

Quality of Student Outcomes: Concepts and Issues of Measurement

Ng Gan Che, Ken Heskin and Raj Sharma

Introduction

The current government's concern in the quality of tertiary education was triggered by public disquiet towards the deteriorating state of Australian tertiary education (Lindsay, 1992). This public perception of the tertiary education environment can be attributed to several factors. These include the trend towards mass tertiary education, a dramatic increase in demand for tertiary places as a consequence of high retention rate of year 12 students, economic rationalism, and the demand for public accountability in tertiary education as the national coffer goes deeper into the red. As a consequence, there is public demand for efficiency and effectiveness in the tertiary education sector. Student outcomes assessment is one of the processes that has been recently identified as being more relevant in indicating an institution's performance and accountability.

Assessing the quality of tertiary education is by no means a simple procedure. There are no absolute or simple values either within the tertiary education system or, indeed, over time in relation to the concerns of any of the key player groups or stakeholders. The ground keeps shifting. The Higher Education Council's (HEC) document (1992) *The Quality of Higher Education — Discussion Papers* has recognised both the conceptual and practical problems and issues of assessing the quality of tertiary education. Of late, these problems and issues, which generally have been centred on the definition and measures of quality, have been widely debated in Australian academic circles in particular (refer to the HEC Discussion papers for a comprehensive account of the debate by the different stakeholders: students, academic staff employer unions, etc.).

Following Kerschner (1987), who suggested that the increased diversity of students commencing programs mean that input measures alone will be insufficient to obtain a complete picture of quality, Hall (1992) argued for a focus on outcomes in quality assessment. Hall averred that the main concerns of the clients -the students and the employers - are on quality outcomes. Hence, institutional aim at satisfying the needs of our clients (Layer, 1992). This view is apropos to the HEC's current emphasis on the quality of student outcomes as a fundamental pathway 'to understanding how each of the processes within institutions are organised and evaluated ...' (HEC, 1992, p.6). Underlying this focus, the HEC has adopted three broad principles as the basis for assessing the quality of student outcomes. These are listed as follows:

- the attribute acquired by graduates provide the ultimate test of the quality of the system to which they have been exposed;
- the judgments about the value of the individual processes that combine to lend to quality outcomes rests with the universities - the internal stakeholders, and to an extent, with the peers; and
- the major criterion to be applied to the judgement of the quality of the individual elements of learning programs should be linked to the contributions that it makes to the staged development of the students (HEC, 1992, p. 7).

Aim And Objectives

The overall aim of this paper is to discuss the concepts and issues of assessing quality of student outcomes partially in the light of experience in conducting a pilot study and informed by some initial feedback from

that process. The pilot study was undertaken at Swinburne University of Technology (SUT). Underlying this central aim, the paper will address the following themes on quality in order to provide a comprehensive framework for a better understanding of the concepts and issues involved:

- an overview of quality assessment in the tertiary education sector;
- the theoretical concepts of quality assessment, especially relating to student outcomes;
- the issues involved in assessing quality of student outcomes;
- methodology in outcomes assessment as applied to the SUT pilot study.

Quality Assessment: An Overview

The concern for quality in Australian tertiary education dated back to the late 1970s (The Australian College of Education, 1978). Other countries are also concerned about quality within their tertiary educational systems. In fact, the quality of education has become an international issue and has drawn much attention from governments and tertiary education institutions worldwide in the 1980s following the appearance of a spate of alarming reports about the declining quality of general education in socio-political and pedagogical publications (Malkova, 1989, p.34). Debate and research activities on quality definition, measurement, and issues abound in the United States (for example, refer to Bogue and Saunders, 1992; and Pascarella and Terenzii, 1991), the Organisation for Economic Cooperation and Development (OECD) countries (Kalkwijk, 1991; Vroeijerstijn, 1991) and other parts of the world. The question of what is quality is best put by Pirsig (1974) as follows:

“Obviously some things are better than others ... but what’s the ‘betterness’? ... so round and round you go, spinning mental wheels and nowhere finding any place to get traction. What the hell is quality? What is it?” (Pirsig, 1974, p.1 84).

For the past twenty years, this question of what is quality has been asked in both the business and education circles. Some say that quality, like beauty, is in the eyes of the beholder. Some say one knows quality when one sees it. Different definitions have been proposed (see for example, Astin, 1991; Mayhew, Ford, and Hubbard, 1990; Goodlad, Soder, and Sirotnik, 1990), and Fisher, 1990). Thus the concept of quality is connotative and varies in meaning according to the context of use. This appears to be true within stakeholder groups as well as between them.

While recognising the multidimensional character of quality, Bogue and Saunders (1992) however warned that one should not detract from the fact that quality can be defined, measured, and “can be used to improve our impact on students and to their growth as well as to enhance programs and services”. To them, quality is about the “conformance to mission specification and goal achievement — within publicly accepted standards of accountability and integrity” (Bogue and Saunders, 1992, p.20). They advised that the debate on quality should not contribute to the proclivity of inaction by institutions or “acting on the possible while awaiting perfection” (Bogue and Saunders, 1992, p.19). Instead, as a learning community, it should be part of a discovery process for all educational institutions.

Within the Australian context, there has been great interest in quality assessment issues. As mentioned in the introduction of this paper, several government sponsored documents on quality and performance of higher education were released over the last three years for public debate. Seminars and workshops were also held to deliberate upon the multidimensionality of the quality issue. Different models of quality assessment, especially those used in the OECD countries, were examined in the HEC (1992) document. The HEC was of the opinion that quality assessment in Australia should be based on: (a) “the need to encourage the search for excellence” (HEC, 1992, p.13), and (b) objective evaluation by an independent audit body to comment and report on the effectiveness of quality assessment practices adopted by different institutions. It is also acknowledged that there is a positive relationship between the level of finding of higher education and its quality. As such there is no cheap way to safeguard quality (HEC, 1992, p.15).

In a comprehensive review of the literature and empirical research studies on the assessment of quality in higher education, Tan (1986) identified three main approaches that were commonly adopted for quality assessment studies. These comprise reputational studies, objective indicator studies, and quantitative correlates studies. Reputational studies are subjective evaluations from academic staff, heads of departments, or deans as a basis for rating programs. Objective indicator studies assess programs through the use of objective variables such as academic research productivity, financial resources or student outcomes. Quantitative correlates studies have a primary purpose of identifying variables that are correlated with academic staff or program quality and they include department size, the amount of federal funding, library resources, academic salaries and academic research productivity. It is suggested that all three approaches have contributed towards a better understanding of quality in higher education.

Outcomes Assessment: Some Theoretical Concepts

The mounting criticism of the quality of tertiary education has spurred research on student outcomes over the last decade. The emphasis on student outcomes assessment is due to stakeholders' or consumers' interest in knowing what the return is on their investment since the costs of tertiary education are now identifiable and measurable (Terenzini, 1989, p.645). In fact as far back as 20 years ago, Mortimer (1972) had predicted that outcomes assessment would be one of the main concerns in American higher education.

Given the diversity of institutions in the tertiary education system in terms of their histories, missions, goals, values, and environments, it would be surprising to find a single concept for student outcomes that fits into all these institutions. Pascarella and Terenzini (1991, p. 5) recognised the difficulty of developing a conceptual framework for presenting the different types of institutional outcomes. A recent study in the United Kingdom by Johnston (1991) partly illustrates this point. Johnston obtained the retrospective evaluations of nearly 9000 1980 British graduates of their degree courses. The findings indicated two possible routes of analysis, one based on graduate's satisfaction in becoming an educated person, the other related to their ability to obtain a good job. Graduates differed in their perspective of the relative importance of these outcomes. So even within the ranks of one major group of players, namely students, at one particular point in time, the picture is not uni-dimensional.

After reviewing the various literature dealing with the taxonomies of student outcomes, Pascarella and Terenzini (1991) considered Astin's model as being the most influential in defining the scope and content of student outcomes, although others like Brown and De Coster (1982), Ewell (1984, 1985a, 1985b, and 1988), Hanson (1982), and Parker and Schmidt (1982) had developed a similar genre of taxonomies.

According to Astin (1973a), student outcomes can be conceptualised along three dimensions, namely type of outcomes which can be cognitive or affective, type of data which can be psychological data or behavioural measures, and the time span. Cognitive outcomes are concerned with the use of higher-order intellectual processes as exemplified by knowledge acquisition, decision making, synthesis, and reasoning. Affective outcomes deal with attitudes, values, self-concepts, aspirations, and personality dispositions. Psychological data describe the traits or internal characteristics of the individual. These traits are usually assessed indirectly by means of a test or examination. For example, an individual's level of skill in critical thinking can be inferred from responses to a questionnaire. Behavioural measures are obtained from direct observations of the individuals. Although Astin's taxonomy illustrates the complexity of measuring student outcomes, nevertheless, it can be used as a guide to define the parameters of student outcomes.

In the case of Australia, the HEC document similarly suggests that an important focus in the characteristics of quality of higher education would be its outcomes. The document indicates that graduate outcomes can be assessed in terms of the acquisition of the following attributes:

- generic skills which all graduates should acquire from their respective institutions irrespective of their discipline or field of study and these include learning skills, effective communication, logical and lateral thinking, problem solving, intellectual rigour, and other socially relevant qualities such as knowledge, team building, and life-long pursuit of knowledge;

- “a body of knowledge” which should be mastered in a particular discipline or field of study for the level and type of awards and which facilitates the acquisition of new knowledge;
- professional/technical or job related skills which can be applied by graduates immediately to their employment and these include ability to work with minimum supervision in the specific field, and to apply learning to the workplace.

Once the parameters of the quality of student outcomes have been defined, the techniques for its measurement can be developed.

Issues In Measuring Quality Of Student Outcomes

In assessing the quality of student outcomes, it should be recognised from the outset that this task must be ongoing to take account of changing circumstances, that it is multi-faceted, and that different key groups may have different perspective determined by their different needs and agendas at a given point in time. In addition, the entire situation is inherently dynamic rather than static since the pursuit of new outcomes, arrived at by whatever process, will inevitably create new perceptions of success and failure within the system by all interested parties. For example, the employers' wish for more readily useable, job-related skills in graduates can only come at the expense of other attributes which current graduates possess. The value of these more latent attributes may only become evident when graduates lacking in them enter the workforce, or have been in the workforce for some time and face new and unpredictable challenges to their resources. Then the whole cyclical process will begin again *ad infinitum*.

Terenzini (1989, p.646) has identified three major areas of concern that need to be considered in assessing the quality of student outcomes:

- definitional issues,
- organisational and implementation issues, and
- methodological issues.

Definitional issues

Historically, the term assessment implies more than one measurement or approach to evaluation. An assessment of student outcomes, then, involves the acquisition of multiple evidences of those outcomes. These evidences, as suggested by Pascarella and Terenzii (1991), are provided by the following: verbal skills, quantitative skills, and knowledge of specific subject matter; general cognitive competence and cognitive skills; self-conception and self-evaluation; psychosocial characteristics and personality traits; attitudes and values; moral reasoning, moral judgement, and moral behaviour; educational attainment; career choice and career attainment; economic values and benefits; and non-monetary benefits, life satisfaction, and quality of life. Furthermore, the definition of assessment is connotative and this may contribute a significant threat to the success of any assessment effort. In any assessment, there are three important questions to be considered, viz, what is the purpose of the assessment?, what is to be the level of assessment?, and what and who is to be assessed? The clarification of these questions needs attending to before commencing the assessment. Otherwise, misunderstanding and conflict may emerge to mar the progress of the assessment procedure.

Organisational and implementation issues

The assessment of quality should be fully supported on an institutional-wide basis and as such proponents should mobilise support right from the beginning through collegial consultation. It is important that the assessment is seen by those concerned as a vehicle for individual and institutional improvement rather than as a hidden agenda to evaluate individual faculty staff or to cut budgets, or retrench staff and programs. To gain the confidence of those concerned, it is essential to clarify and to specify the data to be collected, by whom, and for what purposes. It is also useful to indicate how the data will be used, to whom it will be available and under what conditions it will be made available. Also the assignment of coordinating responsibility should be made clear. Overall, the whole tone of the procedure should be made positive and less threatening as much as possible.

Methodological issues

Bogue and Saunders (1992) have listed the following issues which relate to the design and development of a student outcomes assessment process:

Matching assessment to outcomes definition

Given the great variety of outcomes assessment models available from both the institutional and commercial domains, a fundamental question to ask is whether the assessment model or approach matches the outcomes desired and defined by the institution. Trade-offs may be necessary for an institution using a commercially published instrument.

Ensuring validity and reliability

This involves the scientific rigour of the research design, especially in the sampling procedure. To ensure validity and reliability of the data to be collected, the sample design should provide a clear description of the target population, the objective procedures used to select the sample, the stratification decisions, and the procedures used to minimise the dangers of bias through non-response. It is important to involve the faculty in the design and implementation of the process as well as in the interpretation of results. For any reasonable level of reliability in assessing a particular student's outcome, at least 8-10 items have to be attributed to that outcome variable.

Selecting score options

This depends on the research objectives and the survey instrument used. Since multiple measures will be used in the assessment procedure, the scores must be compatible so that the results can be meaningfully be interpreted.

Evaluating time and money costs

There are four areas to be considered in evaluating time and costs. They are the assessment instruments, administration of the assessment procedure, analysis and coordination, and the startup costs.

Isolating sources of bias

Bias may arise from design limitations, measurement difficulties, and statistical hazards. Sample size must not be too small to encounter problems of validity or statistical stability. Hence it is prudent to start small, perhaps on a pilot basis, to explore the problems and issues of conducting the assessment.

Examining levels of difficulty

The pilot study should be able to elicit any inherent difficulties that have been overlooked in the design stage.

The HEC (1992, p.9) has recognised the difficulty of measuring the generic attributes of graduates. In principle, outcomes like higher-level conceptual skills, independence of thought, and intellectual curiosity are less readily measurable than the more specific skills and knowledge. As a result, it is not surprising that the more readily measured parameters of competency-based learning have been recommended for outcomes determination of higher education. To what extent competency measures are valid determinants of quality of student outcomes is still vigorously debated in the higher education sector.

Methodology Used in SUT's Pilot Study

The National Board on Employment, Education and Training (NBEET) report (1991) entitled "The quality of higher education" identified three main interested parties in the outcomes of higher education, namely, employers of graduates, staff and students of the higher education system, and finally, governments on

behalf of the community. The NBEET's second category is not appropriate for assessing outcomes as it confounds providers of the service, namely staff with recipients, namely students, and is therefore essentially two distinct interested parties. Government's interests too, certainly at present, are essentially dependent on and subsidiary to the satisfaction of the other groups. We therefore believe that we take note of the views of the three groups in our research, namely the primary providers of higher educational services, academic staff, the primary recipients of the service, students and the secondary recipients of the service, graduate employers. At a seminar which we organised at SUT to foster debate and cooperation on these issues under the auspices of the Australasian Association for Institutional Research (AAIR), some participants pointed out that students are not only primary recipients, they are also secondary providers since they bring resources to the educational process in the form of individual ideas, experience, and support for fellow students. We gratefully acknowledge that perceptive contribution and accept its validity.

Having clarified these general issues, our first task is to identify which generic skills are perceived to be:

- important in the eyes of our three interested parties and
- the proper concern of universities.

To this end, we have looked at overseas attempts to measure outcomes (for example, see Pascarella and Terenzini, 1991; and Bogue and Saunders, 1992), at the specific concerns expressed in the discussion documents such as the HEC document on quality and so forth. We then put together a pilot questionnaire which specified possible outcomes in the form of generic skills. We will also seek the views of respondents within the three groups in the pilot work to enable us to extend or otherwise modify this list as appropriate.

The results from this questionnaire will enable use to identify common concerns among our three interested parties as well as differences in perceptions of the proper outcomes of university education. It seems probable that there will be areas of both agreement and disagreement on the propriety of specific concerns for outcomes of higher education and these need to be known in order to inform and further the debate on quality.

To the extent that the three interested parties lack consensus on the perceived proper outcomes, then there will be endemic frustration with the system and/or its products. These issues will then have to be debated and resolved until it is appropriate for the higher education system to resolve them and to the extent that it is capable of resolving them. One of the problems one might anticipate is that it is perhaps inevitable that any perceived shortcomings of graduates will be attributed to their university experience, which is the final exit point in a three-tier system of education. University academics are unlikely to regard that as an equitable perception if the complaints refer to knowledge and skills which, in their view, students should have acquired at an earlier point in their education.

The next stage in the research focuses on specific generic skill outcomes. We endeavour to measure the perspectives of our three interested parties on these outcomes. We have developed a pilot scale to measure communications skills, which is a generic skills outcome commonly mentioned as desirable in previous work and clearly fundamental to success at almost any level of graduate functioning. The final page of the questionnaire is an open-ended question which seeks the stakeholders' thoughts on extensions or improvements.

One of the things that eventually struck us in our deliberations and determined the proposed shape of our research is that enhanced communication skills is an outcome that one would predict that all interested parties would claim as a desirable outcome. None the less, employers constantly refer to skills in communication as a common deficit among graduates. The problem of definition in measuring the quality of outcomes has been noted as endemic. We therefore have assumed in our approach that there may be a problem of definition with this commonly used term and indeed that this will probably be the case with other so-called generic skills. These may transpire to be a series of skills with limited spheres of general application. Thus different parties may focus on different communication skills while all claiming allegiance to the same flag in this respect.

We have attempted to define communication skills outcomes in three ways. Our first and second ways basically refer to “the words” of communication — speaking to and writing for other major groups of individuals. Following this logic, we thought it appropriate to tap what we might call “the music” of communication, namely the ability to mix socially with others in the various major categories as a fundamental lubricant of communication. What we are trying to do is cover all the logical possibilities for major categories of communication skill and we intend to use the responses from our three interested parties to construct a matrix of communication skills outcomes from the perspective of the three interested parties. Of course, we intend to develop scales to assess other outcomes in the future, informed by our experience in looking at communication skills. We believe that our approach is suitable for general application within the higher education system and we hope that it will provide a useful methodology to assist the evolution of the debate on the quality of outcomes in higher education in Australia.

Problems With Measurement

A pilot study to trial and evaluate possible measurement instruments to probe the issue of quality of tertiary outcomes was conducted in mid-1991. At the time of writing the data is being processed for analysis. However, preliminary feedback from respondents to our questionnaire and observations by our research assistant indicate that the instrument which we are developing to measure the importance which members of our three stakeholder groups attach to various generic skills outcomes of university education is a useful and valuable one which will require minor modification only to yield good quantitative data.

An instrument was also developed to endeavour to tease out the meaning of “communication” to members of the stakeholder groups in terms of the modes of communication involved (verbal, written, social) and the kinds of target groups for whom the communication is intended (people of the opposite sex, people of a different educational background and so forth).

The rationale for developing a measure of one generic skill at this stage was:

- * to pilot a measure of generic skill in order to facilitate and guide future research into other generic skills;
- * in recognition of the frequency with which the skill of communication has been raised as an issue in the debate on the quality of outcomes in tertiary education;
- * in anticipation of the expected result from the first scale investigating generic skills in general that all stakeholders would endorse communication as a desired outcome.

Our feedback on the communication skills questionnaire highlights the general difficulty which we had anticipated, namely that all stakeholders tended to endorse every aspect of communication as important and therefore found the questionnaire repetitive and some respondents commented on the “motherhood” nature of the questions.

The problem is vexing since

- all stakeholders appear to endorse the importance of communication skills; the students say they regard them as important, the academic staff say that they endeavour to impart them and the graduate employers emphasise their importance in the workforce;
- despite this unanimity, however, the employers consistently mention communication skills as notably problematic among the graduates they employ.

Our best guess is that a similar state of affairs will probably emerge with other “broad-band” generic skills such as problem-solving, which has also emerged with some regularity in the quality debate. Our assessment is that the various stakeholders are using the same words in respect of some generic skills, but are perhaps referring to essentially different processes under the same rubric or different aspects of very broad generic skills.

It was our anticipation of this problem which led us to attempt to break down the broad concept of communication into a number of component parts in our pilot study, as described above, and to endeavour to quantify the importance which stakeholders attributed to these aspects as outcomes of tertiary study. This essentially logical approach, however, still does not break through to the core of the problem since all parties still appear to regard the component parts with relatively equal and high desirability.

For these reasons, it seems that the most productive approach to unravelling the intricacies of this paradoxical difficulty will lie in a more qualitative rather than quantitative approach. What we need try do is to allow the stakeholders to express the perceived nature of their understanding of various generic skills as they operate within the ecology of their own environments. An analysis of these responses should indicate the particular skills which different stakeholder groups are focusing on under the rubric of a given generic skill. In turn, such an analysis should point to means of solving the difficulties which graduate employers are now experiencing, although one should not prejudge either the locus or the nature of such solutions until further data become available.

Conclusion

Changes in the world economic order as effected by global political power shifts and technological advancement have brought about changes in employment patterns, education, social interactions, and lifestyles. In the United States, it is common to have several job changes within one's working career. As such, academic specialisation in a discipline could handicap a person's career path where the acquisition of generic skills and multi-skilling is in vogue. In Australia, employers are saying that the area of generic skills is a problem. The implication is that the educational system is not producing the right kind of product. Those responsible for the management of tertiary education have definitely heard the message clearly, as reflected by the release of several government-sponsored documents on the quality of higher education.

Acknowledgments

The authors wish to thank the Deputy Vice-Chancellor of Swinburne University of Technology, Mr Frank Bannon, for his enthusiastic support for the project. Our gratitude is also extended to the Victorian AAIR members who attended the local seminar on quality and made some suggestions to the authors regarding the study. We wish to acknowledge the services and contribution of Mr Steve Andrew who conducted the pilot study referred to in the paper and to its evaluation. Finally, Gan Che wishes to thank the Department of Social Sciences, La Trobe University College of Northern Victoria for providing partial funding to present this paper.

References

- Astin, A. (1973a) "Measurement and determinants of the outputs of higher education" in L. Solmon and P. Taubman (eds) *Does college matter? Some evidence on the impacts of higher education*, Academic Press, New York.
- Bogue, E.G. and Saunders, R.L. (1992) *The evidence for quality*, Jossey-Bass, San Francisco.
- Brown, R. and DeCoster, D. (1982) *Mentoring-transcript System for Promoting Student Growth*, JosseyBass, San Francisco.
- Ewell, P. (1984) *The Self-regulating Institution: Information for Excellence*, National Center for Higher Education Management Systems, Boulder.
- Ewell, P. (1985a) "Assessing educational outcomes", *New Directions for Institutional Research*, no. 47, Jossey-Bass, San Francisco.

- Ewell, P. (1988) "Outcomes, assessment, and academic improvement: In search of usable knowledge" in J. Smart (ed). *Higher Education: Handbook of Theory and Research*, Vol. 4, Agathon, New York.
- Feldman, K. and Newcomb, T.L. (1969) *The Impact of College on Students*, Jossey-Bass, San Francisco.
- Hall, W. (1992) *Getting to Grips with Quality*, TAFE National Centre for Research and Development, Leabrook.
- Hanson, G. (ed.) (1982) "Measuring student development", *New Directions for Student Services*, no. 20, JosseyBass, San Francisco.
- Higher Education Council (1992) *The Quality of Higher Education: Draft Advice*, Better Printing Service, Queanbeyan.
- Johnston, R.J. (1991) "Graduate evaluation of British higher education courses: an exploration", *Studies in Higher Education*. Vol. 16, No. 2, pp. 209-224.
- Kalkwijk, J.P.Th. (1991) "Quality assurance in higher education in The Netherlands", paper presented at Conference on Quality Assurance in Higher Education, Hong Kong, July.
- Lindsay, A. (1992) "What is the agenda we call quality?" *The Australian*. 11 March, 1992.
- Malkova, Z.A. (1989) "The quality of mass education" *Prospects*, vol. 19, no. 1, pp. 34-45.
- Parker, C. and Schmidt, J. (1982) "Effects of college experience" in H. Mitzel (ed.) *Encyclopedia of Educational Research*, 5th edn, Free Press, New York.
- Pascarella, E.T. and Terenzii, P.T. (1991) *How College Affects Students*, Jossey-Bass, San Francisco.
- Ross, K.N. and Posthethwaite, T.N. (1988) "Planning the quality of education: different information for different levels of decision-making" *Prospect*, Vol. 18, No. 3, pp. 315-331.
- Tan, D. (1986) "The assessment of quality in higher education: a critical review of the literature and research" *Research and Higher Education*, Vol. 24, No. 23.
- Terenzii, PT. (1989) "Assessment with open eyes: pitfalls in studying student outcomes" *Journal of Higher Education*, Vol. 60, No. 6, pp. 644-664.
- Vroeijenstijn, A.I. (1991) "External quality assessment: servant of two masters" paper presented at Conference on Quality Assurance in Higher Education, Hong Kong, July.