Our programs are good... because our students say they are

Tom Molyneaux
School of Civil, Environmental & Chemical Engineering, RMIT University, Melbourne, Australia
Tom.molyneaux@rmit.edu.au

Margaret Jollands
School of Civil, Environmental & Chemical Engineering, RMIT University, Melbourne, Australia
m.jollands@rmit.edu.au

Lesley Jolly
Strategic Partnerships
ljolly@bigpond.net.au

Abstract: The program in Civil and Infrastructure Engineering in the School of Civil Environmental and Chemical Engineering at RMIT commenced in 2004 following an intensive period of curriculum design/development. The program was based on graduate capabilities and sustainability principles that acknowledge the full life cycle of infrastructure. The graduate attributes were based on graduate capabilities as defined by Engineers Australia in their accreditation process augmented by consultation with industry. The basic teaching paradigm was also changed to encompass project-based learning – widely recognised as an effective means of developing the graduate capabilities required. Now with the first graduates active in industry, a project to evaluate the effectiveness of the program is raising questions as to the very process of curriculum evaluation. This paper describes the approach being adopted and the rationale behind it. Current evaluation practice and the value of several commonplace quality indicators are discussed with a view to stimulating discussion.

Introduction

General

In 2003, following a sequence of years in which quality indicators showed that the current Civil Engineering program was not attractive to industry or to students, it was decided to close the program. This served as a trigger for program renewal and it was decided to develop a more modern and attractive program with a broader focus and an emphasis on sustainability. Consultation was undertaken with Industry, RMIT Applied Science Schools (such as Physics, Mathematics, Chemistry, Geospatial), the School of Constructed Environment, the University Library, the engineering vocational education and training (VET) sector, the RMIT Centre for Infrastructure Research and Management and with students through forums and questionnaires. The new program was based on graduate capabilities and sustainability principles that acknowledge the full life cycle of infrastructure. The graduate attributes were based on the graduate capabilities for an education program as defined by the Institution of Engineers Australia (2007) in their accreditation process augmented by the consultation with industry. These attributes are in line with the requirements of a professional graduate engineer (Engineers Australia, 2009)– defined in terms of Engineer Australia’s “Stage 1 competencies” (first published at this time). The basic teaching style was also changed to encompass project-based learning – recognised as an effective means of developing the graduate capabilities required.

This program and (in subsequent years) programs in Environmental and Chemical Engineering were rolled-out over the following 4 years. Now there are graduates in industry from all the disciplines and the time is right to ask whether the programs are effective. This question has raised discussion as to how to evaluate the programs – how do we know our programs are effective?
Program evaluation – how do we know our programs are effective

There are several commonly adopted criteria by which academics and education managers measure the effectiveness of their teaching processes and curricula.

Our programs are good… because our students say so

Course Experience Survey

Student evaluations of teaching as performance indicators have been used in the USA since the 1920s and have been commonplace worldwide since the 1970s (Popham, 1988). The basic form of the Australian student response evaluation (the Course Experience Questionnaire, CEQ) was developed for graduates, and tested in Australian universities during 1989 then added to the Graduate Destination Survey in 2004.

Government funding is linked to such indicators of institutional performance and the results are used by students (National and International) to select universities/programs. Thus it is not surprising that a quick survey of university web sites demonstrates that the vast majority have developed similar ‘in-house’ versions of their own and apply these at the individual course (subject) level. Many such systems focus on a particular extraction from the results that provides a simple quantitative assessment based on 6 questions relating to feedback - the Good Teaching Scale (GTS). Marsh (1987) reviewed student evaluations in higher education and commented that the methods were generally very sound and produced useful information – however he warns about the combination of results for the sake of expediency. He advises that despite the generally supportive research findings, student ratings should be used cautiously, and there should be other forms of systematic input about teaching effectiveness, particularly when they are used for tenure/promotion decisions. He concludes that, while there is good evidence to support the use of students’ evaluations as one indicator of effective teaching, there are few other indicators of teaching effectiveness whose use is systematically supported by research findings. The study undertaken by the Mathematical Sciences Education Board (MSEB, 2004) comments on the “corruptibility of indicators” in situations where the performance of teachers or administrators is judged by such systems.

Course Experience Survey – the RMIT experience

Undergraduate programs at RMIT comprise 8 courses (subjects) per year each of which are evaluated using the RMIT Course Experience Survey (CES). The CES employs the same questions as those used to assess programs in the national CEQ and involve assertions answered by a five point scale between ‘strongly agree’ and ‘strongly disagree’. Many new initiatives (Othman et al 2010, Li & Molyneaux 2009, Molyneaux 2009) have resulted in significantly improved student feedback. As a consequence the Good Teaching Scale has consistently improved over the years – with many values up at 80-90%. The current mean score of approximately 66% suggests that two thirds of all students agree with the six factors that contribute to the GTS (the Overall Satisfaction Index (OSI) is generally even better). So clearly the students’ perception is that the programs are ‘good’. Does this imply that are programs are successful/effective?

The value of student perceptions

There is an extensive body of research that supports the view that the CEQ and such student surveys are indicative of good teaching practice (for example Cohen 1981, Marsh 1987, Elphinstone 1990). However the implied relationship between the students’ perceived level of quality of teaching and the quality of learning is much more complex. More recently Lizzio et al (2002) investigated the relationship between university students’ perceptions (through CEQ) of their academic environment, their approaches to study, and academic outcomes. They found that the strongest predictors of satisfaction were a learning environment which was perceived as involving (good teaching) with clear expectations (clear goals), and allowing of a degree of choice to pursue individual interests (independence). In terms of their academic achievement - only three aspects of the academic
environment (clear goals, good teaching, and appropriate assessment) were significant predictors of their level of academic achievement. Hence, higher achievers perceive their teaching to be good.

Modern engineering curricula are characterised by the use of student centred learning, Project (and Problem) Based Learning (PBL) with the development of many skills embedded in courses and mapped throughout the program. These attributes mean that students taking a university course (subject) cannot easily comment on the applicability of the course content with regard to vocational relevance or take a longer term view (over the 3 or 4 years). In addition, within the engineering profession, undergraduates are at the very start of their career - and hence are not best qualified to comment on such matters as the relevance of content and such issues as effectiveness of various teaching paradigms. Consequently their perceptions based on surveys at the course (subject) level at best reflect on the teaching at that level. Can the amalgamation of all the results for a program be used to adequately evaluate that program? Do they reflect the overall effectiveness of the program?

Given the widespread use and with the encouragement/pressure to achieve good outcomes it should come as no surprise that such results are presented as evidence in many research publications in teaching and learning. It could even be considered in some cases that an improvement in student perception (in particularly, GTS) has been the main goal of an intervention or teaching initiative. This opens up the possibility that some initiatives may improve survey outcomes but not affect learning - a case of treating the symptoms. The observation that courses with good teaching practice receive favourable student response does not necessarily mean that innovations that produce improvements in response also improve learning. In fact there is a wealth of examples and evidence where this does not occur. Goldfinch et al (2009) conducted a literature review on factors influencing learning of mechanics and discusses innovations that improve engagement and interest commenting that they often receive positive feedback but demonstrate no improvement in capability (Crawford & Jones, 2007, Balascio et al, 2007). Experience at RMIT supports these findings with several cases where initiatives have improved the GTS but not improved student performance in assessments. It would appear that the students perceive that their ability has improved – the gap between their perceived ability and their actual ability may have actually widened.

As a stakeholder the student choosing a university and program is taking the first step in becoming a professional engineer – the student has a vested interest in the quality of the program/university and uses published data (student/graduate perceptions, CEQ etc) to make this decision. Consequently the CEQ and similar processes play an important and recognised role. However their role in curriculum evaluation is not so clear.

Our programs are good… because Engineers Australia says so -

The accreditation process

In Australia, Engineers Australia is responsible for the accreditation of engineering degrees under the various accords that define the three levels of engineering professionals. The Washington accord covers the education of professional engineers - signed in 1989 it was the first of three recognising equivalence in the accreditation of qualifications in professional engineering, normally of four years duration (International Engineering Alliance, 2009). The Sydney and Dublin accords cover engineering technology and technician engineering. All the accords are overseen by the International Engineering Alliance. The Alliance approves national organisations as being responsible for the accreditation of degrees in that country – the EA in Australia. Outcome based criteria for evaluating university programs have been developed for each accord and several regulatory bodies (including EA) have developed competency based standards (outcomes of the programs). The description of competency profiles (International Engineering Alliance, 2009) states that a professionally or occupationally competent person has the attributes necessary to perform the activities within the profession or occupation to the standards expected in independent employment or practice. The professional competency profiles for each professional category record the elements of competency necessary for competent performance that the professional is expected to be able to demonstrate in a holistic way at the stage of attaining registration. The main focus of the Engineers Australia accreditation process is assessment of how adequately the competencies expected of a graduate (Australian Engineering Competency Standards - Stage 1) are developed throughout the duration of
the program. In addition the accreditation process is used to stimulate innovation and to disseminate best practice and diversity in engineering education. The accreditation process does not prescribe detailed program objectives or content, but requires engineering education providers to have in place their own mechanisms for validating outcomes and continually improving quality (Institution of Engineers Australia, 2007). The evidence used to assess the programs includes published program rationale, management structure, staffing levels, library, IT and laboratory facilities, syllabi, course descriptions, a mapping of the assumed development of competencies through the program, examples of students’ work, interviews with staff, students and industrial liaison bodies.

Accreditation as a means to evaluate education was historically the most prevalent form of curriculum evaluation of schools in the USA but has declined in popularity since the 1980’s. The main reason for this is that despite the intuitive support for the proposition that certain input factors are linked to outcomes of an educational process, there is a scarcity of well evidenced research (Popham, 1988). The accreditation process in effect undertakes a quality assurance appraisal of the system (the machine, the process) against criteria that (in the case of engineering) are predetermined by the profession and the educational institution. It does not directly appraise the product but assures that they have been through a process that was designed to produce the desired output.

**RMIT and accreditation**

The schools’ programs were developed in a collaborative process with Engineers Australia, industry, graduates and taking into account modern educational principles at a time when EA were developing their Stage 1 Competencies. Thus at our first full accreditation of the running programs in 2009 it came as no surprise that the school’s programs were rated highly. Consequently – our machine/process is in good order – does this imply that we are achieving our educational targets? The curriculum is being delivered as intended – but is it having the desired influence on the students as they pass through? Are our graduates work-ready?

**Our programs are good... because well accepted methods of education evaluation have demonstrated so**

Education evaluation is a well established process spanning back to the 1930’s in the USA (Popham 1988, Kellaghan 2003, Stufflebeam & Shinkfield 2007). The process has become a major field of research and practice driven by governments desire to demonstrate value for money, control quality, guide improvement, and protect consumers. In the 1960s the realisation that the evaluation was not contributing to or assisting curriculum development became a driving force. In particular Cronbach (1963), with a view to improving education, advised that evaluations should examine processes and report on information that could be used to guide curriculum development – an opinion that took some time to become accepted. By the 1970s evaluation was becoming a recognised profession with guidelines, and ultimately standards.

There are five models of evaluation recognised by the profession (Popham 1988)

- Goal attainment models
- Judgemental models – focussing on process
- Judgemental models – focussing on product
- Decision facilitation models
- Naturalistic models

The goal attainment models reflect Tyler’s approach (Popham, 1988). Goals for the student, society and the subject matter are interpreted as measurable changes in behaviour of the students. These changes in behaviour are then measured and attained goals reflect a successful education program. A danger of this approach is that goals can be changed or that goals might be inappropriate.

The next two categories involve assessment where professional judgement plays an important role. Engineering education accreditation (above) is a judgemental process focussing on process - employing professionals and professionally developed criteria. The success of judgemental evaluations depends on the view of the evaluator. Alternatively, judgemental approaches that focus on outputs or product may include such approaches as:

Proceedings of the 2010 AaeE Conference, Sydney, Copyright ©, Tom Molyneaux, Margaret Jollands, Lesley Jolly, 2010
• A judgement of the merit of the goals – if the goals are not relevant then it is of no interest as to whether they are met.
• Goal free evaluations – in which the focus is on the actual outcomes of the program – intentional and unexpected. The approach involves drawing inferences about the likely outcomes based on a judgemental (professional) view of the components of the program. The approach reveals a wider range of outcomes (including incidental outcomes) than when the evaluator is focussed on pre-determined goals. The approach includes such research/evaluation techniques as observation and unstructured interviews.

Decision facilitation models are evaluations aimed at supporting educational decision makers. Such processes and evaluators are concerned with providing data for the decision making process. Within this group there are internal (formative) and external (summative) focuses. Formative evaluation is intended to improve courses through evaluation and should provide information that is of use to curriculum developers. Summative feedback is retroactive and is driven by the need for accountability – arguably the Australian national survey (the CEQ) is in this category.

The final category of naturalistic or qualitative is profoundly different from that of the classical experiment/investigation approach. The approach has some similarities to that of grounded theory in so much as rather than developing a hypothesis and then seeking data, data are collected and patterns and outcomes sought/ explored. Anthropological/ethnographic principles are adopted. The approach would typically use interviews, observations, career histories, and surveys of behaviour, opinions, knowledge, preference, values, feelings etc.

A more focussed approach to the design of methodologies for educational program evaluation is presented by Owen (2007). The author distinguishes evaluations which are generally commissioned and hence have clear stakeholders from other investigations/appraisals or research motivated by the desire to improve understanding and accountable more to the education/scientific community at large. Five categories of evaluation are defined: proactive (at the program concept stage), clarificative (exploration of relationships between objectives and outcomes), interactive and monitoring (appraisal of delivery process, QA), and impact evaluation (assessment principally for accountability). The clarificative and impact approach is of relevance to engineering curriculum evaluation - how do the outcomes (intentional and incidental) relate to the program objectives and approach. The author describes a program logic approach to clarificative evaluation whereby the causal thinking that led to the program development is examined and an explicit conceptualisation of how the program causes the outcomes developed - this is then used to guide the evaluation. The analysis would typically split the process into sequential components such as inputs (resources, students, academics), activities/processes (teaching, learning activities), outputs (immediate educational outputs) and finally outcomes (longer term consequences). These are presented in the form of tables or flow charts. Owen avers that it is necessary to conduct this type of study before monitoring or impact evaluations are undertaken. This approach is widely used to assess the impact of projects beyond those of curriculum delivery (for example, Wolf 2008, Australian Learning and Teaching Council - ALTC 2010).

**Approaches to program evaluation relevant to engineering education**

An engineering student’s learning is a complex phenomenon involving interaction of various educational influences many of which take place outside of formal teaching activities and result in outcomes that were not intended (Walther, 2008). This implies that in terms of curricula development the learning cannot be controlled in a deterministic manner and in turn any evaluation of the effectiveness of such curricula should examine actual outcomes – intended and incidental.

Table 1 illustrates the wide range of approaches used in attempts to ask deeper questions about the actual effectiveness of educational programs – that is to understand what outcomes can actually be attributed to educational interventions.
Table 1: Research approaches adopted in evaluating curricula

<table>
<thead>
<tr>
<th>Source</th>
<th>Methods</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meier at al (1999)</td>
<td>Industry surveys and focus groups</td>
<td>Identified skills gaps</td>
</tr>
<tr>
<td>Collins (2008)</td>
<td>Survey of 1622 employers</td>
<td>Perceptions of pre and post PBL graduates</td>
</tr>
<tr>
<td>Seidel et al (2006)</td>
<td>Questionnaires and observations in class</td>
<td>Effectiveness and development of incidental competencies</td>
</tr>
<tr>
<td>Shi et al (2006)</td>
<td>Linguistic analysis</td>
<td>Development of research skills, time management, leadership and teamwork</td>
</tr>
<tr>
<td>Bons and McLay (2003)</td>
<td>Interviews with employers and academics, Surveys of 467 graduates</td>
<td>Factors in preparing graduates for employment</td>
</tr>
<tr>
<td>Brunetti et al (2003)</td>
<td>Interviews of students and employers about awareness of sustainability issues</td>
<td>Effectiveness of PBL in teaching social and environmental responsibility</td>
</tr>
<tr>
<td>Tilli and Trevelyan (2008)</td>
<td>Interviews, diary studies</td>
<td>Effect of non-technical skills on technical development</td>
</tr>
<tr>
<td>Kennedy (2009)</td>
<td>Structured interviews</td>
<td>PBL and retention of engineers in profession</td>
</tr>
<tr>
<td>Grippa and Gloor (2009)</td>
<td>Social network analysis</td>
<td>Workplace networks and professional development</td>
</tr>
<tr>
<td>Lin and Dumin (1986)</td>
<td>Workplace interviews, SNA</td>
<td>Network development and professional advancement</td>
</tr>
<tr>
<td>Schmidt et al (2006)</td>
<td>Questionnaires</td>
<td>PBL graduates compared with non-PBL graduates</td>
</tr>
<tr>
<td>Tamblyn et al (2005)</td>
<td>Test of current knowledge</td>
<td>Retention of expertise in PBL graduates Vs non-PBL graduates</td>
</tr>
</tbody>
</table>

Very few studies have compared graduates of PBL and non-PBL programs from similar programs and none (to the knowledge of the authors) can boast such a significant number of participants as the last three examples in Table 1 – so it is disappointing that the results appear to be contradictory. The differences in the approaches adopted are subtle. In one case a peer review is used to detect the ability (of doctors) to maintain competence (continued professional development) while in the other they were assessed directly in a field where knowledge is changing (by a test). Norman et al postulate that maintaining competence may be a consequence of many factors as opposed to maintaining knowledge in a particular field. They concluded that to adequately investigate the link between undergraduate education and practice would require the availability of defensible measures of workplace performance gathered over long periods of time after graduation together with significant triangulation from alternative studies looking at different aspects of behaviour (multi-source feedback). All the evidence so far, then, suggests it is very difficult to evaluate how good programs are. We conclude that no one method is effective in program evaluation. It seems likely that a range of evaluation methods needs to be used, that assess both teaching and learning outcomes.
Discussion

Driven by feedback from the profession, the last decade has seen a change of focus in engineering education to an outcome-based approach with wider educational goals that include concepts such as teamwork, communication, environment and sustainability. Besides the focus on outcomes, modern engineering programs are characterised by such pedagogical approaches as inclusive, student-centred teaching, integration with industry, inclusion of real-world problem solving, PBL, etc. However, whilst there are systems in place to assure that the sub-program components (courses/subjects) are delivered well (student survey feedback etc); it is unclear how well the program educational targets (workplace requirements) are being achieved in practice. This latter question is a more complex issue and consequently good practice has not yet been established. The sections above suggest that for an approach to adequately evaluate such programs it should examine not only how students develop throughout their undergraduate program but also during their early exposure to the work place – as it is there that the effect of the curriculum will become apparent.

The RMIT engineering program evaluation

An assessment of the programs within the school is underway. The approach focuses on the main product of the new programs (characterised by the PBL approach) – the graduate and their acceptance into the profession. A direct comparison between graduates of the ‘new’ PBL courses and the old programs is being sought. The measure of ‘success’ being investigated is the degree to which the programs serve the need of both the graduate in terms of developing their career and the needs of industry. Graduate behavioural characteristics are being explored/classified with a view to attributing the behaviour to particular teaching interventions/influences throughout their university program. The mechanism of the research is interviews and focus groups of students throughout the program, interviews of graduates of our new (PBL) and our old (pre-PBL) programs and their employers.

The approach being adopted has much in common with the program logic/theory approach as described by Owen (2007) and the Mathematical Sciences Education Board (MSEB, 2004) together with aspects of goal free evaluation (Popham 1988, Kellaghan 2003, Stufflebeam & Shinkfield 2007). The program logic is well articulated as a result of the thorough curriculum design process that took place in 2002-3. That process developed a rationale for the development of the desired outcomes through the four years of the program – indications of these outcomes will be sought by examining the interview transcripts of the graduates and their employers (in-line managers). Recognising however that there may well be outcomes (desirable and undesirable) that are incidental and not attributable to the program logic the interviews are of an open nature with a view to applying a goal free approach to reveal such phenomenon.

The authors are conscious of the lack of well accepted practice in this particular area; however they are aware of plans to conduct similar evaluations of PBL based curricular by other institutions. As advised by the Mathematical Sciences Education Board study (MSEB, 2004) and indicated by studies of medical graduates (Schmidt et al 2006, Norman et al 2008 and Tamblyn et al 2005) it is by such triangulation that reliable conclusions can be made regarding causal relationships between curriculum and outcome.

Conclusions

1. Course evaluation questionnaires, whilst serving as a means of Quality Assurance, are not a reliable means of curriculum evaluation.
2. An assessment of an engineering program should include studies of the graduates in their workplace.
3. A program logic approach is an effective approach to evaluate the outcomes of modern engineering curricula because the development of the outcomes is mapped through the duration of the program.
4. There is no one best practice approach to the evaluation of engineering curricula. Reliable results can be obtained by comparing evaluations of the same curricula delivered elsewhere and assessed independently.

5. Evidence of both intended outcomes (in-line with the program logic) should be sought, together with unintended outcomes – possibly using a goal free approach.

Acknowledgements

The authors would like to thank Andrew McLean for his contribution to this study in conducting and transcribing interviews, formulating questions and undertaking an extensive literature review. The study has been supported by funding from the School of Civil, Environmental and Chemical Engineering.

References


Cronbach, L. (1963). Course improvement through evaluation. Teachers College Record, 64, 672-683


Grippa, F. & Gloor, P. (2009). You are who remembers you: Detecting leadership through accuracy of recall. in Social networks (in press), accessed 01/10/09 from
Our programs are good… because the students say they are…

Association for Engineering Education: Creativity, Challenge, Change; Partnerships in Engineering Education, Auckland University of Technology


Copyright statement

Copyright © 2010 Tom Molyneaux, Margaret Jollands, Lesley Jolly: The authors assign to AaeE and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to AaeE to publish this document in full on the World Wide Web (prime sites and mirrors) on CD-ROM or USB, and in printed form within the AaeE 2010 conference proceedings. Any other usage is prohibited without the express permission of the authors.