INSTITUTIONAL EVALUATION: CAN IT CONTRIBUTE TO IMPROVING UNIVERSITY PERFORMANCE?

Alan Lindsay  
School of Education  
Macquarie University

Other Western countries, like their counterparts in other system contexts, are increasingly operating under financial constraints which have accentuated the need for improved planning and administration. These financial pressures have stimulated interest in techniques for assessing system, institutional, and sub-unit performance. Following the report of the Williams Committee with its recommendations for extending research into institutional and system performance,1 and the establishment of the Evaluative Studies Program by the Tertiary Education Commission,2 most universities are engaging in evaluative studies intended to yield information which will assist endeavours to improve institutional performance. Some of the Australian and overseas experience with institutional evaluation and review has been reported by Clare and Birt,3 Holdaway,4 Galvet,5 and Harman and Johnston.6

However, there are substantial obstacles confronting the Williams Report. First, the basic institutions and processes are not well understood and so the meanings of the terms “performance” and “management” in the university setting are not completely clear. Hence, it is not possible to define exactly what is to be evaluated. Second, adequate measurement techniques and judgmental procedures are not available and so there is a substantial gap between what is desirable in an evaluation and what is achievable. These current evaluative studies often provide results which are fragmented assessments relating to a set of arbitrary dimensions of performance.7 The dimensions are generally not in the form that can be readily utilized by decision-makers in attempts to improve university performance.8

The purposes and uses of evaluation  

Chapter 10 of the Williams Report deals specifically with one of the items in the terms of reference which directed the Committee to advise on the “means of evaluating the quality and efficiency of the system”. The Committee associated the increasing concern with evaluating quality and efficiency with the greatly increased financial burden of supporting education, the allegations of mismatches between supply and demand for graduates, and the failure of education to achieve desired social objectives. The Committee supported its recommendations for the extension of evaluative studies by a conceptual analysis and suggestions for evaluation checklists relating to various components of the system. The Committee reported that fewer submissions were received in relation to this item of the terms of reference than in relation to any other, and acknowledged that interest in evaluating the quality and efficiency of the education system has grown faster than the capacity to do the evaluation.

The Committee’s review of current arrangements for evaluation reveals the lack of both evaluative focus and rigour in the current scattered and disjointed evaluation practices. Unfortunately, the Committee’s own discussion of the topic is superficial; it draws upon no theoretical or empirical foundation, and there is a lack of rigour in its analysis of the concepts and problems involved.

One deficiency in the Report is the absence of any discussion of what is meant by “evaluation” and the ways in which evaluation is distinguished from other decision-making activities. Evaluation is a complex process as the substantial literature on educational evaluation attests. Feaslby has recently produced a dictionary of evaluative studies intended to yield information for decision-making in relation to attempts to improve the system or programme. Finally, evaluation may be undertaken merely as a ritual to provide a perception of rationality and accountability which promotes a feeling of security.

Quality and efficiency  

The basis of the Williams Committee’s discussion of evaluation is an analysis of the terms “quality” and “efficiency”. The Committee’s contribution was to provide dictionary definitions of these terms. Quality is defined as a “degree of excellence”. Since the complex concept of “excellence” is not analysed, the meaning of quality is not greatly clarified or specified. Efficiency is defined as “the extent to which resources are converted into educational quality which the Committee could usefully pursue. Sources of the most recent work include Lawrence and Green,10 Lawrence and Solomon11, and Kuhl.12 Astin13 reviews the “five very different concepts of quality that have dominated educational research and policy during the past several decades”. One view is the “mythical” concept which maintains that quality simply cannot be defined or measured because the activities of institutions are too complex and varied, because different institutions are different. Alternatively, there is a “reputational” concept which relates outcomes of higher education to the institutions from which the graduates come. A view promoted by educators is to equate quality with an institution’s reputation and prestige. A third view which maintains that quality is a multidimensional concept is the “value-added” approach. This view which defines quality in terms of financial, economic, and social returns is defined in relation to national economic development and the extent to which the actual outcomes match the objectives or intended outcomes. Alternatively, there is a “productivity” concept which accords by an opinion survey. A view promoted by efficiency with the underlying assumption that the system works for itself. The basis of the Williams Committee’s discussion of evaluation is an analysis of the terms “quality” and “efficiency”.

A second view is the “evaluational” concept of quality which relates outcomes of higher education to the institutional system. The value-added view which defines quality in terms of financial, economic, and social returns is defined in relation to national economic development and the extent to which the actual outcomes match the objectives or intended outcomes. Alternatively, there is a “productivity” concept which accords by an opinion survey. A view promoted by efficiency with the underlying assumption that the system works for itself. The basis of the Williams Committee’s discussion of evaluation is an analysis of the terms “quality” and “efficiency”.

A second view is the “evaluational” concept of quality which relates outcomes of higher education to the institutional system. The value-added view which defines quality in terms of financial, economic, and social returns is defined in relation to national economic development and the extent to which the actual outcomes match the objectives or intended outcomes. Alternatively, there is a “productivity” concept which accords by an opinion survey. A view promoted by efficiency with the underlying assumption that the system works for itself. The basis of the Williams Committee’s discussion of evaluation is an analysis of the terms “quality” and “efficiency”.

The value-added approach directly relates quality and efficiency with the intended outcomes and the establishment of the Evaluation Commission. The deficits in the Report are the absence of any discussion of the terms “quality” and “efficiency”. The Committee’s own discussion of the topic is superficial: the Williams Committee’s discussion of evaluation is an analysis of the terms “quality” and “efficiency”. The Committee’s own discussion of the topic is superficial: the Williams Committee’s discussion of evaluation is an analysis of the terms “quality” and “efficiency”.

The dimensions are generally only tenuous and the Williams Committee with its recommendations are not completely clear. Hence, it is not possible to define exactly what is to be evaluated. Second, adequate measurement techniques and judgmental procedures are not available and so there is a substantial gap between what is desirable in an evaluation and what is achievable. These current evaluative studies often provide results which are fragmented assessments relating to a set of arbitrary dimensions of performance. Following are the definitive works in interpreting these concepts is the extent to which the economic analysis of education is discussed by Cohen and Layton.20 A review of research on the efficiency dimension of institutional performance is provided by the classical models. The main organizational characteristics are different for different groups; are often in conflict or the subject of factorial conflicts, and are not readily translated into clear-cut programmes. The other characteristics include unclear and imperfectly understood educational outcomes which have been defined in decision-making with members taking part only intermittently according to their interest in particular issues; structural looseness with relatively ill-
defined links and relatively high levels of individual and sub-unity autonomy; and widely differing criteria of success. Consequently, even if participants make judgements in the absence of measurement or market prices for outputs.

**Evaluating effectiveness and efficiency**

If the competitive perspective is adopted, the evaluation of effectiveness involves comparing outcomes with goals or making corrections to the pursuit of goals. Both approaches are suggested, for example, by those who make judgements in the absence of measurement or market prices for outputs.

Despite the considerable efforts devoted to "goal-free" competition among activities for resources, producing outcomes from the institutional process, the decision-making for an ideal mix of these outcomes and the decision on an ideal outcome mix or of weighting the value of the actual outcomes in relation to the intended outcomes. This notion of efficiency involves a relationship between two variables regardless of cost. This notion confuses efficiency with the notions of quality and effectiveness. A third approach is to regard the crude proxy input-output measures which are available, as costs per student or student-staff ratios, as adequate measures of efficiency or productivity and to use them in decision-making without regard to their considerable imperfections. A fourth approach is to demand a strict cost-benefit accounting of expenditures in relation to the anticipated outcomes. Hence there is no way of deciding on an ideal outcome mix or of weighting the various outcomes to produce a single composite measure. For example, the relative values of a PhD and a Bachelor's degree, or of a specific professional degree or a general degree, in deciding the trade-offs for an ideal mix of these outcomes. Lindsay and Bailey have stated this problem as follows:

> The outputs of higher education are not solely goods whose value is determined in the market place, so the relative value is only partially determined by a price mechanism. Even if a market for the outputs existed it would only provide one way of measuring value. The values of the outputs of higher education is a complex and normative issue, the existence of varying perspectives on the purposes of higher education makes it difficult views about the value of the outputs and so many different assessments of performance may be made.

Secondly, many of the outcomes of education are intangible and therefore not easily identified or measured. Despite the considerable efforts devoted to outcomes research in the United States, there are still substantial deficiencies in the techniques for measuring outcomes. The development of outcomes research is only one of the many influences in the development of individuals and the progress of society, it is difficult to isolate education's distinctive effects. These considerations present a major challenge to those engaged in evaluating effectiveness in education. The influence of education's complexity and intangibility can also be seen in relation to the application of the notions of efficiency to educational institutions. The most appropriate concept of efficiency is one which requires the comparison of outcomes with the resources required to attain these outcomes. In higher education, several simplistic notions are frequently used. One involves judging efficiency only in relation to cost, by assuming that a lower cost is necessarily preferable to a higher cost. For example, an institution with a cost per student of $5,000 may be judged as less efficient than one with $4,000, regardless of the quality of the service provided to the students. Another notion is to judge efficiency only in relation to outcomes. The concern is, to assume that improved outcomes are desirable regardless of cost. This notion confuses efficiency with the notions of quality and effectiveness. A third approach is to regard the crude proxy input-output measures which are available, as costs per student or student-staff ratios, as adequate measures of efficiency or productivity and to use them in decision-making without regard to their considerable imperfections. A fourth approach is to demand a strict cost-benefit accounting of expenditures in relation to the anticipated outcomes. Hence there is no way of deciding on an ideal outcome mix or of weighting the various outcomes to produce a single composite measure. For example, the relative values of a PhD and a Bachelor's degree, or of a specific professional degree or a general degree, in deciding the trade-offs for an ideal mix of these outcomes. Lindsay and Bailey have stated this problem as follows:

> The outputs of higher education are not solely goods whose value is determined in the market place, so the relative value is only partially determined by a price mechanism. Even if a market for the outputs existed it would only provide one way of measuring value. The values of the outputs of higher education is a complex and normative issue, the existence of varying perspectives on the purposes of higher education makes it difficult views about the value of the outputs and so many different assessments of performance may be made.

The outputs of higher education are not solely goods whose value is determined in the market place, so the relative value is only partially determined by a price mechanism. Even if a market for the outputs existed it would only provide one way of measuring value. The values of the outputs of higher education is a complex and normative issue, the existence of varying perspectives on the purposes of higher education makes it difficult views about the value of the outputs and so many different assessments of performance may be made.
dardisation into this process, the American literature on accreditation and self-study attests to the very limited gains in reliability from this strategy. A more sophisticated approach has been the development of standards-based instruments for measuring institutional dimensions on some institutional dimensions. These instruments include the measures of student learning outcomes developed by the Educational Testing Service (ETS), the American College Testing Program (ACT), and the National Center for Higher Education Management Systems (NCHEMS); the College and University Environment Scales (CUES), the Institutional Functioning Inventory (IFI), the International Group of Institutions (IGI); and the many techniques for analysing costs, staff activities, and space utilisation developed by NCHEMS and other agencies. A number of useful reviews of these instruments are available. For the dimensions concerned, these instruments provide the most suitable means of assessment, although they also are by no means perfect.

There are a number of difficulties with the self-study approach. Questions may be raised in relation to how well the dimensions used in the checklists represent the performance of an institution or department. The inclusion of most dimensions appears to be based on commonsense or administrative folk-lure rather than on the basis of a theoretical or empirical foundation. Even though institutional processes are not well understood, it is still desirable that each dimension be justified in relation to the evaluation of institutional performance. One consequence of the absence of a theoretical foundation underpinning an institutional self-study approach is that the approach does not provide any mechanism for weighting the assessments of dimensions or aggregating them into an overall assessment of performance. Thus a good deal of responsibility and power is given to the individual who is trying to synthesise the results and judge the overall performance.

Making separate assessments in relation to each dimension means that the institutional self-study approach does not necessarily provide information about effectiveness and efficiency. To do so, the criteria for assessing each of the dimensions which conform to the overall category must be found on the checklists, and compared with intended outcomes, and on the other, with costs or resources used. Assessment of the former type is much more commonly included in the dimensions and criteria proposed; the institutional self-study approach is thus largely concerned with assessments of effectiveness rather than efficiency.

Some of the problems of the institutional self-study approach may be anticipated and avoided. For example, the evaluation of an institution with an embryonic stage; it is under conceptual development, an approach which has quite severe limitations. It is now replacing the first in popularity, is to make inter-institutional comparisons and judgements in relation to the trends in the institutional performance. Trends in such quantitative indicators as costs per student and student-staff ratios could be usefully studied in this way, as could the results from different methods of self-assessment. In addition, if carried out carefully, institutional self-assessments might also be useful in identifying trends in performance improvement or decline.

Inter-institutional comparisons

An alternative approach to institutional assessment, which is most powerful when qualitative measures or indicators are available, is to make inter-institutional comparisons. To the extent that a set of dimensions of institutional performance is homogeneous, comparisons can yield useful assessments of both relative effectiveness and relative efficiency. Peterson sees the development of the inter-institutional research field as being dependent on progress in three areas:

- major standardised instruments that have institutional norms and available databases;
- the growth of large-scale institutional data bases that can be used to make institutional databases available, and the availability of the computer networks with more sophisticated software for improved access and ease of data sharing;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
- a detailed data element dictionary; a methodology for collecting the data;
In Australia, however, the sequence of events now to be described in the 1960s, although parallel developments, was sensitive to a possible accusation from the States that it was becoming involved in "wage-fixing" in universities. Throughout the 1960s, with considerable reluctance, the predecessor of the Universities Council of the Commonwealth Tertiary Education Commission (the Australian Universities Commission (AUC)), found itself involved in attempts to resolve the academic salaries issue. Despite its moves to adopt a disinterested stance, the AUC had to take part in what was the component of the recurrent grants devoted to academic and non-academic salaries by each university was so large. The AUT put up to the government successive "claims" based on the comparability principle used by the Civil Service Commission and specifically on an attempted direct comparison with civil service salaries, arguing that the chief alternative employ­ment for academics was the administrative or scientific class of the civil service. But this comparison was never specifically accepted by the UGC or Treasury. A "slabbery minority" developed whereby the UGC listened to the AUT, listened to the Committee of Vice-Chancellors and Principals and then made its own confidential recommendations to the government.

In 1963 the UGC's role was significantly diminished when academic salaries were made the subject of an independent review by the National incomes Commission. There was then a brief reversion in 1966 to the arrangement whereby the government named a (sec 57) "independent UGC which in turn consulted the AUT and the Committee of Vice-Chancellors and Principals. As the revision was not satisfactory, in 1968 and again in 1970 an independent review was conducted by the National Board for Prices and Incomes. During that time, the role of the universities had been established by the Australian Universities Commission (AUC) which had the task of reviewing academic salaries. For the UGC and the Treasury to be bargaining on a trade union model. During that time, the UGC maintained a unique and important role in the process of reviewing academic salaries. There was no workable pattern of negotiation between employers and employees and the UGC did not act on behalf of universities but in its own right.


Z. Cowen, "Problems in assessing efficiency," Paper presented at the Symposium "How efficient are Aus­tralian universities?" held at the University of NSW on November 8th. 1982.


M. W. Peterson, op. cit.

M. W. Peterson, op. cit.


