

Article Title: What universities might learn from professional associations in marking large undergraduate modules

Abstract

Purpose: The purpose of this viewpoint paper is to highlight to universities the approach of professional associations in marking large volumes of assessments.

Design: The issues arising in marking large undergraduate modules in universities are discussed, before describing the approach typically adopted by professional associations. The benefit for universities of adopting such an approach is then examined.

Findings: The key to marking large volumes for professional associations lies in the initial standardisation meeting, on-going discussions between markers, and an ability to check and if necessary stop the marking process at several points in the marking cycle, until agreement and consistency between markers is achieved.

Research limitations: This viewpoint paper is based on the experiences of the two authors in assessing for professional business management associations, and not empirical research.

Practical implications: The approach described here to marking large volumes of undergraduate assessments offers a practical solution to the prevailing issues concerning consistency across teaching teams.

Originality/Value: The application of the approach of professional associations to the issues surrounding marking large undergraduate modules offers a unique proposition for university staff to consider.

Keywords: Assessment, Higher Education, Professional Associations,

Article Classification: Assessment Policy Article

Paper Type: Viewpoint

What universities might learn from professional associations in marking large undergraduate modules

Mass Higher Education

While the growth in students applying for degree programmes at UK universities has recently slowed (Havergal, 2015), the numbers registered to undergraduate level studies in the UK have increased from 1541225 in 2000/01 to 1747855 in 2015/16 (<https://www.hesa.ac.uk/data-and-analysis/students>). This growth in higher education (Mayhew et al., 2004); Scott, 2005; Smith, 2006) has resulted in some large class sizes (Bandiera, 2010) especially in first year undergraduate programmes, although there is some evidence to suggest this scenario is marginally improving (Guardian, 2015).

Current issues with marking assessments

Nonetheless, not only do large class sizes create problems for learning and teaching (Mulryan-Kyne, 2010), issues surrounding assessments especially quality of feedback (Nicol, 2010) and ensuring reliability and consistency in marking (O'Hagan and Wigglesworth, 2015) remain perennial concerns. It would be logical to assume that the progression to mass higher education would result in greater standardisation of knowledge, programmes and assessments (Teschler, 2001) which would resolve any marking discrepancies. However, in modules with large student cohorts, assessments tend to be marked by a team of Tutors, which can result in problems of consistency across all of the members. As Dunn et al. (2005) note, even with criterion-referenced assessment assessors can differ in their respective interpretations, particularly in essays or similar discursive pieces (Brooks, 2012). This

results in assessors tending to mark work in a holistic manner rather than in a precise, analytical method (Bloxham et al., 2011; Kalthoff, 2013; Sadler, 2005) which can thus create discrepancies across marking teams.

These discrepancies can be exacerbated by using sessional staff or graduate assistants who might have little assessment experience (Bearman et al., 2016) or may have received little training (Smith and Coombe, 2006). They can also potentially adopt risk-averse practices due to unequal power relations (Bloxham, 2009; Grainger *et al.*, 2015; Salamonson et al., 2010) manipulating marks to what they think would appease those in senior positions. Even though marking schemes may be tightly enumerated to remove any personal interpretation on behalf of the respective marking tutor, marks can still be adjusted, in accordance with perceptions, rather than the established assessment criteria. Additionally, there is even evidence to suggest that individual marking consistency will change as an assessor works through a large batch of assessments (Dracup, 1997; Pinot de Moira, 2002).

While marked work is usually subject to moderation, some authors have been critical of this practice (see for example, Light et al., 2014) especially regarding the extent of the subjective nature of moderating (Smith, 2012). Moreover, moderation at the end of a module is somewhat akin to a quality control process, in that it is an end-point check - the substantial work of marking has already been undertaken. If it is then found to be erroneous, then a process of re-working (re-marking) has to be carried out or an adjustment made to all students' work to correct the already given marks. What is therefore required is a more robust system of quality assurance with effective communication between team members to examine issues when and where they occur, to ensure consistency of quality is maintained throughout the marking process. This can correspondingly, raise confidence levels of students and staff in the marking process.

While universities therefore wrestle with the issue of consistency of marking across large student modules, professional associations with an awarding function typically have a long history of assessing large volumes of work, and maintaining standards appropriate to the professional discipline and for external regulation, such as meeting RQF requirements (see: <https://www.gov.uk/government/organisations/ofqual>)

The suggestion in this viewpoint paper is for university colleagues to consider adopting the work practices of those professional associations in marking large student modules. This is based on our

personal involvement of assessing for several professional business management bodies during our full-time employment in the higher education sector.

A possible solution to the marking dilemma? The professional association approach

We have found that the typical marking process for a professional body comprises several stages with checks at each stage, and a facility to stop marking should any discrepancies be noted. The assessment and marking outline is usually prepared by a subject examiner. Once an assessment for a module/unit has been completed by students, a sample of the work is copied and together with the outline mark scheme, distributed to the team of markers for marking. The marks awarded for each element or question of the assessment by the markers are then discussed by the markers in a standardisation meeting and agreement to a final mark and the final marking scheme to be universally applied by all members of the team is concurred.

All markers commence marking students' work and once an agreed number of pieces has been completed early in the marking cycle, for example twenty pieces or 10%, a sample of marked work is sent to the subject examiner for moderation. The subject examiner moderates marked pieces across the team and either gives approval for each respective marker to continue marking or communication is engaged to discuss marking differences, before continuation is permitted. In either case, feedback on the marking of each element is provided back to the marker. Similarly, once between half and three-quarters of a marker's allocated batch has been marked, a further sample is forwarded to the subject examiner for moderation. Again, if the feedback is positive, then continuation of marking to completion is permitted otherwise discussion and agreed actions regarding the discrepancies ensues. Any issues arising at any stage in the assessment or marking process are communicated within the marking team, with the facility to engage in discussion between marking team members if needed. The subject examiner's work is moderated by an internal verifier or programme lead, to ensure consistency of marking across all modules at a particular level of qualification.

Discussion of the professional association approach

While practices will undoubtedly differ across higher education institutions, programmes and teaching teams, the professional association approach denotes a marked difference to our experience of working in several, but not all, higher education providers where the process is

seemingly driven by the need for a speedy turnaround time. This not only inhibits any potential discussions and collective agreements regarding marking, it reduces moderation to an end-point check.

While there has been some doubt regarding the effectiveness of a standardisation meeting (Greatorex and Bell, 2008) the agreement of a marking scheme and model answers available at the outset are deemed to be important to the marking process (Lamming, 2003) particularly at guiding those less experienced members of the marking team. It is however, the continuous cycle of communication and feedback on the performance of all markers that is clearly evident in the professional associations' approach which drives quality assurance. Wyatt-Smith (2010) feels it is the sharing of interpretations of the assessment tasks and agreeing commonality not only at the outset but throughout the assessment process, that is vitally important.

Similarly, it is through the interaction between members of the marking team that not only supports colleagues in consistently assessing requirements and outcomes but provides a more holistic perspective of the assessment process (Beutel et al., 2017). Additionally, meetings and communication between markers facilitates the development of relationships (Grainger et al., 2015) and allows the exchange of expectations between the assessment setter and those marking (Crimmins et al., 2016). It is this development of an assessment-community that Bloxham et al (2016) feels can lead to lead to greater reliability of marking, more than merely focusing on assessment criteria. Moreover, the encouragement of critical discussions between markers can help improve overall reliability rather than relying on moderator-based marking (Garry et al., 2005). In addition, interaction and engagement in dialogue gives consistency of marking assessment tasks (Watty et al., 2014). This will however, be aided by having a regular marking team something that is not always possible in higher education institutions where staff and/or curriculum changes are increasingly frequent occurrences.

However, it is not just change in staffing and curriculum that are create challenges to assessments in universities. Increased teaching loads, outcomes orientation especially with research and administrative burdens, align with concerns of managerialism (see: Huang et al., 2016; Kok et al., 2010; Reid, 2009). These will potentially reduce staff's willingness to embrace new marking regimes, such as those professional practices, particularly where delays and further work might occur.

Those assessing for professional associations are normally paid directly, on a piecework basis for such work and if not performing up to an agreed standard can have their services discontinued. University staff in contrast, are not paid anything additional to their salary, but absorb the marking

into their normal workload. Compressed turnaround times for marking students' assessment across a number of modules in a semester increase pressure and therefore stress on university staff (Biron et al., 2008; Gillespie et al., 2001). As Knott et al. (2015) note perhaps, training and development might help staff especially part-time colleagues adjust to any new marking regime and also improve the reliability in assessment marking? This is however, subject to the availability of training budgets to facilitate this happening.

Conclusion

We have outlined the prevailing issues concerning marking of assessments in higher education and based on our experience with working with professional management associations, described a possible solution. Nonetheless, it remains our viewpoint until it is tested in the university sector and its value scrutinised. There is however, further investigation of assessment practices of other sectors that universities might consider, such as schools, 'A' levels, as well as those pertaining to professional accreditation that might be usefully transferred. Other assessment practices such as negotiated student self and peer assessment (Poon et al., 2009; Vista et al., 2015) and those supported by technology, such as online discussions (Wood and Henderson, 2010) may well provide solutions not only to assessment consistency, but also to help reduce staff workloads?

References

- Bandiera, O., Larcinese, V. and Rasul, I. (2010) The impact of class size on the performance of university students, available at: <http://www.voxeu.org/article/impact-class-size-performance-university-students>.
- Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E., Boud, D. and Joughin, G. (2016) How university teachers design assessments: a cross-disciplinary study, *Higher Education*, DOI 10.1007/s10734-016-0027-7.
- Beutel, D., Adie, L. and Lloyd, M. (2017) Assessment moderation in an Australian context: processes, practices, and challenges, *Teaching in Higher Education*, 22:1, 1-14.

- Biron, C., Brun, J. and Ivers, H. (2008) Extent and sources of occupational stress in university staff, *Work*, 30, 511-522.
- Bloxham, S. (2009) Marking and moderation in the UK: false assumptions and wasted resources, *Assessment & Evaluation in Higher Education*, 34(2), 209-220.
- Bloxham, S., Boyd, P. and Orr, P. (2011) Mark my words: The role of assessment criteria in UK higher education grading practices, *Studies in Higher Education*, 36(6), 655-670.
- Bloxham, S., den-Outer, B., Hudson, J. and Price, M. (2016) Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria, *Assessment & Evaluation in Higher Education*, 41(3), 466-481.
- Brooks, V. (2012) Marking as judgment, *Research Papers in Education*, 27(1), 63-80.
- Crimmins G., Nash, G. Oprescu, F., Alla, K., Brock, G., Hickson-Jamieson, B. and Noakes, C. (2016) Can a systematic assessment moderation process assure the quality and integrity of assessment practice while supporting the professional development of casual academics? *Assessment & Evaluation in Higher Education*, 41(3), 427-441.
- Dracup, C. (1997) The reliability of marking on a psychology degree, *British Journal of Psychology*, 88, 691-708.
- Garry, J., McCool, M. A. and O'Neill, S. (2005) Are Moderators Moderate?: Testing the 'Anchoring and Adjustment' Hypothesis in the Context of Marking Politics Exams, *Politics*, 25(3), 191-200.
- Gillespie, N. A., Walsh, M., Winefield, A. H., Dua, J. and Stough, C. (2001) Occupational stress in universities: Staff perceptions of the causes, consequences and moderators of stress, *Work & Stress*, 15(1), 53-72.
- Grainger, P., Adie, L. and Weir, K. (2015) Quality Assurance of assessment and moderation discourses involving sessional staff, *Assessment & Evaluation in Higher Education*, DOI:10.1080/02602938.2015.1030333.
- Greatorex, J. and Bell, J. F. (2008) What makes AS marking reliable? An experiment with some stages from the standardisation process, *Research Papers in Education*, 23(3), 333-355.

Guardian (2015) UK university class sizes falling, study shows, *The Guardian* [online] of 27 April 2015.

Havergal, C. (2015) Growth in student applications to UK universities slows, *TES* [online] 9th July 2015.

Huang, Y., Pang, S. and Yu, S. (2016): Academic identities and university faculty responses to new managerialist reforms: experiences from China, *Studies in Higher Education*, DOI: 10.1080/03075079.2016.1157860.

Kalthoff, H. (2013) Practices of grading: An ethnographic study of educational assessment, *Ethnography and Education*, 8(1), 89-104.

Knott, G., Crane, L., Heslop, I. and Glass, B. D. (2015) Training and Support of Sessional Staff to Improve Quality of Teaching and Learning at Universities, *American Journal of Pharmaceutical Education*, 79 (5) Article 72, 1-8.

Kok, S., Douglas, A. , McClelland, B. and Bryde, D. (2010) The Move Towards Managerialism: Perceptions of staff in “traditional” and “new” UK universities, *Tertiary Education and Management*, 16(2), 99-113.

Lamming, D. (2003) Marking university examinations: some lessons from psychophysics, *Psychology Learning and Teaching*, 3(2), 89-96.

Light, G., Cox, R. and Calkins, S. (2014) *Learning and Teaching in Higher Education. The Reflective Professional*, (2nd Ed.), London, Sage.

Mayhew, K., Deer, C. and Dua. M. (2004) The move to mass higher education in the UK: many questions and some answers, *Oxford Review of Education*, 30(1), 65-82.

Mulryan-Kyne, C. (2010) Teaching large classes at college and university level: challenges and opportunities, *Teaching in Higher Education*, 15(2), 175-185.

Nicol, D. (2010) From monologue to dialogue: improving written feedback processes in mass higher education, *Assessment & Evaluation in Higher Education*, 3(5), 501-517.

O'Hagan, S. R. and Wigglesworth, G. (2015) Who's marking my essay? The assessment of non-native-speaker and native-speaker undergraduate essays in an Australian higher education context, *Studies in Higher Education*, 40(9), 1729-1747.

Pinot de Moire, A., Massey, C., Baird, J. and Morrissy, M. (2002) Marking consistency over time, *Research in Education*, 67, 79-87.

Poon, W., McNaught, C., Lam, P. and Kwan, H. S. (2009) Improving assessment methods in university science education with negotiated self- and peer-assessment, *Assessment in Education: Principles, Policy & Practice*, 16(3), 331-346,

Reid, I. C. (2009) The contradictory managerialism of university quality assurance, *Journal of Education Policy*, 24(5), 575-593.

Sadler, D. R. (2005) Interpretations of criteria-based assessment and grading in higher education, *Assessment & Evaluation in Higher Education*, 30(2), 175-194.

Sadler, R. R. (2010) *Beyond feedback: Developing student capability in complex appraisal*, *Assessment & Evaluation in Higher Education*, 35(5), 535-550.

Salamonson, Y., Halcomb, E. J., Andrew, S., Peters, K. and Jackson, D. (2010) A comparative study of assessment grading and nursing students' perceptions of quality in sessional and tenured teachers, *Journal of Nursing Scholarship*, 42(4), 423-429.

Scott, P. (2005) Mass higher education – ten years On, *Perspectives: Policy and Practice in Higher Education*, 9(3), 68-73.

Smith, C. (2012) Why should we bother with assessment moderation? *Nurse Education Today*, 32, e45-e48.

Smith, E. and Coombe, K. (2006) Quality and qualms in the marking of university assignments by sessional staff: an exploratory study, *Higher Education*, 51, 45-69.

Smith, K. (2006) The conflicts and contradictions of working in a mass higher education system, *Perspectives: Policy and Practice in Higher Education*, 10(4), 97-101.

Teichler, U. (2001) Mass higher education and the need for new responses, *Tertiary Education and Management*, 7(1), 3-7.

Vista, A., Care, E. and Griffin, P. (2015) A new approach towards marking large-scale complex assessments: Developing a distributed marking system that uses an automatically scaffolding and rubric-targeted interface for guided peer-review, *Assessing Writing*, 24, 1-15.

Watty, K., Freeman, M., Howieson, B., Hancock, P., O'Connell, B., de Lange, P. and Abraham, A. (2014) Social moderation, assessment and assuring standards for accounting graduates, *Assessment & Evaluation in Higher Education*, 39(4), 461-478.

Wood, E. H. and Henderson, S. (2010) Large cohort assessment: depth, interaction and manageable marking", *Marketing Intelligence & Planning*, 28(7), 898-907.