Improving Dissertation Assessment

Chaminda Pathirage, c.p.pathirage@salford.ac.uk
Richard Haigh, r.p.haigh@salford.ac.uk
Dilanthi Amaratunga, r.d.g.amaratunga@salford.ac.uk
David Baldry, d.baldry@salford.ac.uk
Catherine Green, c.m.green@salford.ac.uk

Abstract

During the last decade, there have been increasing calls for higher education to improve standards, increase the quality of assessment, and for greater accountability of lecturers. It is recognised that consistency in assessment is even more important where assessment is through one large piece of work, such as a dissertation, and where the assessment outcome will have a significant impact on the final grade of students. Dissertation modules typically pose further problems for assessment consistency due to the large number of students and the resultant need for large numbers of lecturers to participate in its assessment. This paper synthesises the initial literature findings from an on-going research project that aims to identify good practices for dissertation assessment, in an attempt to improve the quality and consistency of assessment.

Introduction

This paper synthesises the initial literature findings on dissertation assessment practices, based on work undertaken as part of an on-going research project that is examining assessment practices for undergraduate dissertation modules. The research is being undertaken by the School of Construction & Property Management (SCPM), at the University of Salford. A range of assessment practices will be examined, across disciplines and universities, and therefore the findings will have wider applicability to Schools across the University and beyond. The 12 month project – due for completion in April 2005 – is funded under the University of Salford Teaching and Learning Quality Improvement Scheme (TLQIS).

The paper is divided into three main sections. Firstly, it explores the increasing concerns in the UK higher education system regarding assessment practices. Secondly, the paper discusses the literature findings on dissertation assessment practices, while also presenting the reasons for undertaking this research. This section is further subdivided into three parts, in which, quality, consistency, and criteria of assessment are outlined. Thirdly, the paper focuses on the findings from the analysis done on the existing practices of the dissertation assessment and highlights the various assessment practices followed by different schools and universities. Finally, the paper sets out the way forward for the research.

Increased Concerns in Higher Education

During the last decade, a period of considerable change within the higher education system, there have been increasing concerns regarding the quality of teaching, learning and assessment practices within higher education institutions. In the more centralised political culture of the UK there have been strong pressures, even in the context of the rapid expansion of higher education, to hold on to the principle of high academic standards (Lucas & Webster, 1998). Thus the Higher Education Quality Council (HEQC)

and now the Quality Assurance Agency for Higher Education have defined academic standards as 'explicit levels of academic attainment which are used to describe and measure academic requirements and achievements of individual students and groups of students' (HEQC, 1997, p. 5). As contended by Webster et al (2000) not only within the UK, but internationally, there are moves that focus the internal and external quality assurance of teaching on assessment, and particularly on the standards (or 'competencies') of graduating students to ensure that public and private investments in higher education are both accountable and provide 'value for money'. Indeed, many states are mandating graduate assessment strategies and evidence of student accomplishment as a condition of state financial support (*ibid*, p 72).

However, concerns about the rigour of assessment standards have grown with the rapid growth in UK higher education student numbers, class sizes and student-staff ratios, and with a concurrent increase in the proportion of students getting first and upper second class degrees (Chapman, 1994). Consequently several reports have been published, addressing different aspects of assessment in higher education, including, The Reynolds Report (1985), The Harris Report (1996), and The Dearing Report (1997). The Harris Report's (HEFCE, 1996) discussion on quality and standards in higher education highlighted the importance of assurance as to the methods used for assessment and the need for greater innovation in assessment techniques, although it was primarily concerned with postgraduate education. One vital aspect of the Dearing model (Dearing report, 1997) was its emphasis on the need for university teachers to ensure effectiveness in assessing students and in giving feedback.

In addition, several educational committees and agencies have been established due to this widespread interest in higher education. The Institute for Learning and Teaching in Higher Education (ILTHE, now the Higher Education Academy) was established as a response to the requirements highlighted in the Dearing Report. As a consequence of criticisms raised by academics on the Teaching Quality Assessment (TQA) process for being expensive and intrusive, a new Quality Assurance Agency (QAA) for Higher Education was established, which sought to deliver and maintain high standards, particularly through focusing on student assessment and also by promoting transparency. Section 6 of code of practice "Assessment of Students" (May, 2000), published by QAA, has stipulated a number of requirements and expectations in assessing students which are to be followed by higher education institutions, further emphasising the necessity for increased accountability of lecturers, improved quality of assessment and greater consistency of standards.

Particularly important at undergraduate level are assessments that contribute to degree classification, and which thereby present to employers, as well as postgraduate admissions tutors, staff judgements of the standard of student work. Having identified the increased concerns placed for assessment in higher education, the following section examines the assessment of undergraduate dissertation which has a large bearing towards the ultimate degree classification of the students.

Dissertation Assessment

The necessity to ensure quality, consistency and improved criteria of assessment is greatly emphasised with modules where assessment is through one large piece of work such as a dissertation. It is widely acknowledged that the undergraduate dissertation is special both to teachers and to students. From the students' point of view, the dissertation

is the single most substantial, and independently worked upon, piece of work they will undertake while at the university (Webster et al, 2000). It is not surprising then, that dissertations are surrounded with anxiety and uncertainty, so that students look for reliable maps that will guide them through dissertation process (*ibid*, p 72).

From the assessors' perspective, the assessment of a dissertation is also significant since such modules can account for up to 30 percent of marks awarded in a year. Therefore any inconsistencies in assessment will almost certainly be reflected in students' overall grade for the year and ultimately the final degree classification (Saunders and Davis, 1998). Dissertation modules typically pose further problems in consistency of assessment due to the large number of students and the consequential need for large numbers of lecturers to participate in its assessment. As the size of the team expands so the difficulties associated with achieving and maintaining consistency of assessment between lecturers becomes more apparent. However, in spite of the dissertation's status within degree courses and its perceived educational value and challenges, the assessment of the dissertation appears to be relatively under-explored within the published research literature in the UK (Todd et al, 2004). Three major areas were highlighted in the literature in relation to dissertation assessment, namely quality, consistency and criteria of assessment. The succeeding sections outline the literature findings on these major areas.

Quality and Consistency in Assessment

The literature survey revealed the increased concern in terms of quality of the assessment practices which emphasised the maintenance of the 'gold standard' of current assessment practices by individuals, departments and institutions involved with higher education (Webster et al, 2000; Saunders and Davis, 1998). This is further highlighted by the HEQC:

> Student assessment is clearly central to standards. If the work of students is not assessed by valid and reliable methods, standards cannot be rigorous.
>
> (Higher Education Quality Council, 1997, pp. 8, cited in Webster et al, 2000)

As previously mentioned, the QAA code of practice (Section 6) on assessment of students can be perceived as a means of regularising the assessment of undergraduate students, which is directly applicable for undergraduate dissertation assessment as well. The following list details some of the requirements stipulated within this code of practice:

- The principles, procedures and process of all assessment to be explicit.

- Publication of clear rules and regulations governing the conduct of assessment.

- Publication & implementation of consistently clear criteria for the marking and grading of assessment.

- Appropriate feedback to students on assessed work.

- Competent staff to undertake roles and responsibilities in assessment work.

It is questionable how far higher education institutions adhere and follow these stipulated requirements of QAA, at least when it comes to the assessment of dissertation, which has a large bearing towards the ultimate degree classification of the students.

Recent concern in higher education has also focused on the need for greater accountability of lecturers and on ensuring consistency of standards (Aper et al, 1990; Brown et al, 1995; Norton, 1990). Consistency of standards in assessment is important

for all assessed work, as it incorporates issues such as the subjectivity of the individual lecturer, uniformity between lecturers for a single piece of work and ensuring the same standards across pieces of work from similar modules for different courses (Saunders and Davis, 1998). However, the literature reveals several important factors which directly contribute on the consistency of dissertation assessment.

Scepticism of the lecturer's on their own decision is believed to be a major contributor for the inconsistency in dissertation assessment (Rowntree, 1987). The following comments made by several assessors highlight this issue.

'Real evidence of awareness of the various perspectives', mark awarded 46%

'results section unclear', mark awarded 57%,

'this is a clear, well presented [dissertation]… which fulfils its specific aims', mark awarded 49%

(cited in Webster et al, 2000)

In addition, time spent on assessment, relative experience of the lecturer, lecturer's attitude/ values and ownership of the criteria were considered to be the other leading determinants of the consistency in dissertation assessment. It was apparent that, in general terms, the longer a lecturer had spent assessing a dissertation, the lower the grade it received. As such it is argued that a lecturer should not revisit a piece of work that has already been rigorously assessed against the criteria.

Relative levels of experience of assessing dissertations were also felt to have been an important contributory factor. As Balla and Boyles (1994) and Brown et al (1995) contend, lecturers need to be involved in the development of criteria so as to create the ownership of the criteria used for the dissertation assessment. As such, criteria designed carefully and used with clear procedures can reduce inconsistency in assessment and joint development of criteria by those assessing the work provides a useful start for ensuring that each lecturer understands them in the same way. This enables lecturers to be more certain as they are following the same process and judging each piece of work against the same criteria, thereby assessing each student in the same way. Having discussed about the factors affecting quality and consistency of dissertation assessment, the following section outlines the literature pertaining to assessment criteria.

Criteria in Assessment

Assessment criteria are widely used in the education system when students' work is being marked. It is good practice to publish, explain and clarify on what base students are assessed, treating each student similarly, fairly and with consistency (as stipulated in QAA code of practice). Two different types or extremes of assessment criteria practiced in dissertation assessment were unearthed, namely impressionistic/ holistic and analytic (weighting method). The grade or the final marks for the dissertation was arrived on the basis of impression made in the holistic method where as in analytic method marks were given against each category based on a predetermined mark (Harris and Bell, 1994). It is argued that students' awareness about the relative importance attributed by markers to each criterion used is of immense importance for the students to get the maximum out of the assessment. Yet, a holistic framework, using criteria to rationalise an overall mark has the considerable advantage of maximising flexibility from the assessors' point of view.

Adding to this dilemma, much concern is expressed in the literature against considering the assessment criteria as a "straight-jacket" (Balla and Boyle, 1994) which hinders the students' creativity and individuality. It is argued that by having an analytic or weighted method of criteria, the process of assessment is much more standardised than having an impressionistic based criteria. As contended by Webster et al (2000), if the dissertation is a very individual piece of work presented by students, surely it is the last piece of work which anyone would want to standardise by insisting the same or similar criteria and approaches. However, this has already been manifested in the scholarly literature between those who argue for professional autonomy, and those who emphasise the need for public accountability; between those who see a need for explicit criteria and performance standards in assessment, and those who regard assessment as akin to wine tasting (De Vries, 1996; Wright, 1996).

Furthermore, Hands and Clewes (2000), whilst acknowledging the value of criterion referencing, have pointed out that too many criteria, specifically to the marking of dissertations, could diminish the importance of tutors' judgments and lead to an increase in 'marking fatigue' which itself is a cause of much variability found in assessment quality. Nevertheless, assessment criteria can be seen as an important tool for giving new assessors confidence to take part in the assessment process. This is important as many academics report feelings of discomfort and fear when participating in exam boards or when double-marking work (Hand and Clewes, 2000). Partington (1994) has gone so far as to suggest that explicit assessment criteria that are freely available to staff and students should negate the need for double-marking.

Two marking strategies which need to be avoided are also highlighted within the literature, namely the 'defensive marking strategy' and 'game theory'. In defensive marking strategy the assessors avoid giving very high or low marks for the students making them unnoticeable to stakeholders (colleagues, external examiners). Game theory suggests that staff may try to anticipate the reaction of other stakeholders in the process, thereby marking dissertations to have marks close to the average with a very narrow range of marks. It was observed that assessors deploy these strategies especially when double marking is followed.

This on-going research project aims to identify good practices for undergraduate dissertation assessment, by addressing the quality, consistency and criteria of assessment as discussed above.

Research Methodology

The research is being carried out according to four work packages, as Figure 1 illustrates. This paper reports the outcomes of work package one (WP1) which reviewed the literature and existing practices pertaining to undergraduate dissertation assessment. Outcomes and the understanding obtained from the literature review stage (WP1) will be fed into the next pilot study phase (WP2), in which a series of workshops will be organised. These workshops will be used to pilot a range of assessment approaches and criteria in an attempt to measure and ultimately improve assessment consistency within the School's dissertation module on undergraduate programmes. Further, a sample of students – that includes graduates from previous years and current final year undergraduates – will be interviewed to ascertain student understanding of dissertation requirements and assessment criteria. The project will culminate in the publication of good practice guidelines (WP3), outlining good practices from other universities and

disciplines, as well the results of pilot studies undertaken as part of the research. Finally, the project's findings will be disseminated (WP4) to inform the teaching and research community, both internally and externally.
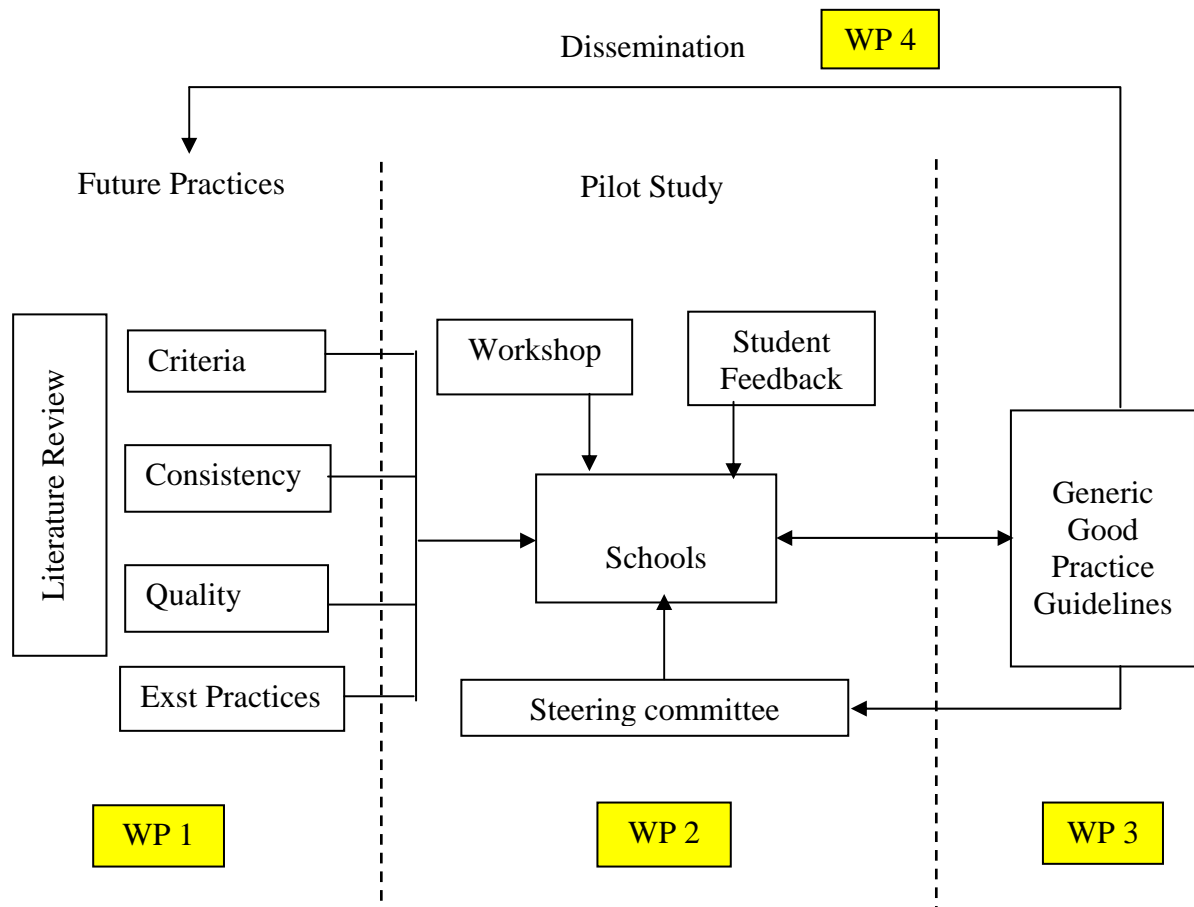
WP1 – Literature review

WP2 – Pilot studies

WP3 – Development of generic good practice guidelines

WP4 – Research Dissemination

Figure 1: The Project's Research Methodology

Analysis of Existing Assessment Practices

The selection of the dissertation practices were done to reflect the procedures adhered by different countries, different universities and by different disciplines. Accordingly, 30 dissertation practices were scrutinised, based on the dissertation module handbooks obtained online, including those from England, Australia, United States and Sri Lanka. In addition to the courses offered (Built Environment) by the School of Construction and Property Management, University of Salford, practices followed by disciplines including Social Work Studies, Business and Management, Geography, Languages, Economics, Environmental & life Science, History and Art & Design were chosen for analysis. Table 1 provides an overview of existing dissertation practices, scrutinised according to the

country and the discipline. The most commonly covered areas within the practices were the assessment procedure, guidelines/ instructions for dissertation production and the assessment criteria.

Table 1: An Overview of Existing Practices Scrutinised

Country \ Discipline	Engineering Science	Business, Management & Economics	Social Science, Languages & Environmental Studies	History, Art and Design	Total
England	7	6	4	2	19
Australia	1	3	1	1	6
United States	1	3	-	-	4
Sri Lanka	1	-	-	-	1
Total	10	12	5	3	30

Assessment Criteria

Approximately 70% of the analysed practices had explicit criteria, out of which, two thirds represented holistic or impressionistic methods of assessment (refer Criteria in Assessment section for explanation). As such 30% of the practices had provided just a style manual, which did not specify any assessment criteria for the student. This clearly contradicts with the requirement of "Publication & implementation of consistently clear criteria for the marking and grading of assessment", stipulated in the QAA code of practice as mentioned elsewhere. It was observed that the impressionistic method was the most common method of assessment, which negates the argument of considering the assessment criteria as a straight-jacket.

The number of categories within the criteria varied from four to ten with an average of six. Being parallel to the argument put forward by Hands and Clewes (2000) on too many criteria (refer criteria in assessment), Laming (2003) offered some interesting evidence from his comparison of findings on judgment in psychophysical experiments to highlight that human markers find it difficult to reliably distinguish between more than five discrete categories. As such it is questionable to have too many categories as revealed in actual practice. The most frequently found categories within the dissertation assessment criteria together with their relative importance placed by the courses are depicted in Table 2.

Assessment Process

Several different approaches for dissertation assessment were revealed from the analysis. In summary, the dissertation assessment process comprised of four different forms or methods of evaluation, these were research/ dissertation proposal, written dissertation, performance of the student and an oral presentation. All courses, either purely or substantially, based their assessment of the dissertation module on the written outcome (i.e. dissertation). Interestingly some practices assessed the performance of the student when deriving the marks for the dissertation module. The criteria for the assessment performance of the student included categories like enthusiasm and self motivation, time

management, communication, record keeping and so on. This inclusion may justify the argument that the process the student has gone through should be reflected in dissertation module assessment and not only the final outcome of the student. Table 3 indicates the relative importance placed on different forms of assessment.

Table 2: The Range of Relative Importance Apportioned to Criteria across Disciplines

Category	Relative Importance
Introduction *(Abstract, Objectives, Background, Context)*	10-25%
Knowledge in relevant Discipline *(Sources, Use & Analysis of Lit, Theories)*	20-30%
Methodology *(Experimental methods, Research design, Ethical dilemmas)*	10-25%
Analysis & Discussion of result *(Presentation, Clarity, Logical arguments)*	20-40%
Conclusion & Recommendations	5-10%
Presentation & Communication *(Structure, Organisation, Referencing, Language)*	10%
Others *(Relevance, Originality, Contribution, Future work, Scope & Difficulty)*	10-20%

Table 3: The Range of Relative Importance Apportioned to Forms of Assessment across Disciplines

Form of Assessment	Relative Importance
Research/ Dissertation Proposal	10%-25%
Written Dissertation	60%-100%
Performance of the Student	20%-35%
Oral Presentation	20%-30%

As a written dissertation was found to be the only common form of assessment for all the courses, this will be analysed to highlight the range of assessment processes adhered across disciplines and schools. Even though most of the schools appointed one supervisor for a dissertation student, noticeably dissertations that involved more than one discipline required two supervisors. Also some schools, as a matter of policy, operate this double supervisory mechanism even within the same discipline. In a majority of courses, the written dissertation was double marked - assessed by the supervisor and at least by one other staff member - and moderated by members of the supervisory group. Although Partington (1994) argued that explicit assessment criteria when freely available to staff and students should negate the need for double-marking, in practice the double marking mechanism was found to be very common. Some practices further extended this double marking system by deploying two blind markers to eradicate the bias of the supervisor.

When disagreements occur between two markers, these are generally resolved between the two assessors of staff and where this is not possible they are referred either to a third examiner within the staff or to an external examiner. Interestingly some used the viva mechanism to resolve the disagreement within the two markers instead of referring it to a third examiner. These different procedures followed in written dissertation assessment process are depicted in

Figure 2.

Supervision		
	Supervisor 1	Supervisor 2

- -

Submission Written Dissertation

- -

Assessment Supervisor/ First Mk   Second Marker

- -

Agreement ———— Yes ————

Moderation

No

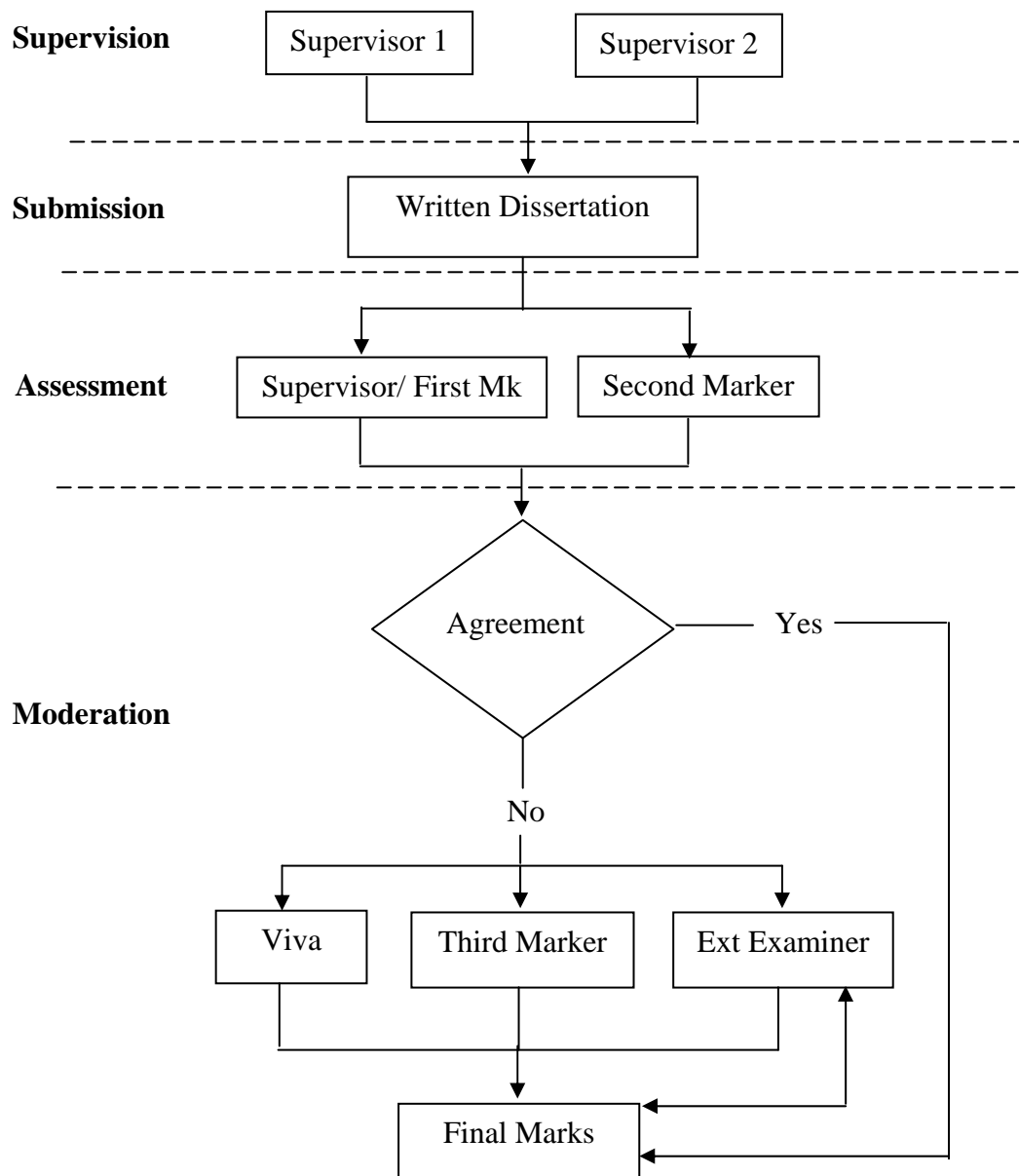Viva   Third Marker   Ext Examiner

Final Marks

Figure 2: A Flow Chart - based on the survey of existing practices that illustrates the range of written dissertation assessment processes across disciplines and schools

Way Forward

This paper is based on the early findings of a research project that is attempting to identify good practices for dissertation practices on undergraduate programmes. It summarises the literature pertaining to dissertation assessment across a range of disciplines and universities, and in doing so, highlights the many challenges that a programme leader faces, when devising an assessment strategy for a dissertation module. The project's future work includes a series of workshops, primarily within SCPM, and obtaining student feedback as discussed in the research methodology section. The project will culminate in the publication of good practice guidelines to disseminate the project's findings.

References

Aper, J. P. Cuver, S. M.. and Hinkle, D. E., (1990), Coming to terms with the accountability versus improvement debate in assessment, Higher Education, Vol. 20, pp. 471-83

Balla, J. and Boyle, P. (1994), "Assessment of student performance: a framework for improving practice", Assessment and Evaluation in Higher Education, Vol. 19 No. 1, pp. 17-28

Baume, D., Yorke, M and Coffey, M., (2004), what is happening when we assess, and how can we use our understanding of this to improve assessment?, Assessment & Evaluation in Higher Education, Vol 29 (4), pp 452-477

Brown, S. Race, P. and Rust, C., (1995), Using and experiencing assessment, in Knight, P. (Ed.), *Assessment for Learning in Higher Education*, Kogan-Page, London

Chapman, K., (1994), Variability in degree results in Geography in British universities, 1973–1990: preliminary results and policy implications, *Studies in Higher Education*, Vol 19 (1), pp. 89–102

Cowan, J., (2004), Plus/ Minus marking: A method of assessment worth considering?, ILTHE (Incorporated into the Higher Education Academy) Assessment Article, Vol 5 (1)

De Vries, P., (1996), Could 'criteria' in quality assessments be classified as academic standards?, *Higher Education Quarterly*, 3, July, pp. 193–206

Hand, L. and Clewes, D., (2000), Marking the difference: an investigation of the criteria used for assessing undergraduate dissertations in a business school. *Assessment & Evaluation in Higher Education,* Vol. 25, 5-21

Harris, D. and Bell, C. (1994), Evaluating Assessing for Learning, Kogan-Page, London

Harris, M. (1996) *The Harris Report: Review of Postgraduate Education.* HEFCE

Higher Education Quality Council, (1997), *Graduate Standards Programme*, Final Report, (London, HEQC)

Joughin, G and MacDonald, R., (2004), A model of assessment in higher education institutions, ILTHE (Incorporated into the Higher Education Academy) Assessment Article, Vol 5 (1)

Laming, D., (2003), Marking university exams. Presentation at one day seminar on Assessment in Psychology degrees, St Barts Hospital , London, 21 March, 2003

Lucas, L. & Webster, F. (1998) *Maintaining standards in higher education? A case study*, in: David Jary & M.Parker (Eds) The New Higher Education: issues and directions for the post-Dearing university (Stoke-on-Trent, Staffordshire University Press), pp. 105–113

Norton, L.S. (1990), Essay writing: what really counts?, Higher Education, Vol. 20, pp. 411-422

Partington, J. (1994) Double marking students' work. *Assessment & Evaluation in Higher Education,* 19, 57-60

Pepper, D., Webster, F and Jenkins, A., (2001), Benchmarking in geography: some implications for assessing dissertation in the undergraduate curriculum, Journal of Geography in Higher Education, Vol 25 (1), pp 23-35

Rowley, J and Slack, F., (2004), What is the future for undergraduate dissertation?, Education and Training, Vol 46 (4), pp 176-181

Rowntree, D., (1987), Assessing Students: How Shall We Know Them?, Kogan-Page, London

Rust, C., Price, M and O'Donovan, B., (2003), Improving students' learning by developing their understanding of assessment criteria and process, Assessment & Evaluation in Higher Education, Vol 28 (2), pp 147-164

Saunders, M. and Davis, S., (1998), The use of assessment criteria to ensure consistency of marking, *Quality Assurance in Education*, Vol. 6 (3), pp162-171

Section 6: Assessment of Students, (2000), Code of practice for the assurance of academic quality and standards in higher education, Quality Assurance Agency for Higher Education, May 2000, Gloucester

Stowell, M., (2004), Equity, jusice and standards: assessment decision making in higher education, Assessment & Evaluation in Higher Education, Vol 29 (4), pp 495-510

Todd, M., Bannister, P and Clegg, S., (2004), Independent inquiry and the undergraduate dissertation: perceptions and experiences of final-year social science students, Assessment & Evaluation in Higher Education, Vol 29 (3), pp 335-355

Webster, F, Pepper, D and Jenkins, A., (2000), Assessing the Undergraduate Dissertation, *Assessment & Evaluation in Higher Education,* Vol 25 (1), pp 71-80

Woolf, H., (2004), Assessment criteria: reflection on current practices, Assessment & Evaluation in Higher Education, Vol 29 (4), pp 480-494

Wright, P., (1996), Mass higher education and the search for standards: reflections on some issues emerging from the Graduate Standards Programme, *Higher Education Quarterly*, Vol 50 (1), pp. 71–85