

# Expectations of student learning quality

## An introductory study

**Dennis Bryant**

University of Canberra

Without a direct measure of learning, universities and lecturers do not have reliable evidence of changes, past or present, in the academic merit of a unit. By using grade data to develop a variable called Academic Merit, all university units were measured for their percentages of academic merit over one semester at an Australian teaching-intensive regional university. Although the results revealed units with excellent percentages, there were others with percentages that were other than excellent. The implication is that an opportunity exists to understand the quality of the learning merit in those units with a view to enhancing student academic learning.

### Introduction

Discussions in higher education that uncritically link teaching and learning into a co-joined mass are simplistic in their acceptance of the assumption that educational improvements in one co-joined principal, stereotypically, the teaching principal, imply that those improvements affect equally the other co-joined principal, here, learning. In the co-joined teaching and learning model, quality assurance efforts that raise the quality of teaching by one point would be expected to raise the quality of learning by a similar margin of one point. Curiously, there is a dearth of published literature that empirically supports the teaching and learning co-joined model's assumption of equivalent co-growth. However, there are claims, not merely against an assumed equivalence of teaching improvements to learning result:gain ratios, but more dramatically against the nature of some co-joined educational models. Biggs (2001) presents a number of impediments to the quality feasibility of co-joined models, arguing that educational success requires the alignment of not two but three principals, with the third principal being assessment. The Biggs constructive

alignment model promises that adherence to teaching and assessment will result in student learning quality excellence; however, in both models the missing ingredient is how best to detect student learning quality excellence (abbreviated hereafter as merit), given that students start from different bases.

A commonly used method of detecting learning merit is to rely on surveys that measure a surrogate variable in place of directly measuring learning merit. In survey devices (Ramsden, 1991), students are invited to reflect on their satisfaction with teaching, from which results are extracted student perceptions of learning. But this approach relies on simplistic co-joined model assumptions and has critics (Denson *et al.*, 2010; Edstrom, 2008; Shevlin *et al.*, 2000).

An alternative approach is to eschew surrogate measures in favour of direct measures. Using a non-surrogate approach with an emphasis on empirical learning results, Bryant (2013a) proposed a variable called 'academic merit', whose values were 'no merit shown' and 'merit shown'. A no merit shown value refers to students who achieved either a failure grade or a P (pass) grade. While it is intuitive that a failure grade be interpreted

as no merit shown, a pass grade is adjudged as no merit shown because it is an expectation that students have a reasonable chance of achieving a pass in units (MacKie, 2001). Completing the picture, a value of merit shown is adjudged to contain credit, distinction and high distinction grades.

With a measurement approach established, discussion turns to expectations, particularly university educators' and university leaders' expectations about the learning quality evident in student grades, which occur in each unit's learning spaces. Ideally, expectations could be set at 100 per cent of a unit's students achieving merit. While this expectation is laudable, there are reasons to doubt that educators and university leaders are likely to be setting their goals at ideal levels. Biggs (2001, p. 236) finds that 'there are factors in the institutional climate or structures that are deleterious to learning'. Specifically, Biggs (2001) registers a number of points and these include concerns about experience levels of validation panels that approve courses, student feedback questionnaires and the measurement of charisma instead of teaching effectiveness, which is a theme also developed elsewhere (Edstrom, 2008), the maintenance of quantitative mindsets that seek to grade on the curve, thereby ranking student learning results against desired norms rather than having a focus on attainment, and, as a final example, Biggs (2001) claims that some educators have adopted a 'Han Dynasty in 4th Century BC China' elitism that seeks to be selecting the real scholars at the expense of perceived student non-scholars.

Accepting the validity of these impediments to the ideal, it follows that expectations would need to be set lower. In this regard, the question becomes: What is the lowest merit point that needs to be achieved? If it is an axiom that students, selected for university courses through entry requirement tests have a better than fair chance of passing courses (MacKie, 2001), then the lowest merit point might be near a half way average point, that is, at least 51 per cent of student grade distributions should be in the merit range. But if this were to be accepted as the merit goal point, it might be considered to be disappointingly low since it would give acceptance to units with 49 per cent student non-merit rates. At first glance, 49 per cent non-merit might seem reasonable because it implies a 24.5 per cent P rate; but it implies also a 24.5 per cent failure rate, on average. Accepting a one student in four failure rate does not evoke images of universities and students cooperatively reaching the heights of learning excellence. If the rhetoric of the *Bradley Review* (Bradley, Noonan, Nugent & Scales, 2008) is any guide to the

quality of learning required in universities, then the bar would need to be raised. Setting expectations at 60 per cent of students attaining merit grades might be a workable solution. Although a 60 per cent rate might seem a mediocre goal, it encompasses expectations of a lower 20 per cent failure rate, together with a 20 per cent P rate, on average.

In summary, while rates of 100 per cent merit are likely to be idealistic and unachievable in some university units, an alternative benchmark point exists at 60 per cent. This benchmark is interpreted to mean that to be considered mediocre as regards attained learning merit, a unit's learning results need to reach this point. Beyond this point, merit might be referred to as good, and beyond good as excellent.

When it comes to the number of students enrolled in a unit (class size), the Biggs constructive alignment model remains a unified model, reiterating the value of assessment alignment at all times with educational objectives. On the same topic of class size, the co-joined model splits into two variants, the second variant dealing with perceived operational problems with large class sizes (AUTC, 2003). In Version 2 of the co-joined model, assessment options are often restricted to formal timed, usually multiple choice examinations and, on occasions, can comprise 100 per cent of the assessment tasks. Without delving into the shortcomings or discussing the justifications for this assessment approach, it is sufficient to note the literature's concern with large classes. Due to this concern, some effort will be expended to differentiate large classes' merit from small and medium classes' merit.

## Research questions

1. Is a failure of quality assurance of learning, evidenced in units with low merit, visible across universities, and are there any indicators of its degree of prevalence?
2. Is a failure of quality assurance of learning evident only in large classes?

## Methodology

### Data

For this analysis, an available but secondary dataset, containing unit learning results for all 9,000 students enrolled in all units in Semester 1 of 2009 and falling within the confines of an existing ethics agreement, was used. Units that assessed students using competency-based assessment were removed, since such assessment does not

**Table 1. Merit distribution for Semester 1 of 2009 units**

Rating	Units	Students with merit %	Learning
1	8	Up to 20	Pitiful
2	47	Up to 40	Dismal
3	57	Up to 50	Disappointing
4	60	Up to 60	Mediocre
No merit total	172		
5	178	Up to 80	Good
6	199	Up to 100	Excellent
Merit total	377		

recognise merit (Booth, 2000; Ennis, 2008; Gonczi, 1994; Guthrie, 2009). This left 549 traditionally assessed units or 26,891 rows of unit data. A few words concerning the unavailability of competency-based assessment data are appropriate at this point. The dataset did not include a unit type attribute that would have described a unit's assessment as either competency based or traditionally based assessment. Additionally, for Semester 1 of 2009, the university did not employ a central repository for paper or electronic copies of unit outline documents. This absence of an attribute and the then non-availability of outlines, meant that competency-based assessment units had to be detected from grades such as 'ungraded pass' (UP in the grade attribute), and from the non-presence of merit grades in a unit's reported learning results. Although not likely, it must be acknowledged that some competency-based assessment units could have become unintentionally included, with traditionally based units.

### Process

The data were loaded into database tables (Microsoft Access, 2007) which enabled the construction of SQL queries to operationalise academic merit. This was done using two attributes, called 'merit shown' and 'no merit shown'. For each student record that indicated merit, the merit shown count was incremented; alternatively, the no merit shown count was incremented. The total count for merit shown for each unit was converted into a percentage of records for that unit. In the final step, and needing the addition of a range attribute, merit shown was segmented into six percentage ranges, as shown below, the six ranges increasing from pitiful merit through dismal, disappointing, mediocre and good merit to excellent merit, and then populated using the earlier calculated percentage counts, which yielded Table 1.

As regards class size in the dataset, descriptive statistics showed the mean to be 49, the median to be 24 and

the mode to be two students (indicating the incidence of small units). But this distribution does not sit comfortably with the university's definition of small at 20, medium at 100 and large above 100 students. A better fit was achieved when quartiles were used. When the quartiles were calculated, the first quartile cut-off point for small was 10; with the second and third quartile representing medium classes at 24 and 58.5 respectively; and the fourth quartile of 423 representing large classes. When mapped, 139 units were small, 272 were medium and 138 were large sized. In terms of statistics, this distribution parallels a bell shaped curve, with tapered left and right ends.

### Discussion

A total of six merit ranges are presented in Table 1 in the Students with Merit column, beginning with an Up to 20 per cent range and flowing through to Up to 100 per cent range. It would have been possible to show ten ranges, that is, at every 10 per cent cut-off point, but fewer range points convey satisfactorily a clear picture of merit distribution.

As regards interpretation, an Up to 20 per cent range means that there might have been a complete absence of students within a unit who showed merit through to a maximum of 20 per cent of the unit's students who showed merit. Alternatively stated, the Up to 20 per cent range means a minimum of four of every five students in a unit were adjudged as unable to demonstrate learning merit against the unit's learning assessment criteria. This level of learning has been labelled as pitiful. While a label of pitiful might be seen to be emotive, it does serve to direct attention to poor learning performance, which is an issue often avoided in higher education literature, although poor performance, and especially failing performance, is 'costly for both individuals and universities' (McInnis, 2001, p. 106) and is easily underestimated since often they are revealed as discontinuation or failure only in later years (Pargetter *et al.*, 1998).

In summary, Table 1 shows that every third unit was below learning merit expectations. This means a total of 172 of 549 units had sub expectation merit levels, and the 172 total can be broken down into eight units wherein student learning merit was pitiful, 47 wherein student learning merit was dismal, 57 wherein student learning merit was disappointing and 60 units wherein student learning merit was mediocre. The good news is that within two of every three units, student learning merit was good or excellent. While this is a solid starting point, it does not deny that quality assurance efforts are needed with learning underperformance in one-third of units.

**Table 2. Units where merit distribution was pitiful**

Unit	Merit students	Non-merit students	Total Students	Merit %
1	0	1	1	0
2	0	1	1	0
3	0	1	1	0
4	0	2	2	0
5	0	5	5	0
6	9	59	68	13
7	17	74	91	19
8	25	130	155	16

**Dismal unit learning merit case study**

The eight units with dismal learning merit are now presented in more detail, via two tables.

Table 2 shows that Unit 1 to Unit 3 had only one student each, effectively making the teaching a full semester of personalised tuition. In spite of the teacher:learner ratio favouring success, the learning results were pitiful, with all three units registering zero merit levels, when it would be reasonable to expect that singularly personalised tuition levels would produce merit results. Even if the ratio were to be increased to become one teacher:two students or one:five students, it would not be unreasonable to hope for merit results, but Table 2 results show two such units in which there was a complete absence of student learning merit. Those units are Units 4 and 5.

A common explanation given anecdotally by teaching staff for low learning merit revolves around students being the culprits for their own downfall. This anecdotal explanation of assumed student under commitment is known as the student deficit discourse (Lawrence, 2002) but it has not been shown to be reliable. When underperforming Unit 5 was investigated further, it was found that there were four first year students of whom three had attained a P grade, with one attaining a fail grade due to not completing all the assessment. These numbers might be taken to represent a picture of new students

struggling with the university discourse and being unable to produce any merit. While this picture of new student total inability seems dubious, it cannot be dismissed as impossible because the small number of students might have been unrepresentative of the overall student population. However, the fifth student's learning journey details suggest strongly that there are institutional impediments blocking student learning in the unit, affirming Biggs' (2001) findings in a small way. Coming into Unit 5 as a continuing student, the fifth student had a grade point average score of seven points, which can be attained by achieving a high distinction in all previously taken units. In this unit, the fifth student attained a non-merit P grade. It must be admitted that this fall is a long drop from highest merit ranking to non-merit attainment. Taken together, the students' grade results suggest the presence of institutional impediments to student learning merit that could be investigated in order to assure quality. The presence of impediments to learning merit, resulting in unit merit underperformance, should not be considered necessarily as a chance event. Unlike the wide ranging claims made by Biggs (2001), no doubt predicated soundly on wide higher education experience, Bryant (2013b) has documented situations occurring at the unit level where learning was negatively influenced by institutional factors. There are three as yet undiscussed pitiful learning units, Unit 6 to Unit 8, for which grade details and some description of year level and course are provided below.

Unit 6 is a postgraduate management unit in the business stream; Unit 7 is a third year accounting unit, again in the business stream; and Unit 8 is a first year administration unit in the sports science stream. Although the three units in Table 3 share a common feature which the count of non-merit students dramatically exceeds the count of merit students, they share a further feature: no students have been awarded a high distinction grade. Unit 6, for example, awarded distinction and credit grades but no high distinction grades. This action is difficult to understand because the number of students in each unit is significantly sized to statistically imply that there should be students awarded a high distinction merit grade (Moore,

**Table 3. Final three units where merit distribution was pitiful**

Unit	Unit	HD	DI	CR	Merit students	P	Fail	Non-merit students	Total students
6	Management	0	3	6	9	37	22	59	68
7	Accounting	0	1	16	17	32	42	74	91
8	Administration	0	0	25	25	118	12	130	155

2007). Furthermore, Unit 7 awarded few distinction grades, and Unit 8 awarded no distinction grades. These figures would seem to indicate what Biggs (2001, p. 235) refers to as 'distorted priorities' in the alignment of educational objectives.

**Regression to non-meritorious assessment**

While it can be argued that the source of misalignment rests with the institution, and while it has implications for student entry procedures, for curriculum design, for teaching, for assessment (Biggs, 2001) or for all of these factors, there is a worrying pattern in Table 3's figures, which could be an indication of a growing trend. Is university merit assessment regressing to competency-based assessment, in which there are no indicators of merit? Units 6 and 7 use the lower two of the three merit indicators minimally, whereas Unit 8 regresses to the lowest possible merit grade as the lone indicator of merit. This situation is reminiscent surely of competency-based assessment. The grades for Unit 8 are ironic because it is the largest of the three units and in statistical distribution terms, should best reflect a student population (Moore, 2007). Additionally, it is easier to attain a fail grade in Unit 7 than to secure a P grade, with failures (42 of 74 non-merit grades) exceeding P grades (34 of 74 non-merit grades). Against these numbers, it is difficult to conceive the beneficial learning opportunity being given to students.

**Answer to research question 1**

Research question 1 asked: Is a failure of quality assurance of learning, evidenced by units with low merit, visible across universities, and are there any indicators of its degree of prevalence? This paper's analysis of one semester's learning results suggest there is a failure of quality assurance, and the failure in learning is happening in up to one-third of the university's units. To the extent to which

this university's attributes are shared by other universities, it is proper to conclude that these findings might be relevant to merit levels in other universities. The following section addresses the second research question, which regards class sizes.

**Discriminating by class size**

Recapping, for size analysis, classes were divided into one of three groups (small, medium and large) based on enrolment numbers, which allowed for different sized groups to be compared by using the six learning merit ranges of dismal, representing poor merit, through to excellent where merit expectations were exceeded. Table 4 shows a pattern of rising class sizes being paralleled by a fall in learning merit. Units with small class sizes, for example, achieved 82 per cent merit levels, but this fell to 71 per cent for medium class sizes and, in a sombre continuation of the falling trend, to 51 per cent for units with large class sizes.

An economic interpretation of the percentages shown in Table 4 might conclude that this university, as a generalisation of all universities, is running a two speed learning economy. In this metaphor, small and medium sized classes are productive centres of learning, with about three out of four units achieving merit. By contrast, large sized classes produce merit in only two out of four units. Furthermore, when the quality of large sized classes is internally examined for learning, merit rated as good (43 per cent) is found five times more frequently than merit rated as excellent (8 per cent), and might reflect the operation of Han Dynasty agendas (Biggs, 2001). While the ratio of good:excellent learning merit is close to par in medium sized units (37 per cent compared to 34 per cent), the ratio in small sized units is the inverse of large sized units, at almost five:one (14 per cent to 68 per cent) in favour of merit excellence.

**Table 4. Units by Class Size and Merit**

Rating	Learning	Small	Small %	Medium	Medium %	Large	Large %
1	Pitiful	5	4	0	0	3	2
2	Dismal	9	6	20	7	18	13
3	Disappointing	9	6	25	9	23	17
4	Mediocre	2	1	34	13	24	17
No merit	Total	25	18	79	29	68	49
5	Good	19	14	100	37	59	43
6	Excellent	95	68	93	34	11	8
Merit shown	Total	114	82	193	71	70	51

### Answer to research question 2

Research question 2 asked: Is a failure of quality assurance of learning evident only in large classes? This paper's analysis of one semester's learning results, done empirically using lecturer-assigned grade data in preference to collection of student perceptions, suggests that there is a failure of quality assurance across units in all three size ratings, but this failure of quality appears to be dire within large sized units and is seen in low merit attainment by large units.

For intellectual universities, large sized below par unit learning must become a point of contention and resolution. Some centuries ago, Plutarch admonished his colleagues by claiming it was an intellectual duty to fire, not fill, student minds. The admonition might be taken to be an exhortation to produce learning merit, ideally at excellence levels, in students. While there is scope for improved learning rates in small and medium sized units, there must be greater scope to uncover the reasons for low merit performance in large sized units.

While it might have been fashionable to accept the culpability of students in all matters of failure, there was never convincing empirical proof (Lawrence, 2002). In fact, there is research that supports a case of institutional culpability, especially in its short sighted avoidance of understanding assessment's impact on learning. The Biggs (2001) alignment model is an argument detailing institutional, not student, blindness. Additionally, Brennan *et al.* (2008) argue for improvements in university excellence. Finally, and apart from this paper's findings on units with below expectations learning merit, there is empirical research that reveals areas of institutional underperformance in understanding learning, including the role of formal withdrawals impacting learning, student strategies in determining unit workloads after unit failure, students who enrol but do not participate in unit attendance and assessment, the presumed quality of teaching inputs, and the fairness of procedures that exclude students from their learning journeys even if the student failures were in units with dismal learning levels (Bryant, 2013a, 2013b; Bryant & Lyons, 2013; Bryant & Richardson, 2010; Bryant *et al.*, 2013). These concerns are indicators of where and how research can be employed to enhance institutional direction of student merit attainment.

### Conclusion

The thesis was that universities and lecturers are restricted to surrogate measures of measuring learning, for example, employing surveys concerning student per-

ceptions of the quality of teaching, and preclude themselves from the benefits of a direct method for measuring academic learning merit for their units. The implication was proven that, without a direct measure of unit merit, universities and lecturers are unlikely to have reliable evidence of merit excellence or merit non-excellence, past or present, in academic units. An empirical methodology was employed at a teaching intensive university, where all units were measured over one semester for their percentages of merit shown which was the operationalisation of unit grade data to measure academic merit.

Even though the results showed units with good or excellent percentages, there were other units that were placed in the non-merit portion of the scale, being ranked as pitiful or dismal or disappointing because of their low merit percentages. When units were ranked by class size, that is, by the number of student enrolments such as small to medium to large, a non-merit pattern was seen to be prevalent in large units. This pattern does not imply student underperformance and hence student culpability. Rather, it implies factors that are outside the control of students and hence points at institutional responsibility, which has been discussed in higher education literature in alignment models by Biggs and inevitably might concern assessment choices made prior to student arrivals into units. The implication is that an opportunity exists to investigate the quality of non-merit units with a view to enhancing their academic merit. This could be an empowering step towards firing student minds, thereby enhancing student learning, not to mention strengthening student retention with resulting reductions in student recruitment replacement costs for universities, as well as more effectively progressing students on their degree journeys.

**Dennis Bryant has an abiding interest in student learning failure and thrives on designing non-surrogate metrics to measure student learning.**

### References

- AUTC. (2003). *Teaching Large Classes Project 2001 Final Report*. Brisbane: The University of Queensland.
- Biggs, J. (2001). The reflective institution: assuring and enhancing the quality of teaching and learning. *Higher Education*, 41(3), 221–238.
- Booth, R. (2000). *Competency based assessment – One minute wonder or here to stay? Practitioners' attitudes to CBA & complexities of implementation*. Retrieved from [http://www.avetra.org.au/abstracts\\_and\\_papers\\_2000/rb\\_full.pdf](http://www.avetra.org.au/abstracts_and_papers_2000/rb_full.pdf).
- Bradley, D., Noonan, P., Nugent, H. & Scales, W. (2008). *Review of Australian Higher Education*. Retrieved from [http://www.deewr.gov.au/he\\_review\\_final-report](http://www.deewr.gov.au/he_review_final-report).



- Brennan, J., Enders, J., Musselin, C., Teichler, U. & Valimaa, J. (2008). *Higher Education Looking Forward an Agenda for Future Research*. Strasbourg: European Science Foundation.
- Bryant, D. (2013a). Retaining students who 'did not participate': two case studies and lessons for universities. Paper presented at the 20th The Learner International Conference on Learning, Rhodes: University of the Aegean.
- Bryant, D. (2013b). Students and their learning journeys: a study of failure in an Australian university. Unpublished PhD thesis, University of Canberra, Canberra.
- Bryant, D. & Lyons, K. (2013). Academic retention or exclusion: student failing learning journeys. Paper presented at the 20th The Learner International Conference on Learning, University of the Aegean, Rhodes.
- Bryant, D. & Richardson, A. (2010). Rising unit withdrawals can signal student concerns with learning results: an exploratory study. Unpublished paper.
- Bryant, D., Richardson, A. & Lyons, K. (2013). Student academic learning failure and workload strategies: what does it mean for universities? Paper presented at the 20th The Learner International Conference on Learning, University of the Aegean, Rhodes.
- Denson, N., Loveday, T. & Dalton, H. (2010). Student evaluation of courses: what predicts satisfaction? *Higher Education Research & Development*, 29(4), 339–356.
- Edstrom, K. (2008). Doing course evaluation as if learning matters most. *Higher Education Research & Development*, 27(2), 95–106.
- Ennis, M. (2008). Competency models: A review of the literature and the role of the Employment & Training Administration. 1–25. Retrieved from [www.careeronestop.org/competencymodel/info.../OPDRLiteratureReview.pdf](http://www.careeronestop.org/competencymodel/info.../OPDRLiteratureReview.pdf).
- Gonczy, A. (1994). Competency-based assessment in the professions in Australia. *Assessment in Education*, 1(1), 27–44.
- Guthrie, H. (2009). Competence and competency-based training: what the literature says. Occasional paper by The National Centre for Vocational Education Research 1–32. Retrieved from <http://www.ncver.edu.au/publications/2153.html>.
- Lawrence, J. (2002). The 'deficit-discourse' shift: university teachers and their role in helping first year students persevere and succeed in the new university culture. Paper presented at the 6th Pacific Rim FYHE conference. Retrieved from <http://ultibase.rmit.edu.au/Articles/march03/lawrence1.htm>.
- MacKie, S. (2001). Jumping the hurdles – undergraduate student withdrawal behaviour. *Innovations in Education and Teaching International*, 38(3), 265–276.
- McInnis, C. (2001). Researching the first year experience: where to from here? *Higher Education Research & Development*, 20(2), 105–114.
- Moore, D. (2007). *The Basic Practice of Statistics* (4th ed.). New York: W. H. Freeman and Company.
- Pargetter, R., McInnis, C., James, R., Evans, M., Peel, M. & Dobson, I. (1998). *Transition from Secondary to Tertiary: A Performance Study*. Canberra: DETYA.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: the course experience questionnaire. *Studies in Higher Education*, 16(2), 129–151.
- Shevlin, M., Banyard, P., Davies, M. & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: love me, love my lectures? *Assessment and Evaluation in Higher Education*, 25(4), 397–405.