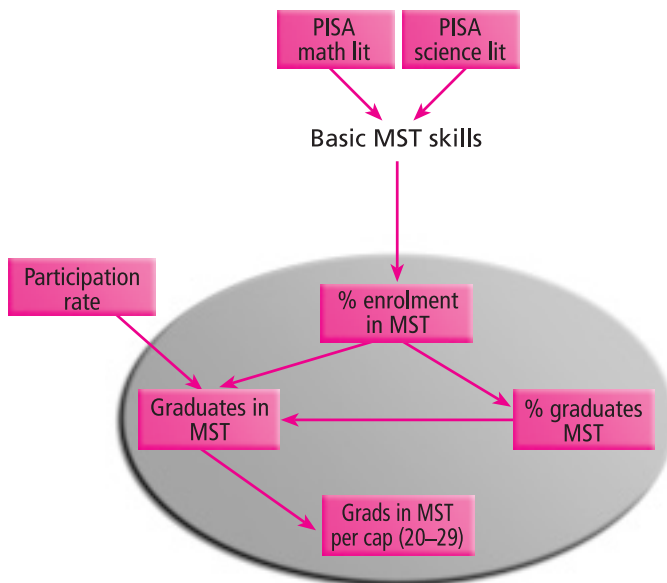
The indicators identified for this objective are:

- ▶ Students enrolled in mathematics, science and technology (MST) as a proportion of all students in tertiary education (ISCED 5A, 5B and 6).
- ▶ Graduates in mathematics, science and technology (ISCED 5A, 5B and 6) as a percentage of all graduates (ISCED 5A, 5B and 6).
- ▶ Total number of tertiary (ISCED 5A, 5B and 6) graduates from mathematics, science and technology fields.
- ▶ Share of tertiary graduates in mathematics, science and technology per 1 000 inhabitants aged 20–29 – broken down by ISCED levels 5A, 5B and 6.

Given the rationale, the main indicator among the four listed is the share of tertiary graduates in MST per 1 000 inhabitants. The number of tertiary graduates is merely an auxiliary variable to that indicator (and may be skipped from the list). The first two indicators indicate the interest in MST among higher education students. The relation between the two refers to the relative drop out rate in MST programmes. The overall basic understanding of MST, as mentioned in the DWP, is not covered by these indicators.

We assume that some of the other indicators on the list also have an impact on the scores of objective 1.4. We assume that the number of graduates is determined not only by the percentage of MST enrolment, but also by the rate of participation in tertiary education. In addition we assume that the percentage of MST enrolment may be positively related to the level of basic MST skills, which may be positively related to the indicators on math literacy and science literacy. When using the scores of the indicators of objective 1.4, we have to take into account those “external” circumstances as well.

FIGURE 3: Increasing recruitment to scientific and technical studies



Similar mappings to the one shown above can be made for the other objectives, which can then be compiled into one table in which all assumed relations are specified. Such a tool may then be used to improve the interpretation of the indicators. It may also be used to improve the definition and calculation of composite, more contextualised



indicators that take into account more characteristics of the system than the simple or isolated indicator. However, we need to resolve a number of issues before we can do that. Detailed specification and a clear view of the relevant time frames are only two of those issues, yet not the easiest to resolve.

DISCUSSION

The role of indicators in the Open Method of Coordination (OMC) in the education and training area is still limited. This is partly due to problems in availability of trend data which makes it very difficult to monitor progress. It may also be due to the fact that their *raison d'être* is not yet implemented. As mentioned above, the indicators are intended also as an instrument for stimulating the exchange of good experience and new ways of thinking about policy approaches. The analysis of good or best practices is not yet implemented. The focus is still on the indicators and not on what member states can learn from them. If this aspect receives more attention, as is expected, the need for good indicators (to find the best practices) will grow.

The complexity and dynamics of the policy context also seem to have an impact on the role of indicators. Complex and dynamic contexts may lead to large sets of indicators, in an effort to capture as much of the relevant phenomena as possible. If we are trying to create an indicator system that will be used by policy makers, the creation of large indicator sets is counter-productive. Policy-makers will pick a small number of indicators, for which data are readily available or, to be a little cynical, suit their purpose best. The selection of the five European benchmarks is a good example of that.

In a turbulent environment, the message the indicators convey must be loud and clear. Therefore, a small set of indicators, for which clearly specified targets are set, seems to be the best way to increase the role of indicators in policy-making in current European policy setting. The targets, however, have to be realistic and the set of indicators must have a broad scope. The latter is clearly not yet the case. The development of composite indicators may be a major contribution in this respect.

The list of indicators used needs to be constantly reviewed. Not only do we need to check whether the list still covers all relevant aspects of the goals towards which we want to measure progress, we also need to explore the relations between the indicators that may enhance our understanding of the phenomena we are trying to measure. In time, this may lead to a concise set of composite indicators that meet most of the technical criteria mentioned.

PERFORMANCE INDICATOR DEVELOPMENTS IN SIX AFRICAN COUNTRIES

MOZAMBIQUE – ARLINDO CHILUNDO

Mozambique has developed a long-term strategic plan divided into three time-phases: 2000–2004, 2005–2008 and 2009–2010. Emphasis in phase one has been on the building up of a database for higher education institutions. Its main objective has been to achieve uniformity in the information coming from higher education institutions. Institutional strategic plans and institutional goals and benchmarks were consulted in the development of this national database, although some institutions refused to release their strategic plans and information.

Currently, Mozambique has no way of measuring performance and progress that has been made in the higher education system. However, consideration is being given to the establishment of a committee to examine the feasibility of producing indicators for measuring quality in higher education institutions.

NIGERIA – SAMBO ABDULRAMAN

The use and development of higher education indicators and benchmarks in this country has occurred primarily in the context of quality assurance and the accreditation of higher education programmes. In the period after 1983, minimum academic standards were established for all programmes/disciplines, and these were then developed into benchmarks used to judge and accredit programmes in all higher education institutions.

Nigerian universities, irrespective of ownership, have to be accredited and teams are drawn from universities to assess programmes at universities for accreditation. Two systems of external accreditation are used – one for academic programmes and one for professional programmes. Accreditation lasts six years, and interim accreditation is valid for two years. Also, institutions are given the opportunity to assess their own performance.

Six broad indicators have been used to measure the performance of higher education institutions in Nigeria. These indicators relate to academic standards, staff, students, employers' ratings, library services and funding. Benchmarks such as those for staff-student ratios, apply to the same programmes in all institutions. Other benchmarks relate to ranks of staff, expenditure within units, female enrolments,



geographical spread of admissions, science enrolments and annual growth rates. A set of management performance indicators are also used in annual audits of higher education institutions.

EGYPT – KAMAL KAMEL

In 2000 Egypt started a quality assurance project, which is run by a national committee on quality assurance. This committee undertook, as a first step, a study of four universities. The committee is now considering the assessment of higher education institutions at a programme level, and has embarked on a system of capacity building which will run until 2007.

Egypt has established a National Quality Assurance Agency (NQAA) which is part of the office of the President, and not the Higher Education Ministry.

BOTSWANA – RICHARD NIELL

The single university in Botswana operates in a very resource-rich environment. A Tertiary Education Council oversees higher education and it has been involved in the benchmark project of the Association of Commonwealth Universities (ACU). Benchmarks are to be used to develop performance indicators for comparisons across commonwealth countries. The performance indicators relate to current performance within certain focus areas and the benchmarks represent projections into the future. A system has been developed that includes a profile of twelve “vital signs” of an institution’s performance. The “vital signs” are derived from the performance indicators and can be used on a regular basis to test the current performance of the institution.

KENYA – DAVID COURT

Not much data has been produced on the performance of higher education institutions in Kenya and quality issues are only beginning to be considered. While there is a recognition of the need to set some targets, at present there is no legal framework governing higher education programmes and no regulative framework for trans-national provision. Kenya is supportive of the revitalisation of the Inter-University Council for East Africa, and will put quality high on its agenda. It is also considering establishing a regional body for higher education quality assurance.

TANZANIA – DANIEL MKUDE

The University of Dar es Salaam (UDSM) has embarked on a programme of developing quality performance indicators for the prime functions of teaching, research and community service. In a highly participatory process of establishing widely accepted quality indicators for research, the starting point was to identify six broad indicators based on a survey of relevant literature. These indicators relate to research planning, research training, staff participation in research, research students, research outcomes and research impact. University staff suggested some additional indicators that would measure customer satisfaction, which included, for example, society's perception of research at the UDSM and the number of local/international research assignments that UDSM wins in a competitive environment. Staff were asked to rate the suitability of these indicators on a five-point scale and the results from this survey were used to identify and rank the top ten indicators. The process of developing reliable performance indicators at UDSM is still in its infancy.

SUMMARY OF RESPONSES TO INTERNATIONAL PRESENTATIONS

In discussing presentations made from an international perspective, participants at the March 2004 seminar made some of the following observations.

It appeared that in many countries academics had a "gut" distaste for and scepticism about performance indicators, because they consistently rated themselves as "good" and therefore did not believe that sets of performance indicators were needed to regulate their behaviour. This kind of reaction had also been noted in South African discussions of performance indicators.

It appeared also that loss of trust has in recent years characterised the relationship between higher education and governments. This led to academics assuming that performance indicators have been introduced by governments to address "bad attitudes" and "incompetent" institutions.

Performance indicators are often viewed negatively because they are assumed to be part of regulatory mechanisms which are placing limits on the autonomy of higher education institutions. This kind of argument has been used in South African discussions of government-higher education relations.

The use of performance indicators could be defended on the basis that, in countries



such as South Africa, they are needed to demonstrate that the higher education system is being “remade”. However, while performance indicators can be part of mechanisms for steering a higher education system, they should not be used by government for the micro-management of institutions

A further argument in favour of the use of performance indicators is that they could assist with evaluations designed to increase public confidence in a higher education system. Their use could draw investment into the higher education system and assist in its growth.

South Africa should take note of the European Union discussions, and should absorb the lessons which can be drawn from this experience in developing higher education performance measures.

A specific feature of the European Union project which is relevant to South African debates is the drawing up of criteria which performance indicators have to satisfy. The seven criteria used in this project should be applied to any South African development of performance indicators:

- ◆ The phenomena to be indicated should be quantifiable.
- ◆ The indicator should have content validity.
- ◆ The indicator should have face validity.
- ◆ The data underlying the indicator scores should be reliable and trustworthy.
- ◆ The indicators should produce information that is up to date.
- ◆ Collecting data and calculating indicators that meet all requirements must be feasible.
- ◆ Indicators should not be used in isolation. To combine the information from indicators, a map of relations between the indicators is necessary.

The EU project also shows that the major challenges facing the development of any system of indicators are: to provide detailed specifications of indicators that have a broad scope; to provide clearly specified and realistic targets; and to provide relevant time frames in which to achieve and measure targets.

A final lesson to be drawn from the EU experience is that the development of a national system of performance measures and performance indicators requires time, and will involve considerable research, discussion and consultation.