Higher Education Research and Development Society of Australasia, Inc

Research and Development in Higher Education: Reshaping Higher Education Volume 33

Refereed papers from the
33rd HERDSA Annual International Conference

6–9 July 2010
Melbourne, Australia

Johnson, E., Maddox, L., Quinton, J. & Burke da Silva, K. (2010). Attitudes to assessment in university Science education. In M. Devlin, J. Nagy and A. Lichtenberg (Eds.) *Research and Development in Higher Education: Reshaping Higher Education*, *33* (pp. 347–357). Melbourne, 6–9 July, 2010.

Published 2010 by the
Higher Education Research and Development Society of Australasia, Inc
PO Box 27, MILPERRA NSW 2214, Australia
www.herdsa.org.au

ISSN 0 155 6223
ISBN 0 908557 80 9

This research paper was reviewed using a double blind peer review process that meets DIISR requirements. Two reviewers were appointed on the basis of their independence and they reviewed the full paper devoid of the authors' names and institutions in order to ensure objectivity and anonymity. Papers were reviewed according to specified criteria, including relevance to the conference theme and sub-themes, originality, quality and presentation. Following review and acceptance, this full paper was presented at the international conference.

Attitudes to assessment in university Science education

Elizabeth Johnson
La Trobe University, Melbourne, Australia
E.Johnson@latrobe.edu.au

Leone Maddox
Flinders University, Adelaide, Australia
leone.maddox@flinders.edu.au

Jamie Quinton
Flinders University, Adelaide, Australia
jamie.quinton@flinders.edu.au

Karen Burke da Silva
Flinders University, Adelaide, Australia
karen.burkedasilva@flinders.edu.au

This study explores the use of assessment in the science disciplines of the higher education sector. The data are drawn from parallel studies at two Australian universities, namely La Trobe University (Melbourne) and Flinders University (Adelaide). Published documents on subjects (topics) at each university were examined to survey assessment practice. Staff attitudes toward assessment practices were gathered through surveys and interviews. The study shows that the major form of assessment in science disciplines is examination, which is usually administered at the end of the teaching period. Staff reported that examinations were the assessment tool of choice because they avoid collusion, plagiarism and have less impact on staff workloads. Furthermore, staff believe that examinations are an appropriate assessment method particularly where there is an emphasis on acquiring fundamental information and concepts in a discipline. However there is also support for a reduction in the emphasis on examinations and an acknowledgement that examinations often encourage shallow learning and do not represent real world situations. The attitudes of educators will be a key factor in development of sustainable shifts towards more student-centred assessment.

Keywords: assessment, science, educator attitudes

Introduction

It has been well-established that assessment plays a crucial role in determining the attitudes of students toward their learning (Ramsden, 1997). Most students are strategic about where they choose to invest their time (Prosser & Trigwell, 1999). Since the results of students' study are most often expressed to an external audience in the form of marks (and they expect to be judged by these marks) there is a strong incentive to tailor their study to suit the assessment criteria. If the assessment simply asks for recall of facts, students are likely to put their energies into memorisation rather than developing a more profound understanding of the principles of the subject or developing skills in application or creative analysis and problem-solving. This leads us to ask "What do science educators want students to gain from their learning experience and does this align with the type of assessment and learning activities that

are implemented?". This is captured by the idea of constructive alignment (Biggs & Tang, 1999), which suggests that alignment of the objectives of the subject, the learning tasks and the assessment gives the clearest message to the student about how they should study and what the desired outcomes are. Constructive alignment occurs when all components direct the student towards deep, and high quality, learning.

When considering assessment within the sciences, anecdotal experience suggests formal examinations are a common form of assessment in science, particularly in large undergraduate classes. This poses a problem as research suggests that examinations are generally a poor vehicle to foster deep learning (Gibbs & Simpson, 2004–5; Scouller, 1997). In a traditional written examination, the student has little or no access to external resources, has little or no time to refine or revise their work and may have little opportunity to explain or expand on their view of a subject. This is particularly problematic in examinations that use multiple choice questions where the student is very restricted in their ability to communicate their learning and the question design concentrates on recall of content. Some methodologies for multiple choice questions are designed to test higher order learning (for examples see Harper, 2003; Williams, 2006), but these innovations may not be widely used and are much more time consuming to design. In 2003, the Quality Assurance Agency for Higher Education in the United Kingdom (UK) noted in relation to assessment in UK universities that:

> reviewers commented frequently on the narrow range of assessment methods employed, suggesting that an over-reliance on examinations did not allow students to demonstrate their achievement of all learning outcomes. Although there had been clear development in this area, in recent review visits, the reviewers continued to point to the narrow range of assessment methods employed in a number of institutions (QAA, 2003, p. 27).

Nor do examinations reflect professional practice of research or applied science. This raises the question of why science educators would select examinations as the preferred mode of assessment. Is the use of examinations forced by concerns surrounding plagiarism, scarce resources and a drive to efficiency, or do science educators actively choose examinations as the most appropriate form of assessment? If we are to investigate, lead and implement change in the current assessment practice, it is important to understand what colleagues think about the assessment they use, why they use it and what they see as the obstacles to changing assessment (van Driel, Beijaard, & Verloop, 2001).

In this study, we have investigated attitudes to assessment modes amongst staff teaching undergraduate science at two Australian universities. The study produced a snapshot of the educator's perspective, which provides some indication of the factors staff take into consideration when selecting assessment tasks. The study has also collected information about science educator conceptions about assessment and the limitations of working with large classes. Since selection of the assessment task has profound effects on student learning, this study provides insight into construction of the teaching and learning environment.

Methods

This study was conducted through parallel projects undertaken at La Trobe University in Melbourne and Flinders University in Adelaide. La Trobe University was established in the 1970s and Flinders University in the 1960s. Both are members of the Innovative Research Universities Group (IRU) and are characterised by considerable diversity in the student body.

At La Trobe University, information on current practice in assessment for undergraduate subjects in Years 1, 2 and 3 was gathered from the university handbook for 2009 across the range of disciplines in the Faculty of Science. At Flinders University, information on current practice in assessment for undergraduate education in 2008 and 2009 for physics, nanotechnology and biology (including honours) was gathered from handbooks and topic (subject) proposal documentation.

Staff attitudes were measured by anonymous surveys at La Trobe University and Flinders University in 2009. At Flinders University this was carried out Faculty wide, with a second stage that focused on staff in biology specifically. Staff involved in nanotechnology and physics were independently interviewed using the survey questions. A total of 45 surveys and interviews were recorded. At La Trobe University, a short anonymous survey was sent to all continuing academic staff in the Faculty of Science, Technology and Engineering with 35 staff responses collected. Staff were asked to respond to statements about assessment in their own classes using a Likert scale and then to reflect on assessment through open comments. Although the survey instruments were not identical at the two universities, similar questions (tailored to local contexts and understanding) were asked. This snapshot does not attempt to compare representative data from all disciplines but rather to uncover factors and views influencing assessment in science subjects.

Trends from the quantitative items in each survey set were identified, open responses were classified into categories and the incidence of comments in each category was recorded. The rich data from the open comments was captured in illustrative quotations which were representative of either a widespread view or an exception.

Results

Purpose of assessment

Both studies investigated staff attitudes about the value of assessment. At Flinders University, the faculty survey was used to ask directly about the purpose of assessment (Table 1).

Table 1: Summary of staff beliefs about the purpose of assessment at Flinders University

Purposes of assessment identified by staff	No. of comments
Grading/warranting/certifying	7
Learning level of the students	10
Encouragement/engagement in learning	6
Feedback for staff and students	4

It was notable that most of the comments connected assessment with helping students learn and with recognition of the learning outcomes for both the teacher and the student. For example, a typical comment was “assessment examines student progress in grasping concepts”.

Examinations are heavily used in undergraduate science subjects

Data for the use of assessment modes in undergraduate subjects was gathered from the subject handbook for each university. This data was collated across all year levels to give a picture of

the type of assessment that students are exposed to throughout the entire degree. Examinations in undergraduate subjects at La Trobe University dominate assessment both in the frequency of use and the relative contribution towards the final mark (Table 2). This was also true for the topics investigated at Flinders University (Table 3 and Table 4).

Table 2: Frequency and weighting of assessment type in individual subjects at La Trobe University

Assessment type	% of subjects using this assessment tool	Weighting (% of final mark) for assessment type in a subject	
		Average	Range of weightings for this assessment
Examinations	90	63	25–85
Written reports/assignments	64	34	0–100
Lab/field work (including prac reports & prac examinations)	63	28	5–80
Tests/quizzes during teaching	11	21	5–60
Presentation	4	13	5–30

Table 3: The extent to which examinations are used and the weighting distribution categorised by discipline

Discipline	Flinders University		La Trobe University	
	% Subjects using examinations	Weighting of examinations across the degree	% Subjects using examinations	Average weighting of examinations in subjects[#]
Mathematics and Statistics	–	–	92	72
Biology	67	31	97	62
Physics*	94	54	100	58
Nanotechnology*	78	34	–	–
Psychology	–	–	100	52

*Note that Physics and Nanotechnology tests (generally online quizzes) have been included as examination style assessment.
#Weighting of examinations across the degree was not calculated for courses from La Trobe University due to the open course structure of the relevant degrees.

The use of examinations does vary between discipline areas (Table 3). This is partly due to cultural differences between disciplines, such as a tradition of using essays in psychology, but also due to targeted curriculum review and development. It is interesting to note the differences in the diversity of assessment used in physics and nanotechnology (Table 3). This can be attributed to the more recent development of nanotechnology as a discipline at Flinders University when compared to physics. Nanotechnology is becoming increasingly used as a

way to link different disciplines and engage students in science (Gooding & Shapter, 2008; Shapter, Ford, Waclawik, & Maddox, 2002).

When the assessment structure is investigated it was often observed that a significant amount of the assessment is summative at the end of a topic (subject) and usually appears in the form of an exam. At La Trobe University, respondents across the Faculty were asked how they would prefer to weight different assessment components if no limit was placed on resources. Approximately one half of the staff said they would prefer to reduce the weighting on formal examinations and about one quarter would replace this with assessment using assignments. This was particularly interesting since staff were quite positive about the value of formal examinations in the open text responses. When asked to suggest a better weighting for assessment the median decrease in weighting for examinations was –10%. However staff did feel it was important to maintain a high weighting for examinations and to concentrate on command of subject knowledge as the key learning outcome in many cases.

At Flinders University, 90% of staff thought that the assessment weighting should vary over the course of the degree. In Biology, approximately 80% of staff thought that formal examinations should decrease and 70% thought that practical work should increase over the course of the degree. Table 4 shows the distribution of assessment with a focus on examinations and practical work across year levels in the Faculty of Science and Engineering at Flinders University. It is interesting to note that there appears to be a discrepancy between what people believe should happen and what happens in actuality.

Table 4: Distribution of assessment weighting from the Faculty of Science and Engineering at Flinders University across year level

Year level	Exam	Practical work	Group work	Individual work
1	46	18	6	30
2	48	15	2	35
3	38	13	3	46

Learning outcomes and assessment

At La Trobe University, respondents were asked to list the learning outcomes measured by formal written examinations, practical reports, self-test quizzes and assignments (Table 5). Educators believe that formal written examinations assess the acquisition of information (facts) and fundamental concepts and problem-solving. Practical reports, unsurprisingly, were expected to assess the applied skills associated with laboratory work. Self-test quizzes were seen as measuring quite low-level learning outcomes such as “rote learning of basic facts” and “small definitions and concepts” and were thought to be of most value in providing feedback to students about their progress. Assignments were the only assessment form felt to significantly address higher-level learning outcomes such as research, analysis and synthesis.

Table 5: Learning outcomes assigned to assessment tasks

Assessment task	Learning outcomes measured	No. of comments La Trobe	No. of comments Flinders – Biology*
Formal written examinations	Information (facts), fundamental concepts	14	–
	Solving problems, applications	12	–
	Integrated view of unit	4	–
Practical reports	Application, practical skills	8	–
	Communication skills	6	15
	Integrated understanding	1	18
Self-test quizzes	Self-evaluation of learning	8	–
	Information, facts	7	–
Assignments (essays, reports, case studies)	Research, analysis, synthesis	12	–
	Problem-solving, design, practical concepts	6	–
	Communication skills	4	20
	General or integrated understanding	3	20

*Only representing responses to questions that were used in common between Flinders University and La Trobe University.

Open text responses showed that staff at La Trobe University and Flinders University were confident that examinations could measure more than recollection of facts although they were certainly aware that students can (and do) prepare for examinations by rote learning of material.

> If the exam questions are good, then students can show in their exam what they have learned in the way of concepts and ideas and importantly can show how these can be connected to real world experiences.

> [Exams] can evaluate a student's ability to synthesise information, recall, recognition and sometimes novel application.

However there is a:

> danger that "parrot" learning is rewarded, especially in the first year where basic concepts are taught.

and that:

> unless carefully worded can test memory rather than understanding.

Staff also pointed out that "a good assessment tool takes a lot of time to develop" and acknowledged that it is "not so easy to assess/encourage graduate attributes" particularly when students are stressed and do not present their best work. This is represented by the following comments:

> Another disadvantage is that very rarely in the real world do people have to show their knowledge under exam-type conditions.
>
> Students cannot access a range of information sources (e.g. internet and colleagues) which you would in a real situation.

Effectiveness of modes of assessment

At La Trobe University educators were asked to rate the effectiveness of different modes of assessment on a Likert scale: 1 – 'little match with student learning', 2 – 'some match with student learning', 3 – 'good match with student learning', and 4 – 'excellent match with student learning' (Table 6).

Table 6: Summary of staff rating of the effectiveness of measuring learning at La Trobe University

Assessment mode	Effectiveness of assessment (weighted mean)	
	Knowledge/content	Skills
Formal written examinations	3.5	3.3
Assignments (essays, reports, case studies)	3.4	3.5
Prac reports	2.9	3.0
Self-test quizzes	2.6	2.3
Multiple choice examinations	2.4	1.7

Interestingly, despite their reservations about what learning outcomes can be measured in examinations, educators believe that examinations are as effective as assignments or written work in assessing both skills and knowledge for their students. Multiple choice examinations were seen to have the least value in measuring learning outcomes with scores indicating some or little match with student learning. Open text responses to survey questions reveal that staff at both La Trobe University and Flinders University place considerable value on those skills that they believe can be measured by examinations, such as problem-solving or presenting an integrated view of the subject matter.

> The student needs to study all the unit material in an integrated fashion, thus facilitating the linking of concepts and a deeper understanding of the unit material.
>
> Some things the students actually have to sit down and learn, our lecturer used to say "Neurox it don't xerox it". There's a lot of truth in actually making sure you understand something enough so that you can recall it where and when you need it.

> [The] examiner can see that the students actually give evidence of having understood some of the material.

Educators recognise this is only true of well-constructed examinations and that the conditions of the examination are limiting.

> Many exam questions are very boring and encourage students to regurgitate facts.

> The learning of material in this unit is much less dependent on rote learning. The types of questions addressed in assignments often require a long time for reflection and thought and much more than the 3hrs available in a formal written exam. A 3hr exam would necessitate rather trivial computational questions.

> Students with good memories can "cram" and forget all information after the examination. That is why projects, reports and assignments have a better chance of "deep" learning than can be tested in a formal examination and is difficult to "cram".

> Some students get stage fright/mental blocks etc.

There was also some indication that there is inertia associated with past practice and efficiency.

> Many years of experience with this form of assessment (i.e. most of the existing academics in the world have been assessed in this way) suggests that this has been a successful method to teach information that has 'stuck' in their minds and informed further learning.

> Examinations have the advantage of scale [as] many people can be tested at once.

Why choose examinations?

The major reason for the selection of formal examinations for assessment was to minimise cheating and plagiarism. Educators believe that formal examinations are the most effective way to ensure that assessed work is completed independently by each student. This was particularly emphasised for those subjects that rely on solving calculated problems (mathematics, statistics) where it is more difficult to detect and manage plagiarism. Table 7 lists the features of examinations that staff find important and is drawn from the open text responses from surveys at both universities.

Table 7: Characteristics of formal written examinations reported by science staff at La Trobe University and Flinders University

	No. of comments		
	La Trobe (Faculty)	Flinders (Faculty)	Flinders (Biology)
Advantages of formal written examinations			
Independent work (no cheating or plagiarism), fair	19	9	16
Measure knowledge and understanding	9	5	11
Encourage study	8	2	18
Require integrated knowledge	8	2	
Easy management and organisation, efficient	6	3	13
Can measure other skills	5	1	
Disadvantages of formal written examinations			
Students are stressed and/or lack exam skills	14	4	
Can encourage shallow learning (memorisation)	11	9	13
Constrained by time	10	1	
Poor assessment of graduate attributes	8		
Staff workload in marking is high	5		12
Unrealistic situation		5	
No feedback for student learning		3	

Educators were also asked to comment on external factors that restrict their selection of assessment tasks (Table 8). By far the largest factor limiting the choice of assessment for undergraduate subjects is staff workload. The only assessment method that staff nominated as providing a significant saving in staff workload was the multiple choice examination. This was virtually seen as a requirement for very large classes.

Table 8: External constraints on choice of assessment task by La Trobe University and Flinders University educators

Limitations to use of best forms of assessment	No. of comments La Trobe	No. of comments Flinders
Staff workload	15	6
University's assessment regulations	8	1
Resources for teaching	5	4
Capacity for plagiarism/cheating	4	4

The resource limitations on the practice of teaching in Australian universities have been widely recognised (Bradley, Noonan, Nugent, & Scales, 2008). The ratio of staff to students has decreased at the same time as the sector has been required to demonstrate improved learning outcomes for students. The sector is also being encouraged to expand diversity in the student cohort and to accommodate students with more varied preparation for academic study. These pressures will undoubtedly restrict selection of assessment and particularly of feedback mechanisms for students.

Conclusions

Science educators at La Trobe University and Flinders University use formal written examinations as the major form of assessment. This agrees with earlier analyses of the effect of the science disciplinary context (Ahlberg, 2008). Staff at both universities are strongly biased toward the use of examinations to limit student opportunities for plagiarism. This is at least partly a consequence of the large scale of subjects (particularly at first year) where it becomes impossible for lecturing staff to be familiar with the work of individual students. It is also a product of the pressure on staff to competitively rate students.

Open text responses in surveys from educators at both universities showed that staff felt that examinations produced appropriate outcomes for many undergraduate courses, yet most staff identified that they would like to reduce the weightings placed on examinations. The reported learning outcomes are substantively focussed on the content of the topic area and the ability to solve problems. It is likely that problem-solving in this context refers most often to a calculated solution to a theoretical or applied problem since most of these comments applied to responses from staff in mathematics, physics, statistics, engineering or computer science. Some staff recognised limitations to the outcomes that could be measured by examinations but a number of staff commented that this could be ameliorated by construction of "good" examinations. Others acknowledge the difficulty in constructing "good" examination questions. Staff also felt that decreasing the weighting of exams as students proceed toward graduation was a good idea, essentially to prepare students for science careers. In the case of experimental sciences, more effort should be put toward practical work later in a degree program.

Development of assessment practices appears to be limited by local pressures, primarily staff workload, such that staff can be reluctant to invest in the development of alternative assessments. Academics from IRUs in particular have been found to spend more than twice the number of hours in face to face teaching than are academics from Group of Eight Universities (Burke da Silva *et al.,* 2008) further stressing the problems for this group with respect to workload. It is also interesting that educators offered little independent evidence to support their views although information on awareness of educational research was not specifically sought. This may in part be connected to staff not being aware of the strengths and validity of other forms of assessment and the language of higher education and finding it difficult to invest the time to explore this field (Burke da Silva, *et al.*, 2008). Investment in trials of alternative assessment tasks could increase confidence that learning outcomes can be measured effectively by alternatives to written examinations or by better design of examination questions and that organisational hurdles can be overcome. This may also assist staff to further develop their ideas about what effective assessment is and how it can be developed for local conditions.

Acknowledgements

The authors thank Mr Tony Gleeson for his assistance in gathering data on assessment practice at La Trobe University and with construction of survey questions.

References

Ahlberg, A. (2008). Teaching and learning in hard science research environments: Views of academics and educational developers. *Higher Education Research and Development, 27*(2), 10.

Biggs, J., & Tang, C. (1999). *Teaching for quality learning at university*. Philadelphia: Open University Press.

Bradley, D., Noonan, P., Nugent, H., & Scales, B. (2009). *Review of Australian higher education: final report*. Canberra: DEEWR.

Burke da Silva, K., Buckley, P., Roberts, M., Dent, L., Wood, D., & Gannaway, D. (2008). *Raising the profile of teaching and learning: Scientists leading scientists*. Sydney: Australian Learning and Teaching Council.

Gibbs, G., & Simpson, C. (2004–5). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3–31.

Gooding, J., & Shapter, J. (2008). Teaching undergraduates nanotechnology. In A. E. Sweeney & S. Seal (Eds.), *Nanoscale science and engineering education* (pp. 421–457). Stevenson Ranch, California: American Scientific Publishers.

Harper, R. (2003). Multiple choice questions – a reprieve. *Bioscience Education 2*, 2–6.

Prosser, M., & Trigwell, K. (1999). *Understanding learning and teaching: The experience in higher education*. Philadelphia: Open University Press.

Quality Assurance Agency (QAA). (2003). *Learning from subject review 1993–2001: Sharing good practice.* Gloucester: Quality Assurance Agency for Higher Education.

Ramsden, P. (1997). The context of learning in academic departments. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning* (pp. 198–216). Edinburgh: Scottish Academic Press.

Scouller, K. (1997). Students' perceptions of three assessment methods: Assignment essay, multiple choice question examination, short answer examination. *Research and Development in Higher Education, 20*, 646–653.

Shapter, J., Ford, M., Waclawik, E., & Maddox, LM. (2002). Teaching undergraduates nanotechnology. *International Journal of Engineering Education 18*, 512–518.

van Driel, J., Beijaard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching, 38*(2), 137–158.

Williams, J. (2006). Assertion-reason multiple-choice testing as a tool for deep learning: A qualitative analysis. *Assessment & Evaluation in Higher Education, 31*(3), 287–301.