Programme-Level Assessment:

A review of selected material

Prepared by:
Jane Hudson, Research Assistant
ASKe, Oxford Brookes University

Table of contents

Contents

Programme-level Assessment: A review of selected material

1   Introduction

This literature review supports and contextualises the Bradford PASS group's aim of carrying out case studies of existing practice and surveys of opinions on programme-level assessment (PLA) in a number of partner institutions. It reviews selected literature that can inform the design, implementation and evaluation of 'effective, efficient, inclusive and sustainable assessment strategies that deliver key course/programme outcomes' (Bradford). In a way that is both descriptive and synthetic, the review considers practice-oriented issues such as agreeing goals and evaluating success.

1.1   Project focus

The review seeks to answer the following questions:

> *What constitutes programme-level assessment (5.2, 5.3.3)*
> *What factors should we consider when designing programme-level assessment? (5.3.2)*
> *What models for programme-level assessment have been implemented, and which ones are emerging? (5.3.3)*
> *What enables the implementation of programme-level assessment and what are the potential challenges and obstacles to take into account? (5.4)*
> *How can we evaluate the effectiveness of programme-level assessment? (5.5)*

Responses to the questions are found throughout the literature review, but are most clearly articulated in the sections indicated in parentheses above.

2   Review framework

Books, book chapters and 'grey' literature are all included, but the review emphasises journal articles. The review includes research- and practice-based literature, as well as a few opinion pieces.

The literature review focuses primarily on higher education, but takes into account other literatures where they are more advanced on a specific topic, such as the secondary school literature on collective work. The literature review concentrates on UK-based research, but relevant research based elsewhere is included. An important consideration in international searches is that of terminology, with the American use of the 'assessment' referring to what in the UK is understood as 'evaluation'.

In the literature, three different understandings of the phrase 'programme-level assessment' overlap. In some instances, this phrase describes the evaluation of programmes. In others, it describes the assessment of students at the end of their studies in order to gauge whether or not they have met the objectives set by that programme. A final understanding of programme-level assessment relates to the assessment of students in a coherent way which is integrated

across years of study. This last understanding is that which is of interest to this project. That said, all three understandings of programme-level assessment overlap (for example, portfolios are relevant in all three discussions). Regardless of their primary focus, articles that inform the design or implementation of integrated assessment strategies are included in this review.

Although the topic of this review is programme-level assessment, some authors discussed 'curriculum', 'degree,' 'course' or 'department' assessment strategies. Where these were found to be dealing generally with the same set of issues, they were included.

The literature review focuses on undergraduate, and 'in person' rather than web-based courses, but e-learning as a component of a programme is considered. The literature review generally assumes a modular programme for two reasons: firstly, the universities involved in the PASS pilot projects use modular schemes; secondly, traditional Oxbridge tutorial teaching systems generally already take a programme-level approach to assessment.[1] Gibbs and Dunbar-Goddet point out,

> Oxbridge is also commonly characterised by the study of a single subject in depth and relatively little choice of course units (though there are exceptions to this pattern) and so assessment is usually designed more at the level of the whole programme than at the level of the individual course unit (2009, p. 483).

In modular programmes, on the other hand, 'there may be more emphasis on the design of assessment within individual course modules than on the coherence or integrity of assessment at the level of the programme' (Gibbs & Dunbar-Goddet, 2009, p. 483).

3   Methodology

Because there is not yet a consensus among researchers on a single term to describe 'programme-level assessment strategies', the review took as its starting point the identification of search terms. This was accomplished though a systematic reading of abstracts in eight selected journals published from January 2005 through March 2010. Relevant articles identified were mined for alternative keywords to be used in database searchers, and their bibliographies were reviewed for further publications of interest, published at any date. Journals reviewed in the initial systematic reading included:

> *Active Learning in Higher Education*
> *Assessment & Evaluation in Higher Education*
> *Higher Education*
> *Higher Education Quarterly*
> *Journal of Further and Higher Education*
> *Quality in Higher Education*
> *Quality and Assurance*
> *Studies in Higher Education*

Keywords identified included:

---

[1]Gibbs and Dunbar-Goddet (2009) include a third category, traditional research intensive universities.

Aligned learning outcomes	Final year & assessment
Assessment & progression	First year & assessment
Assessment procedures	Inclusive assessment
Assessment regime	Integrated assessment strategy
Assessment strategy	Linear programmes
Assessment system	Modularisation debate
Capstone modules	Programme-level assessment
Constructive alignment	Programme-wide assessment
Coordinated assessment	Second year & assessment
Critical moments & assessment	Student progressions
Degree-wide assessment	Sustainable assessment
environment	Systematic course design
Effective assessment	Thresholds & assessment
Efficient assessment	Transition & assessment

Four databases – Academic Search Complete, Education Search Complete (both through EBSCO host), SAGE Journals Online and HEA Evidence net – were searched using the above keywords, with no restriction placed on date of publication.

Research into Higher Education Abstracts from 2005 onwards was also consulted using relevant keywords. The Assessment for Learning CeTL website was reviewed for other resources.

4 Results

The review cites 42 sources attributed to 26 primary authors. A full half of the sources reviewed were authored (as primary author) by the following key researchers on the topic: David Boud (4), Graham Gibbs (3), Peter T. Knight (6), Chris Rust (3) and Mantz Yorke (5). The discussion of programme-level assessment seems to be taking place primarily in *Assessment & Evaluation in Higher Education* and *Active Learning in Higher Education*.

Assessment & Evaluation in Higher Education
   o *Cited: 11 articles published 1998-2009*
Active Learning in Higher Education
   o *Cited: 6 articles published 2000-2009*
Higher Education
   o *Cited: 2 articles published 2003, 2006*
Higher Education Research & Development
   o *Cited: 2 articles published 2001, in press*
Innovations in Education & Teaching International 2001
   o *Cited: 2 articles published 2001, 2003*

Eight journals contribute a single relevant article on the topic:

European Journal of Engineering Education (2004)
Journal of Education for Teaching (2008)
Journal of Geography in Higher Education (1994)
Journal of Higher Education (1988)
*Learning and Teaching in Higher Education (2004-05)*
Studies in Continuing Education (2000)

Studies in Higher Education (2002)
Teaching in Higher Education (2001)

Key books include: Peter T. Knight and Mantz Yorke's As*sessment learning and employability* (2003), David Boud and Nancy Falchikov's *Rethinking assessment in higher education: Learning for the longer term* (2007) and Sue Bloxham and Pete Boyd's *Developing assessment in higher education: A practical guide* (2007).

The chart below shows that the academic conversation about programme-level assessment strategies has been taking place, albeit quietly, over the past decade. The context of this debate is reviewed in the following section.

Publications cited:
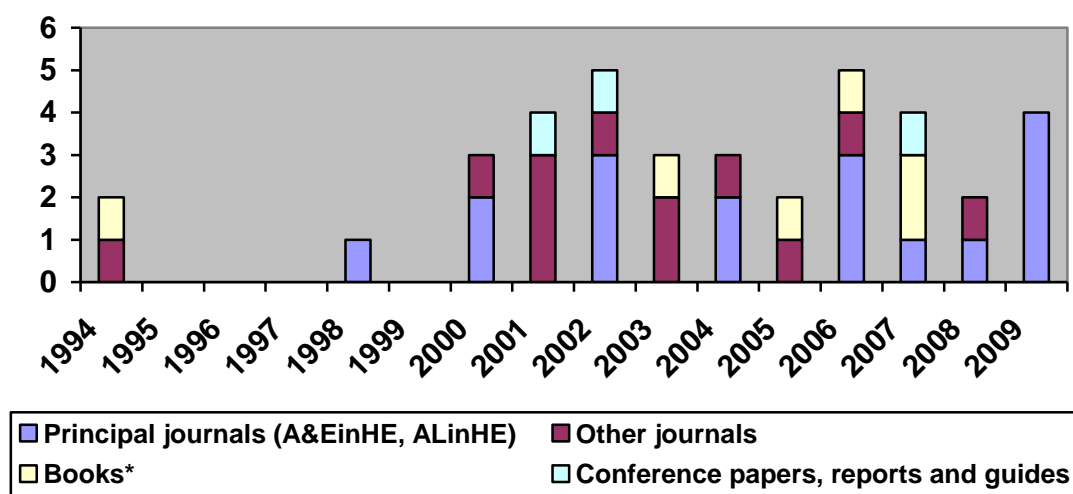
* Multiple chapters from a single edited book (Boud & Falchikov, 2007) not included separately. One early American article (Banta & Schneider, 1988) and one article in submission (Leask & Carroll) are not included.

5   Reviewing the literature

5.1   Context

Currently, UK journals on Higher Education contain far more calls for a more programme-based approach to assessment than practical insights into how to make such changes. These calls have linked programme-level assessment to numerous priorities in UK higher education including improved student learning (Rust, 2002), retention (Yorke, 2001a; Yorke, 2001b), life-long learning and employability (Boud, 2000; Boud & Falchikov, 2006; Knight & Yorke, 2003; Yorke & Knight, 2006), and financial efficiency (Knight, 2000)[2]. These priorities and their links to programme-level assessment are discussed further in section 5.2.

Knight and Yorke (2003, p. 147) explain that before the widespread acceptance of modularity in higher education, curricular coherence and progression were largely understood in terms of

---

[2] Efficiency in terms of staff *time* may be a concern given the demands of the RAE, but this concern is not yet evident in the literature on programme-level assessment.

established disciplinary structures.[3] Knight and Yorke characterise this as a 'supply-side' approach. Although 'supply-side' coherence continues to play a role in the design of programme-level assessment, the focus shifted to 'demand-side' as modular approaches gained currency. Students became more responsible for making sense of the curriculum and for conceptualising progression as they moved through academic years of study. At the same time, 'the move toward modularisation has, in a number of institutions, reduced the opportunity for formative assessment' (Yorke, 1998, n.p.) and contributed to an increase in summative assessments. Semesterisation has amplified the effect (Knight, 2002b).

Knight and Yorke link the current shift back toward a 'supply-side' approach to the government's emphasis on employability. They explain:

> The enhancement of employability requires that 'slow learning' be acknowledged in programme structures, and in the assessment expectations. In the UK, the political interest in employability has caught a tide that was already running in the higher education sector: the interest in programme specifications that had been restimulated by a policy document issued by the QAA (1999). (Knight & Yorke, 2003, p. 147)

A focus on employability (and an associated attention to key skills) and the pursuit of a learning society (along with widening participation and rapid-expansion) are linked to the 1997 Dearing report (Boud, 2000; Leedham, 2009).

Any consideration of assessment practices must take place within what Bloxham and Boyd (2007, p. 5) characterise as the constraints of the academic infrastructure of the QAA and other public policy.[4] As Gibbs and Dunbar-Goddet (2009) explain, the QAA (in line with similar bodies in Europe) 'usually embodies assumptions about what makes for a good assessment regime' (2009, p. 481). Currently, this includes constructive alignment of assessment and learning outcomes (Bloxham & Boyd, 2007, p. 5).

Bloxham and Boyd (2007), who describe the QAA approach as judging assessment practice 'primarily on whether it effectively measures the intended outcomes of a course of study in a valid, reliable and transparent way', underscore that it 'is open to criticism' (Bloxham & Boyd, 2007, p. 5). Bloxham and Boyd (2007, p. 28-29) summarise the arguments against learning outcomes but, despite their reservations, they ultimately decide to fall into stride with this approach because of 'its pervading influence on so much day-to-day institutional practice' (p. 5). The pervading influence of the outcomes-led approach is reflected throughout this review and in the PASS team's project objective of designing 'effective, efficient, inclusive and

---

[3] Yorke (1998) indicates another approach taken to curricula coherence before modular schemes became popular: 'perhaps the closest that higher education has approached to this [extreme rigidity in curricula] has been in connection with some of the behaviourally-based curricula (and hence assessments) that were given impetus by the work of Bloom and his co-workers in the 1950s' (n.p.).

[4] From their study of programme-level and institutional assessment environments, Gibbs and Dunbar-Goddet (2009) conclude that 'at the programme level, assessment regimes appear to differ widely between institutions in relation to these assumptions' (p. 481). That is, 'quality assurance does not seem to have constrained variation in assessment regimes, ensured that quality assurance requirements are met (such as variety of assessment methods aligned to goals) or ensured that characteristics known to support learning (such as formative assessment and frequent, prompt, feedback) are evident' (p. 485).

sustainable assessment strategies *that deliver key course/programme outcomes'* (Bradford, 2009). Knight's (2001) criticisms of the outcomes-led approach are discussed in section 5.3.C of this review.

5.2   What makes assessment effective, efficient, inclusive and sustainable?

As mentioned in the introduction, the PASS project 'confronts a fundamental issue for every HE course/programme leader: how to design an effective, efficient, inclusive and sustainable assessment strategy which delivers the key course/programme outcomes' (Bradford). This section of the literature review seeks to inform the question '*What constitutes programme-level assessment*' by looking specifically at what the literature suggests would constitutes an effective, efficient, inclusive and sustainable assessment strategy.

5.2.1   Effective assessment

Effective assessment is, essentially, assessment that fulfils its objectives, whatever they may be. Objectives may vary between programmes, but will likely be linked to the general purposes of assessment – certification, support of student learning, quality assurance, and development of life-long learning capacity (Bloxham & Boyd, 2007, p. 31-32). For an assessment regime to be successful, it must fulfil these four purposes (although not every assessment or even the assessment plan of every module need satisfy all four – herein lays one benefit of a programme approach).

Bloxham and Boyd (2007) point out that institutions often have distinct goals which can shape the objectives of their assessment strategies. For example, institutions 'promoting research-focused teaching may emphasise the use of inquiry-based assessment, whereas retaining non-traditional students may indicate that the programme should pay particular attention to preparation of students for assessment' (p. 161). Yorke (1998) makes the same point,

> Those institutions espousing a vocational interest, for example, will wish to give a noticeable weighting to generic skill such as team-working, self-reliance, and practical problem solving in respect of the 'messy' problems of the world outside of academe, as well as to the more obviously academic aspects of their studies (n.p.).

These specific objectives will determine how the effectiveness of a given assessment regime should be evaluated.

5.2.2   Efficient assessment

> *Programmatic thinking cannot deliver the chimera of perfect reliability and validity in the assessment of complex learning but it does make it easier to see how to invest in reliability and to indentify where it really matters. (Knight, 2000, p. 239)*

An assessment regime can be efficient in at least two ways; in terms of the quantity of assessments necessary for the assessment strategy to satisfy its purposes and 'by ensuring that no time and effort are wasted as the institution pursues effectiveness' (Yorke, 1998, n.p.).

An assessment strategy can be efficient in the first sense by ensuring that duplication is used only strategically (rather than as a result of poor planning) and by taking a 'minimax' approach – selecting assessment tasks 'so as to focus on the central intentions of the curriculum and require the demonstration of lower order learning outcomes en passant' (Yorke & Knight, 2006, p. 572).

The word 'value' in Knight's (2000) article, 'The value of a programme-wide approach to assessment' hints at his vision of efficiency in assessment strategies. Knight contends that in the assessment of learning in higher education, 'the ideals of usability, low cost, reliability and validity interfere with each other', but that a programmatic approach can reduce the conflicts (p. 237). Knight (2000) proposes a comprehensive model which 'capitalises' on recognising the distinction between those learning outcomes that can and need to be reliably assessed, and those which cannot be assessed *reliably* but which can be assessed in ways that foster student learning and for which achievement can be demonstrated in other ways.[5]

Knight (2000) outlines in four points his strategy for redistributing resources away from reliability in some types of assessment.

1. To recognise that some things cannot be reliably assessed by faculty and therefore to devise alternative ways of making information available to stakeholders;

2. To invest in reliable assessments of what can be and needs to be reliable assessed (this will be called 'high stakes assessment');

3. To recognise that assessment in higher education is not primarily a tool with which to hew robust awards but an aid to student learning (low stakes assessment);

4. To use resources freed up by (3) to do (2). (p. 328-239)

Knight (2000) emphasises that 'these strategies only make operational sense if attention is concentrated upon assessment arrangements in complete degree programmes' (p. 239). A more detailed version of Knight's model is included as an appendix.

5.2.3   Inclusive assessment

Inclusive assessment seeks equity in assessment for 'conventional students' as well as late entrants, school leavers, international students, disabled students and students with families.

Much of the literature on inclusive assessment addresses the need for a programmatic approach that aids academic integration and retention in the first year, and especially in the first semester.

> From the different perspectives of both school-leavers and mature entrants, the first year of higher education (full-time or part-time) is critical to students' success: in the UK approximately two thirds of withdrawals take place during, or at the end of, the first year. (Yorke, 2001a, p. 116)

Yorke argues that summative assessment during the first year should be reconsidered 'since students may need longer than one semester to come to terms with the demands of study as well as other pressures, such as those of living away from home or accommodating the needs of dependents whilst studying' (Yorke, 2001a, p. 121). He argues,

> There is no need, at the end of the first semester, to require that failure at that early stage be redressed through the resitting of examinations or the reworking of

---

[5] Gibbs and Simpson (2004-05) echo this sentiment in their argument that 'we should design assessment, first to support worthwhile learning, and worry about reliability later' (p. 3).

assignments – these are simply extra burdens on students who are manifestly struggling to cope. (Yorke, 2001a, p. 121)

Alternatively,

> If institutions were to emphasise formative assessment at the end of the first semester of full-time study, instead of using summative assessments in a partially diagnostic way, students might feel more supported and not disheartened – with beneficial effect on both students and the institution (whose income is affected by non-completion). (Yorke, 2003, p. 490)

A focus on formative assessment and feedback and student development in the first year, 'is particularly important when students have entered higher education from backgrounds which have left them underprepared for institutional expectations' (Yorke, 2001a, p. 121).

An inclusive programme-level assessment strategy might also take into consideration the needs of international students. There is very little literature available on this specific topic, but Leask and Carroll (in press) suggest that 'fewer, more carefully selected and longer group work assignments which also provide (or draw upon) training and support for cross-cultural interaction' (n.p.) can help international students to integrate into UK academic communities. This may be particularly effective in the first year.

Adams and Brown (2006) write about how the designing assessment to be inclusive of people with different learning needs can benefit all students, not only those with disabilities. They argue that in order to achieve inclusive practice we should

> …stop adopting practices which predominantly focus on *adjustments* and start thinking about inclusive curriculum and assessment *design* which offer all students choices that align with their abilities. All students are likely to benefit from the flexibility in time, mode and place that is often seen as the basis for making reasonable adjustments' (emphasis mine) (p. 187).

Adams and Brown (2006) make a strong case that flexibility (as a rule rather than as an exception) is a key principle of inclusivity (see also Craddock and Mathias (2009) for an example of how building in assessment options can benefit students with learning disabilities). Flexibility or 'assessment empowerment' may also contribute to the sustainability of an assessment strategy (discussed below) because choice may 'engender intrinsic motivation' (Rust, 2002, p. 156).

5.2.4   Sustainable assessment

> *Existing assessment practices are perhaps the greatest influence inhibiting moves towards a learning society* (Boud, 2000, p. 155)*.*

David Boud (2000) identifies sustainable assessment as 'assessment that meets the needs of the present without compromising the ability of students to meet their own future learning needs' (p. 151). Boud and Falchikov (Boud, 2000; Boud & Falchikov, 2006) fear that assessment can send the wrong message to students, compromising them as life-long learners. Yorke and Knight (2006) share the same concern:

> If broader achievements are perceived by students as being of secondary importance (perhaps because they do not attract grades), then students will pay less attention to them than to those that the assessment regime seems to value more highly. Such an

approach might be criticized for its simplism and short-termism, but is best challenged by developing a learning culture in which a variety of achievements is visibly valued. (p. 581).

This idea was argued by Squires in 1990, who suggested that negative effects could result from 'overcrowded curricula which induce "surface" learning, methods of assessment which encourage "strategic" learning, curricular or teaching functions which have become dysfunctions, or the transmission of closed or compact attitudes towards continuing education' (Squires, 1990, p. 146, quoted in Knight & Yorke, 2003, p. 141).

The discourse of assessment can engender unsustainable assessment practice if it focuses on certification and constructs learners as passive subjects (Boud, 2007, p. 17). Boud and Falchikov (2007) support a move away from marks and grades (norm-referenced grading is particularly problematic) to 'a language of judgement that is both more transparent and more varied than the language of current assessment discourses', perhaps through portfolios or qualitative descriptions (p. 24). This move would require 'employers to understand the proposed changes and how these would affect them' (Elton, 2004, p. 420).

Although sustainable assessment is sometimes linked with life-long learning, not all examples of self-assessment fulfil the criteria for sustainable assessment purposes (Falchikov, 2005, p. 79). To promote life-long learning, self-assessment must support students in becoming 'autonomous professionals capable of appraising the quality of their work beyond the prescribed standards of the programme of study' rather than merely helping them to understand the gap between their work and demands of the programme (Tan, 2007, pp. 118-119).

Another aspect of the sustainability of an assessment strategy is economic. The Bradford group may want to consider whether any pilot projects they design can be sustained beyond their pilot periods. This is discussed (implicitly) in the sections of this literature review about efficiency and leadership/working culture.

5.3   Taking stock and design

5.3.1   Mapping the assessment at the programme level

A number of researchers writing about programme-level assessment make passing reference to the need to 'take stock' of what current assessment practices are in use (Bloxham & Boyd, 2007; Brown & Knight, 1994; Knight, 2001; Knight, 2002a). Bloxham and Boyd suggest a mapping system in which programme outcomes are mapped across modules and years in light of whether the outcome is being developed or assessed in that module. This allows programme teams to ensure that outcomes are developed over a few modules before being assessed summatively.

Although a number of researchers include sample grids in their publications (Bloxham & Boyd, 2007; Brown & Knight, 1994), only one article details the mapping process at length. [6] In their

---

[6] Gibbs and Dunbar-Goddet's (2007) approach to 'taking stock' of assessment environments (discussed throughout this literature review, see footnote 3) has the advantage of taking into account formative assessment, but is more suited to comparing programmes/disciplines/ institutions than informing our understating of how students experience assessment as they move though an individual programme.

2009 article, 'Mapping the maze of assessment', Gillett and Hammond (2009) review their experience of mapping the assessment tasks that students are expected to complete at different levels (year 0, 1, 2, 3 and masters) and in different subjects. They devised a grid structured around the *processes* involved in completing various assessment tasks rather than the end products. The grid comprised 22 'features' distributed through six 'key areas': **tasks** (multiple choice, open book, IT based, interactive, group element, role play); **medium** (oral, numeric, diagram/pictorial); **who assesses?** (self-assess, peer assess, self-set element); **cognitive** (analytic, evaluative, skills focus, primary research, theory focus); **time span** (reflective, process/periodic, portfolio); **work-related** (practice focus, case study).

The authors looked at the University's 'Definitive Module Documents' (module guides) and noted which of the 22 'features' applied to the assessment tasks described in the section 'assessment details' (as a result, they may have omitted a range of less formal, formative assessment). Gillett and Hammond were able to use this data to extract a number of graphs depicting patterns in assessment at the University, including 'percentage of modules in schools that involve a particular feature' (such as group work or reflective work) and 'percentage of modules (differentiated by level) that involve particular features'.

Programmes could potentially use Gillett and Hammond's grid to better understand the assessment tasks that are assigned to students as they progress through the programme. Although he discusses learning process in general rather than focusing on assessment, Knight's description of the implementation of the 'Skills Plus' project demonstrates how the mapping processes can be useful in designing curricula, including assessment systems:

> Coherence is assured by taking each module, choosing encounters that compatible with the material and then looking at the distribution of these encounters against the set of modules that make up the programme, or that part of it within a department or team's control. Departments in the North-west of England participating in the 'Skills plus' project to enhance student employability are using this strategy in 'low cost, high gain' approaches to enhancing established programmes. Simple audit forms allow departmental co-ordinators to map learning processes across a programme. This stocktaking shows that some encounters are rare, others are too common and others are clustered in, say, Year 1, but absent from the final year. It leads to collective and individual negotiations with tutors to get a better distribution of a good range of encounters (Knight, 2001, p. 376)

Knight (2001, end notes) cautions that due to the freedom allowed within modular schemes, 'it is not always possible to map out the encounters they will have'. In this case, he recommends looking at the 'combinations of modules that students regularly construct or are expected to construct', which may account for 50% or 80% of the degree'. In an earlier book, Brown and Knight suggest:

> This means that when planning an assessment system, the department has to make certain assumption about the model route though its courses. This may involve a department in making rules that students taking who are taking at least half of their courses within the department must take certain courses, not particularly because of the content of those courses, but because the have been designed to provide certain sorts of important assessment and learning opportunities. It will be appreciated that students who mix courses from several departments may well by-pass all plans for a [sic] coherence and progression in their assessment experiences. (Brown & Knight, 1994, p. 131)

However, if Gillett and Hammond's scheme were adopted and all modules were mapped, it might be possible for a student or academic adviser to quickly (with software) generate a personalised schedule of assessment processes for any student. An adaptation of the mapping process might pay greater attention to formative assessment and might also take into account students' experience of assessment in modules run concurrently, including bunching/log jamming.

5.3.1.1 Collective auditing and interpretation

Although mapping has the potential to be tailored to an individual's goals and selection of modules, mechanistic mapping takes an inherently impersonal approach to progression and it would it be unlikely to be able to address discipline-specific learning thresholds. In the American guide to 'assessing' (evaluating) programmes, *Program based review and assessment* (Stassen, Doherty & Poe, 2001) curriculum mapping is described as providing 'an inventory of the link between your objectives and the curriculum' (p. 22) and is seen as a *starting point* for a departmental (or programme) conversation.

> It can also serve as a catalyst for discussions about the proper sequencing of courses, the degree to which the curriculum really supports student learning, and the extent to which core objectives are appropriately addressed within the curriculum. Discussing the link between learning objectives and the curriculum may lead to a more general conversation about how processes within the major facilitate or hinder accomplishment of program goals. (p. 22)

The guide explains that this conversation might help the team to identify aspects of the programme that require special attention, such as 'key points in the curriculum where it is particularly important to gauge student progress' (p. 22).

Knight and Yorke (2003) argue that institutional, departmental and programme-level cultures will shape *how* a programme-level assessment strategy can be implemented, '...communication and the development of a shared learning, assessment and teaching culture are central issues in the design of departmental assessment systems' (p. 176). Brown and Knight (1994) argue that having departments collectively interpret such a map or grid can help a department or programme to develop shared understandings about teaching, learning and assessment. The benefits of having a department address assessment include:

> Making them consider their aims in some detail. Once aims are linked to assessment, some sharp thinking is needed, since invariably aims have to be refined, tightened, and perhaps reconceived as academics wrestle with the issue of what would constitute a fair test of the aim and a suitable display of performance.
> Making them trace which of their goals is advanced in which part of their programmes. This frequently exposes a highly uneven pattern of coverage, with some goals receiving mass attention, whereas others are distinguished by their rhetorical presence and practical absence.
> Forcing a collective consideration of teaching and learning methods, since assessment questions are teaching and learning questions
> Sharpening understanding of the place and purpose of individual modules and unites within the scheme of things, thereby making it possible to arrive at a more coherent view both of the programme and of the department. (p. 132)

Brown and Knight (1994) argue that collective interpretation can also help the department or programme to negotiate the meaning of its own goals.

> A benefit to a departmental approach to assessment is that it helps academics to become clearer about their colleagues' understanding of what matters. One facet of this is negotiating the meaning of the goals which have been set. If one goal is critical thinking, then what would be taken as sufficient evidence of critical thinking at this level? With a cross-curricular achievement such as this the answer will in part be worked out at university level. None the less, it also needs to be resolved at department level, for the goal takes on its full meaning within the disciplinary context. (p. 133)

Bringing their collective knowledge into the discussion, the programme team can make sense of goals, negotiate what counts as 'effective' and reconcile the competing purposes of assessment in light of the programme's specific context. In addition to the disciplinary/programme context (for example, does the programme feed into professional accreditation bodies? Is it linked to well-established subject-discipline traditions of progression? What are the existing learning intentions and teaching and learning practices), programme teams will also be able to consider institutional priorities (inquiry-based learning, retention) and broader institutional culture.[7]

Brown and Knight (1994) suggest involving students in the design of new assessment systems, 'Not only is this good politics, to give students some ownership, but it is pragmatic, since they will have invaluable points to contribute because of their perspective as the people who will be assessed in this way' (p. 131).

5.3.2   Considerations in designing programme-level assessment

Among other things, this literature seeks to answer the question: *What factors should we consider when designing programme-level assessment?* Researchers have identified a number of 'features' of assessment systems that programme teams may want to consider. Knight and Yorke (2003) list: the consistency of assessment with programme learning intentions, teaching and learning practices; communication; fitness for purpose; differentiation by task, purpose, level and audience; progression; and coherence (p. 173), and other researchers have identified two additional features, variety and log jamming of assessment. It should be noted that these are *considerations* in the design of programme-level assessment, *not necessarily 'principles' of good assessment strategies*. Individual programmes teams need to decide how and the degree to which each should be pursued, given their programme's specific context and objectives.

5.3.2.1 Consistency of assessment with programme learning intentions, teaching and learning practices (constructive alignment)

Bloxham and Boyd (2007) contend that the effectiveness of an assessment strategy can be affected by the degree to which it is consistent with learning and teaching practices. They

---

[7] In their study of programme assessment regimes, Gibbs and Dunbar-Goddet (2007) found that patterns of assessment characteristics 'are fairly consistent across the three disciplines [that they studied] within an institution,' suggesting that 'there are institutional assessment characteristics that are evident across disciplines'. Gibbs and Dunbar-Goddet illustrate how local traditions can influence assessment choices: 'In the post-1992 institution a proposal for a module with 100% assessment by examination would be frowned upon, while in the Oxbridge institution anything other than a very high proportion of marks being derived from examinations would require considerable negotiation'. The way that an institution's focus should shape its assessment system is discussed in the section 'effective assessment'.

illustrate with the example of a module taught using problem-based learning methods, arguing:

> If this same module is assessed by an unseen examination, students are unlikely to take a deep and open-minded approach to their exploration of the subject matter. Instead, they are likely to demand information on what they have to 'learn' for the examination. (pp. 169-170)

Consistency is an important consideration not only in terms of constructive alignment, but also in terms the implicit and explicit messages conveyed in and by assessment systems, as discussed below.

5.3.2.2 Communication

Knight (2002a) argues that we should 'explore assessment as complex systems of communication, as practices of sense-making and claim-making' (p. 285). Assessment needs to communicate to those outside the programme, such as graduate schools and future employers (for a critique of the way communication with employers is currently handled, including a discussion of programme specification and learning outcomes, see: Elton, 2004; Knight, 2006; Rust, 2007). As discussed in the section on sustainable assessment, assessment systems send students implicit messages regarding what is valued by the programme (Boud, 2000, p. 160), and assessment geared toward communicating with outside stakeholders can send messages about assessment which are not 'sustainable'.

Knight and Yorke (2003) argue that programmes may need to send more explicit messages to students about assessment. They argue that elaborate programme assessment plans can be 'paper systems' if students do not 'get the message' (p. 175).

> There are usually slippages between what is planned, what is created and what is understood. The implication is that a carefully planned assessment system may be invisible to students and, more seriously, have no practical effect on assessment, learning and teaching practices. Few higher education institutions are able to sustain the level of coordination needed to ensure that students generally believe that they are working within coherent programme assessment, learning and teaching arrangements. (Knight & Yorke, 2003, p. 175)

Knight and Yorke argue that to counter this, 'explanations of "how we assess round here and why" need to be full and frequent' (p. 181). One aspect of 'how and why we assess round here' that could be explained to students is the concept of *differentiation*, proposed by numerous researchers as a way of reconciling the conflicts between the various purposes and audiences of assessment.

5.3.2.3 Fitness for purpose & differentiation by task, purpose, level and audience

> *There needs to be differentiation amongst assessment methods because those we have grown used to when assessing propositional knowledge in humanities and social sciences, notably essay writing, are barely fit for the purpose of appraising practical competence (Knight & Yorke, 2003, p. 52).*

Bloxham and Boyd (2007, pp. 163-168) argue that in making sure assessment strategies are fit for purpose, it is important to question longstanding traditions of assessment rooted in subject disciplines. Knight and Yorke (2003) argue, 'The adequacy of any assessment method can only be assessed in relation to a purpose. Some methods are much better than other for particular

purposes' (p. 69). For example, 'a method which works well for a summative purpose may not create good enough feedback for formative purposes' (p. 69). Assessment should be differentiated by purpose because summative assessment cannot fulfil student progress toward complex learning. Likewise, assessment for the purpose of certification meets the needs of one audience, employers, but may not be useful as a diagnostic tool for teachers. In a differentiated assessment system, 'some learning outcomes [could be] brought under cumulative summative assessment regimes, whereas others would be handled by formative assessment arrangements' (Knight & Yorke, 2003, p. 53).

5.3.2.4 Variety

Some calling for programme level-assessment express concern that in modular schemes, 'students may be able, through an appropriate selection of modules, to duplicate types of assessment or to avoid some types of assessment completely' (Yorke, 1998, n.p.). This is problematic because in order to avoid certain types of assessment, students may select modules that do not fit well with their longer-term goals. Alternatively, in selecting the subjects they want to study, students may inadvertently expose themselves to a limited variety of assessment methods. In their discussion of evaluating the quality of modular geography curricula, Gardiner and Chalkley (1994) propose that modular curricula could be balanced in terms of methods of assessment rather than simply in terms of content, 'with restrictions on choice of modules being conditioned by restrictions on how many modules can be assessed by particular forms of coursework' (n.p.).

Notably, the concern that students in modular systems may experience a limited number of assessment tasks is rooted in the assumption that a wider variety of assessments is better than a smaller one (see discussion in Leedham, 2009). Although variety in assessment is reported to have a number of benefits, including promoting inclusivity and equity and supporting intrinsic motivation (Bloxham & Boyd, 2007, p. 166), it should be considered critically.[8] Bloxham and Boyd (2007) caution that 'a multiplicity of assessment techniques should be balanced with a recognition that students need to become familiar with different assessment tools in order to demonstrate their learning effectively' (p. 166). Students may struggle to prepare for the range of assessment tasks they encounter because the material available to students on how to prepare for assessment tasks tends to concentrate on conventional assessment tasks (Gillett & Hammond, 2009, p. 133). Bloxham and Boyd (2007) suggest that 'programme assessment strategies should provide a reasonable variety of assessment methods but schedule their introduction across a programme' (p. 166).

5.3.2.5 Bunching/log-jamming of assessment

Yorke (1998, n.p.) uses the term 'bunching' to describe the problem created when students are required to undertake a large amount of assessment during a short time frame. Brown and Knight (1994, p. 147) call this 'log-jamming', and note that it also causes problems for teachers (for example, with marking). While in some cases there may be a good reason for bunching (a challenge that makes up part of scaffolding, perhaps), it often occurs not for pedagogic reasons but because of limited options and a lack of coordination between modules. This can lead to anxiety in students, surface learning, and the neglect of formative work. Rust (2000, p. 126) believes that students sometimes select modules in order to balance assessment tasks so

---

[8] In a study by Gibbs and Dunbar-Goddet (2007), 'Commonly held assumptions about the benefits of frequency, variety, explicitness and alignment were found to be contradicted by the evidence' (p. 25). Their findings underscore the need to pursue the 'features' discussed in this section critically.

that they don't have all course-work or all exams in the same semester. He cites the advice in circulation among students '"Never do two modules with course-work projects in the same term"' (p. 126).[9] Students selecting modules so as to balance the assessment tasks that will be required of them are doing to work of creating a workable assessment profile, potentially at the expense of coherence in their course selection.

When designing a programme assessment strategy, it is possible to verify whether it is likely to lead to a 'bunching' of assessments at certain points in the year. Bloxham and Boyd (2007, p. 173) cite the example of a well-intentioned programme team who – in an effort to limit the number of summative assessments – inadvertently increased pressure on students (to promote attendance, teaching staff all scheduled their assessment points for the last week of the semester, leading to a severe bunching of high-stakes assessment).

5.3.2.6 Progression

> *The advent of unitized curricula means that opportunities for slow learning have to be deliberately planned at programme level, since the individual study units are likely to be too short in duration to permit it to happen, and hence for assessment to capture it. Indeed, unless deliberate effort has been made to encourage slow learning, there may be no locus at which its development can be formally recognised, save, perhaps, by means of a culminating 'capstone' project or dissertation. (Knight & Yorke, 2003, p. 140)*

Bloxham and Boyd (2007) point out that 'progression can be conceived of in different ways, for example, progression in relation to the subject discipline, in terms of becoming a more autonomous learner, or in terms of ethical reasoning' (p. 168). Yorke argues that the 'logic' of the subject discipline still shapes the design of programmes: 'most programmes in higher education are based on a set of general assumptions in which the subject discipline, rather than student development, is dominant' (2003, p. 492).

Knight and York (2003) argue that assessment tasks at the end of a programme should 'make different demands from those at the beginning', that is, 'the problems on which students work should be progressively less well defined and progressively more complicated', and students 'should be asked to transfer their understandings and skilled practices to problems that are progressively further removed from the context in which the learning originally took place' (p. 92). Progression can be ensured by dismantling 'scaffolding' as students move through a programme (Knight, 2001, p. 337). Scaffolding includes the degree of task structure; whether resources are at hand and easily manageable; framing; familiarity; and social resources (endnote Knight, 2001, p. 379).

In Knight's 2001 article on a 'process' model of programme-level assessment planning (discussed in 5.3.C 'Take an outcome-led approach' section), he contends that a scaffolding conception of progression is less psychologically and philosophically problematic then one which would envision progression as 'moving from one learning outcome statement to the next in the taxonomy' (Knight 2001, p. 337; see also Hussey & Smith, 2002, pp. 226-227). Rather than working backwards from outcomes, Knight and Yorke argue that we should see

---

[9] Of course, dodging group work or avoiding exams are only a few of the reasons that students may select inappropriate modules. In his opinion piece, Rust (2001) remarks that students may choose modules to be in the same class as friends or because they do not clash with other engagements such as part time work.

assessment systems as 'sets of arrangements that make certain sorts of outcome more likely' (2003, p. 176).

Knight and Yorke (2003) caution that 'development is rarely linear, people often have to go back to relearn things that they have not kept refreshed through practice' (p. 145). They also argue that a progressive curriculum needs to emphasise 'learning' rather than performance goals, so for certain outcomes students' starting places may be relevant (p. 149).

To the list of things that should progress over the years of a student's participation in university life, Rust (2002) adds participation in the design and choice of assessment tasks. He argues,

> Any consideration of a department's assessment strategies should surely therefore consider at least the inclusion of marking exercises in year one modules or units, as an important part of the students' skills development, and then possibly moving onto involving the students in the design and choice of assessment tasks, and the negotiation of assessment criteria. (p. 152)

Francis (2008) calls this 'assessment empowerment' (see article for research results on how students' receptivity to this responsibility develops over time).

5.3.2.7 Coherence

Knight and Yorke (2003, p. 186) call coherence the 'guiding principle' of a systemic approach to assessment. It is rarely neatly defined in the literature on programme-level assessment and is probably made up of a sound balance of all of the other features listed above (it seems particularly close to the idea of consistency).

Knight and Yorke (2003, p. 182-184) identify three ways in which programmes can achieve coherence: relying on the student to construct coherence; focusing on assessment (programme learning outcomes are the source of coherence); and starting with the subject area. In fact, in the design of an assessment plan it is possible to draw from all three ways, rather than simply choosing between them. The following section of this report outlines the way that programmes have and may organise assessment in order to help students construct coherence. In doing so, it informs the question asked at the beginning of this review: *What models for programme-level assessment have been implemented, and which ones are emerging?*

5.3.3   Models for achieving coherence

5.3.3.1 Support students in constructing coherence

As discussed above, it has been the norm in modular programmes to situate the burden of coherence with individual students.

> Many undergraduates have to create coherence out of curriculum disintegration . . . Arguably, creating coherence is itself good learning, but where some become sense-makers on a grand-scale, others are like flotsam in swirling waters. Since it is a necessary condition for expertise that learners deploy their fledgling achievements in different settings for different purposes, it follows that good curricula need to do more that permit happenstance to bring them to threshold achievements (Knight, 2001, p. 371)

Calls for a more programme-based approach come from the recognition that on their own, students may struggle to find modules that meet their learning needs in such a way as to allow them to progress to their maximum potential. Helping students to 'identify the themes and achievements that the programme promotes' can support students in finding coherence in modular programmes (Knight & Yorke, 2003, p. 182-184).

Another way in which students can be helped to construct coherence in a modular curriculum is through a final year culminating project, be it a dissertation or otherwise.

> The role of these activities is for students to make sense of the often disparate events in which they have taken part for purposes of learning, assessment and practice and to draw from them both portrayals of achievement and indication of how they have been developing their judgement over time (Boud & Falchikov, 2007, p. 194).

Although culminating projects provide for the consolidation of learning (Boud & Falchikov, 2007, p. 194), they do so only in retrospect, so do not help students to construct coherence *throughout* their modularised programme.

Patchwork texts (see Winter, 2003) and portfolios (see Klenowski, Askew & Carnell, 2006) and personal development planning (PDP) offer a less retrospective approach to helping students construct coherence. Personal development planning may increase coherence for students who do not take a traditional route though a programme (taking more than the minimum specified time to complete) by encouraging them 'to apply the programme-level perspective to their particular educational situation' (Yorke & Knight, 2006, p. 574).

Knight (2002a) argues that portfolios could also support the certification function of assessment:

> Employers could satisfy themselves about claims by asking for evidence from the portfolios that students would develop adjuncts to a programme-wide approach to learning and formative assessment. If higher education institutions also provide summaries of the key learning, teaching and assessment processes associated with each programme of studies – of their process standards – stakeholders would be more able to scrutinize claims and warrants alike. (p. 285)

Jackson and Ward (2004) make a similar claim for progress files, arguing that progress files help students 'represent learning and achievement for a world that requires people to understand and demonstrate their capacities for working with disciplinary and transdisciplinary knowledge' (p. 427).

5.3.3.2 Look to the subject discipline

As discussed above, before the advent of modular programmes coherence was normally seen as inherent in the subject area itself. In designing an assessment strategy based on a subject area in the (learning outcomes dominated) modular era, one would:

> Devise tasks that engage learners with important issues and which represent the subject's characteristic inquiry procedures. Consider how task performance might be judged and arrange assessments accordingly. Identify the outcomes that are therefore likely to be promoted, and write the programme the specification. Check it against suitable points of reference – subject benchmarks in England, government requirements (as in the case of teacher education, for example), and professional standards. (Knight & Yorke, 2003, p. 183)

Deakin University (Australia) developed an assessment portfolio linked to professional practice and accreditation. It combines a number of practices discussed elsewhere in this review.

- the use of assessment in the first year as a foundational tool to establish student study habits and skills;
- the evolution of assessment tasks by the fourth year to reflect the world of professional practice and to allow students to demonstrate their integration of knowledge and skills;
- the weighting of assessment tasks to indicate the value attached to particular tasks;
- the structured inclusion of group work;
- a concern for student and staff workloads;
- the recognition of student diversity, in particular the needs of off-campus and mature-age students;
- the matching of assessment tasks to professional accreditation requirements

(Palmer, 2004, p. 200).

The subject-discipline approach is probably more relevant in single-discipline or pre-professional programmes.

5.3.3.3 Take a learning outcome-led approach

As mentioned earlier, Bloxham and Boyd put aside their reservations about learning outcomes for the purposes of their book *Developing Effective Assessment in Higher Education* because they wanted it to be relevant within current trends. In the book, Bloxham and Boyd (2007, p. 159-174) offer six steps to designing a coherent assessment strategy based on learning outcomes.

1. Develop programme learning outcomes

2. Map learning outcomes across a programme

3. Make assessment fit for purpose

4. Provide for progression

5. Match assessment and learning activities

6. Identify constraints on assessment

They point out that if an aggregated approach to marking is taken (if students are able to pass modules with out passing all of the learning outcomes), the assessment system should require that students demonstrate these outcomes at some other point in the programme (Bloxham & Boyd, 2007, p. 174).

In his 2001 article 'Complexity and curriculum: a process approach to curriculum-making', Knight criticises the outcomes-led approach, which he describes as follows:

> In the outcomes model coherence comes though programme goals that are decomposed into statements of learning outcomes. These, in turn, are turned into criteria to describe different levels of end-of-programme achievement, such as a lower second, pass or first class degree and into level-related descriptors that specify model

achievements at different stages of the programme. Differentiation, progression and coherence all designed to the programme by apparently precise descriptions of what should be learnt. (p. 376)

Knight associates the outcome-led approach with 'rational curriculum planning' or RCP, which he defines as 'a systemic approach that begins with specifying goals, and proceeds to objectives, thence to curriculum, instruction, assessment of learning, evaluation and such revisions as are needed to make the system work better next time' (p. 372).[10] Knight agrees with earlier scholars (Stenhouse 1975 & Eisner 1985) that this approach is appropriate for training, 'a process designed to achieve convergence' but that it is not appropriate for education, 'a process stimulating divergence' (p. 373).

Knight's other criticisms are that complex learning 'is not reducible to precise statements predicting what the outcomes will be' and that that '"plans" are often rationalisations of what we have [already] done' (p. 373). RCP 'aims to map an elegant pathway from goals, to objectives, delivery, reception and so on', but this does not leave space for creativity, innovation or 'the undermind's unhurriable slow learning' (p. 374). Finally, an outcomes-led approach does not address 'what needs to be done to get to those outcomes' (p. 376). As an alternative, Knight suggests that curriculum planning should concentrate on programme learning *processes*—encounters, environments, and messages' (p. 378) and begin 'by imagining how to draw together the processes, encounters or engagements that make for good learning' (p. 375).

5.3.3.4 Key researchers' visions for programme-level assessment

In an article about using a programme approach to make assessment more efficient, Knight (2000) proposes that evidence of achievements not be sought until the later stages of an academic programme. Certain course-specific material might be assessed along the way, but, in general, the years leading up to the final year would be spent participating in activities and completing formative assessments that support students' progress towards objectives established by the programme's assessment criteria. The final year assessments would involve 'the application of understandings, skills and qualities to fresh problems in fresh situations in the domain'. Knight details the model in ten points (p. 246-247). This very useful model is included as an appendix.

In an opinion piece, Rust (2000) outlines his vision of a system of assessment for modular programmes 'in which the student takes responsibility for tracking and recording their learning' (p. 127). Rust envisions a system in which

> …summative assessment were to be 'disaggregated from individual modules – in other words, modules cease to be linked to credit – and students were made responsible for the accumulation of a portfolio of evidence that they have met the intended learning outcomes for the programme, and for getting this evidence assessed. (2000, p. 127)

Rust argues, 'Assessment would start to be seen by students as an integrated part of the learning process, clearly linked to learning outcomes, rather than just a series of hurdles to be jumped and forgotten' (2000, p. 128). In terms of progression,

---

[10] An example of a systemic, outcomes led approach is detailed on page 22 of Paul Watson's (2002) article about incorporating learning outcomes into an assessment strategy for a Construction Management course.

…this proposed system would allow much more for development over time. The student could be allowed to take more than one term – and possibly a whole course – to develop sufficiently to be able to produce satisfactory evidence of a particular outcome. (Rust, 2000, p. 129).

Notably, both Knight's (2000) and Rust's (2000) schemes share in common an argument for delaying summative assessment. Delaying summative assessment is one of the most frequently presented pieces of advice in the articles reviewed (Gibbs & Dunbar-Goddet, 2007; Gibbs & Dunbar-Goddet, 2009; Knight, 2000; Rust, 2000; Yorke, 2001a; Yorke, 2001b; Yorke, 2003), and is consistent with the Oxbridge model, which Gibbs and Dunbar-Goddet (2007) found to be associated with a range of positive learning responses (see earlier footnotes for discussion).

Writing in 1994, as modular schemes were still only being implemented in many universities, Brown and Knight suggested a less radical plan for minimising summative assessment:

> Many fear that because individual units or modules have to be discretely assessed, this necessity will mean that students will be assessed more. This does not have to be the case. Universities, such as Teesside, that have already modularised, set some assignments that integrate learning across modules and yet provide discrete elements of assessment linked to specific units. Where problems have been experienced, it has tended to be because course designers have been too ambitious in the scope of their planned assessment. (Brown & Knight, 1994, p. 148)

Although the Teesside strategy sounds innovative, there does not appear to have been any publications on the implementation or success of this strategy.

The models of programme-level assessment described above inform the question '*What constitutes programme-level assessment?'* already discussed in terms of what constitutes effective, efficient, inclusive and sustainable assessment in section 5.2. It is apparent that the spectrum of what can be considered a programme-level assessment strategy is wide; it ranges from the introduction of a culminating project in the final year, to the integration across a programme of a number of research-supported 'good practices' (such in the Deakin University Engineering programme), to a radically integrated programme wherein students are not assessed until learning outcomes have been revisited several times in different modules (as in the models envisioned by Rust (2000) and Knight (2000)). The following section, implementing change, informs the questions: *What enables the implementation of programme-level assessment and what are the potential challenges and obstacles to take into account?*

5.4   Implementing change

5.4.1   Leadership and institutional/departmental/programme cultures

Calls for programme-level assessment strategies are based on the concern that 'there is a risk that complex learning achievements, which need to be fostered across a programme, may not get sufficient attention when curriculum development is in the hands of individual course leaders' (Yorke & Knight, 2006, p. 572). In a separate article, Knight clarifies:

> However, it is not prudent to presume that curriculum coherence will be achieved by making heads of department responsible for it. Heads of department often have had little or no training as leaders, and work in situations where taking a lead in curriculum matters would be unexpected and perhaps unwelcome (Knight & Trowler 2001). Furthermore, even if they were well-placed to make a difference to the curriculum,

their effectiveness would be compromised by bad advice on what to do. (Knight, 2001, p. 372)

This raises the questions, expressed by Yorke in 1998, of 'where the responsibility lies for ensuring that a students' profile of assessment tasks matches what is expected, and where the superior responsibility lies for ensuring that units of the institution (co)operate in a manner consistent with operational aims' (Yorke, n.p.). That is, how can a programmatic approach to assessment be enforced?

Knight (2000, p. 239) acknowledges that any move toward a programme-level approach to assessment may 'impinge on views of academic freedom' among professors, but Yorke and Knight (2003, p. 213) insist that

> …taking a systemic approach to assessment need not be tantamount to 'hard managerialism', with its emphasis on 'command and control' strategies. 'Soft managerialism', with an emphasis on agreeing goals and then helping other to meet them in the ways they judge most fit for purpose and context, lies closer the notion of subsidiary. (p. 213)

Yorke and Knight (2006) suggest that to achieve programme goals, it is helpful to have 'a set of general expectations that guide programme development across the institution' (p. 573). That said, a set of general expectations will not be sufficient if the culture and structures of the institution or programme pushes them into modular 'silos'.

As Grangeat and Grey (2008) explain (writing about secondary school), a collaborative educational community is instrumental in bringing about pedagogical changes.

> The introduction of new pedagogical methods and techniques, such as inquiry-based science education, requires more collaboration within educational communities in order to share professional support and to bring individual practices into line with collective purposes. (p. 179)

The need 'to bring individual practices in line with collective purposes' is, of course, especially true of a programme-approach to assessment. Grangeat and Grey (2008), referring back to an conference paper by Grangeat and Munoz (2006), argue:

> Teachers' sense of agency also mediates their participation in collective activities: when they participate in collective projects they pay greater attention to the temporal (e.g. anticipation of students' trajectories) and spatial (e.g. coherence with other departments' actions) aspects of teaching. (Grangeat and Munoz, 2006, paraphrased in Grangeat and Grey, 2008, p. 180)

Grangeat and Munoz are not referring to programme-level assessment strategies in particular, but their emphasis on progression and coherence suggest that their discussion of the importance of collaboration and agency is also relevant for programme-level assessment systems.

5.4.2 Tuning

In order to increase the coherence of an established programme, Yorke and Knight suggest orchestrating what is already there, or, 'tuning'. This involves auditing existing modules in order to insure the appropriateness of the range of teaching and learning activities, the differentiation of assessment, the relative usage of different learning, teaching and assessment

methods and their distribution across the programme for support of progression (Knight & Yorke, 2003, p. 184). The approach is pragmatic in that it works 'with the grain of colleagues' professional understandings' and does not 'attempt to promote wholesale curricular change, which would entail extensive developmental work and require the implementation of the full institutional quality assurance process' (Yorke & Knight, 2006, p. 575).

Yorke and Knight (2006, p. 575) used 'tuning' to adapt a number of programmes' curricula so that they more strongly emphasised employability as part of the 'Skills plus' project. Although their model provides an example of non-invasive programme-level change, Yorke and Knight (2006) acknowledge that assessment was that aspect of the curriculum least 'tuneable'. They identified two difficulties of altering summative assessment within programmes:

> The first was the difficulty of making significant changes to summative assessments within the existing summative assessment regime, since there were both within-programme and across-programme implications if the summative assessment within a single module was altered. The second, and pedagogically more important, was the limitation places on formative assessment by a semesterized curriculum. (Yorke & Knight, 2006, p. 576)

'Tuning' offers a model for implementing change in established programmes but is not particularly well suited to changing assessment. Many programme teams may have to work within the constraints of already validated modules and programmes. Researchers do agree, however, that it is much easier to make changes to assessment if change is coordinated with validation. Bloxham and Boyd (2007) argue, 'The ideal time to consider the assessment strategy of a programme is when it is introduced or revised' (p. 158). Adams and Brown (2006, p. 187) advise that it is best to consider inclusivity at the beginning, perhaps by engaging disabled [and other] students in the debate that takes place in designing assessment instruments and curricula, rather than waiting and relying on ad hoc adjustments.

5.5   Evaluation

The final section of this literature review seeks to inform the question: *How can we evaluate the effectiveness of programme-level assessment?* The literature search revealed that there is very little research available on evaluating programme assessment strategies.[11] Some of the key researchers discussed throughout the review did touch briefly on the topic of evaluation.

In point ten of Knight's (2000) proposed assessment system (see appendix), he suggests that 'course reviews and inquiries about programme quality should ask about the learning and formative assessment processes (they are perhaps the same thing) as well as explore the security of the high stakes, summative assessments' (p. 249). Gibbs and Dunbar-Goddet (2007) offer a model of how to evaluate an assessment *environment* in terms student learning responses, but this may not be relevant for the evaluation of programme level assessment strategies.

Any evaluation strategy will necessarily need to taken into account the objectives of the programme level assessment, asking, *'was the assessment strategy effective, efficient, sustainable and inclusive?'* Whether the evaluation should be undertaken separately or as part of programme review/validation processes must be considered. Another question to ask is who the stakeholders are, and who should do the evaluation. Yorke (1998, n.p.) argues that

---

[11] A complicating factor is the language, e.g. 'programme evaluation of programme-level assessment'.

evaluation and feedback mechanisms regarding the management of assessment are necessary at the level of the institution, department, programme, programme team and individual.

As with the design and implementation of programme-level assessment, research suggests that evaluation may be more successful if faculty-driven. Brown and Knight (1994) argue that 'assessment programmes are most effective if academic staff are involved in their design, in their operation and in their review', and suggest that one reason why is that planning the evaluation causes staff to reflect on teaching and learning (p. 128). They point to Banta and Schneider's (1988) article about developing a comprehensive examination that would serve the purposes of programme evaluation the University of Tennessee, Knoxville. Banta and Schneider report:

> The *test development process* itself had an effect on departments even before the product was administered to students. A common feeling expressed by department heads was that that faculty were brought closer together in their thinking about the curriculum because they were forced to focus on common learning objectives for students. (p. 78)

This also led staff to establish 'a clear progression of courses from lower- to upper-division levels' (p. 78). One department head remarked that '"…it motivates faculty to correct weakness that they discover in the process"' (p. 82).

Brown and Knight (1994) emphasise that no matter how evaluation data is collected, 'data do not *prescribe* any one course of action. It [sic] provides imperfect data which are to be interpreted. Policy decisions are then negotiated around the interpretations' (p. 145). In short: 'Using assessment data is an art, a political art, in that it involves making value-led decisions' (Brown and Knight, 20004 p. 145).

6   Conclusion

Although many of the articles reviewed espouse specific 'best practices' such as delaying summative assessment, this literature review will not attempt to further distil a list of such practices because it would quickly become apparent that some practices were not realistic or compatible in many instances. Two truly overarching principles of best practice for the design, implementation and evaluation of programme-level assessment strategies are identified in the review. First is the importance of *collectively* interpreting the problem and designing the solution, maybe even including students – perhaps especially non-traditional students – from the beginning. The second overarching principle is *close attention to the programme's specific context and objectives* in agreeing goals and designing, implementing and evaluating the assessment system.

The literature review supports the Bradford PASS group in designing a 'tried and true workshop format which programme teams can use to review/revise their assessment strategies' (Bradford, 2009) by offering ideas for taking stock of what is already there and by listing different features of an assessment strategy to be considered in an inclusive discussion about how to meet the programme's specific objectives for the assessment strategy in light of the programme context.

Works Cited

Adams, M. & Brown, S. (2006) A manifesto for mainstreaming inclusive practice, in: M. Adams & S. Brown (Eds) *Towards inclusive learning in higher education: Developing curricula for disabled students* (Abingdon, Routledge), 187-189.

Banta, T.W. & Schneider, J.A. (1988) Using faculty-developed exit examinations to evaluate academic programs. *Journal of Higher Education,* 59(1), 69-83.

Bloxham, S. & Boyd, P. (2007) *Developing effective assessment in higher education: A practical guide* (Maidenhead, England, Open University Press).

Boud, D. (2000) Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education,* 22(2), 151-167.

Boud, D. (2007) Reframing assessment as if learning were important, in: D. Boud & N. Falchikov (Eds) *Rethinking assessment in higher education: Learning for the longer term* (Abingdon, Routledge), 14-25.

Boud, D. & Falchikov, N. (2006) Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education,* 31(4), 399-413.

Boud, D. & Falchikov, N. (2007) *Rethinking assessment in higher education: Learning for the longer term* (Abingdon, Routledge).

Bradford, U.o. (2009) *Programme assessment strategies (PASS)*. Available online at: http://www.pass.brad.ac.uk/ (accessed 18/02/2010).

Brown, S. & Knight, P. (1994) *Assessing learners in higher education* (London, Kogan Page).

Craddock, D. & Mathias, H. (2009) Assessment options in higher education. *Assessment & Evaluation in Higher Education,* 34(2), 127 - 140.

Elton, L. (2004) Should classification of the UK honours degree have a future? *Assessment & Evaluation in Higher Education,* 29(4), 415 - 422.

Falchikov, N. (2005) *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education* (Abingdon, UK, Routledge).

Francis, R.A. (2008) An investigation into the receptivity of undergraduate students to assessment empowerment. *Assessment & Evaluation in Higher Education,* 33(5), 547-557.

Gardiner, V. & Chalkley, B. (1994) Quality assessment and modular geography curricula. *Journal of Geography in Higher Education,* 18(2), 175-183.

Gibbs, G. & Dunbar-Goddet, H. (2007) *The effects of programme assessment environments on student learning.* Report for HEA (York).

Gibbs, G. & Dunbar-Goddet, H. (2009) Characterising programme-level assessment environments that support learning. *Assessment & Evaluation in Higher Education,* 34(4), 481-489.

Gibbs, G. & Simpson, C. (2004-05) Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education* 2004-05(1), 3-31.

Gillett, A. & Hammond, A. (2009) Mapping the maze of assessment: An investigation into practice. *Active Learning in Higher Education,* 10(2), 120-137.

Grangeat, M. & Gray, P. (2008) Teaching as a collective work: Analysis, current research and implications for teacher education. *Journal of Education for Teaching,* 34(3), 177-189.

Hussey, T. & Smith, P. (2002) The trouble with learning outcomes. *Active Learning in Higher Education,* 3(3), 220-233.

Jackson, N. & Ward, R. (2004) A fresh perspective on progress files: A way of representing complex learning and achievement in higher education. *Assessment & Evaluation in Higher Education,* 29(4), 423-449.

Klenowski, V., Askew, S. & Carnell, E. (2006) Portfolios for learning, assessment and professional development in higher education. *Assessment & Evaluation in Higher Education,* 31(3), 267-286.

Knight, P.T. (2000) The value of a programme-wide approach to assessment. *Assessment & Evaluation in Higher Education,* 25(3), 237-251.

Knight, P.T. (2001) Complexity and curriculum: A process approach to curriculum-making. *Teaching in Higher Education,* 6(3), 369-381.

Knight, P.T. (2002a) Summative assessment in higher education: Practices in disarray. *Studies in Higher Education,* 27(4), 275-286.

Knight, P.T. (2002b) The assessment of complex learning outcomes, paper presented at the *International conference on engineering education*, Manchester, August 18-21. Available online at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.13.3429&rep=rep1&type=pdf

Knight, P.T. (2006) The local practices of assessment. *Assessment & Evaluation in Higher Education,* 31(4), 435-452.

Knight, P.T. & Yorke, M. (2003) *Assessment, learning and employability* (Maidenhead, Society for Research into Higher Education & Open University Press).

Leask, B. & Carroll, J. (in press) Moving beyond 'wishing and hoping': Internationalisation and the student experience of inclusion and engagement. *Higher Education Research & Development*.

Leedham, M. (2009) From traditional essay to 'Ready Steady Cook' presentation: Reasons for innovative changes in assignments. *Active Learning in Higher Education,* 10(3), 191-206.

Palmer, S. (2004) Authenticity in assessment: reflecting undergraduate study and professional practice. *European Journal of Engineering Education,* 29(2), 193-202.

Rust, C. (2000) A possible student-centred assessment solution to some of the current problems of modular degree programmes (opinion piece). *Active Learning in Higher Education,* 1(2), 126-131.

Rust, C. (2002) The impact of assessment on student learning: How can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices? *Active Learning in Higher Education,* 3(2), 145-158.

Rust, C. (2007) Towards a scholarship of assessment (opinion piece). *Assessment & Evaluation in Higher Education,* 32(2), 229-237.

Stassen, M.L.A., Doherty, K. & Poe, M. (2001) *Program-based review and assessment: Tools and techniques for program improvement* (Amherst, Office of Academic Planning and Assessment).

Tan, K. (2007) Conceptions of self-assessment: What is needed for long-term learning?, in: D. Boud & N. Falchikov (Eds) *Rethinking assessment in higher education: Learning for the longer term* (Abingdon, Routledge), 114-127.

Watson, P. (2002) The role and integration of learning outcomes into the educational process. *Active Learning in Higher Education,* 3(3), 205-219.

Winter, R. (2003) Contextualizing the patchwork text: Addressing problems of coursework assessment in higher education *Innovations in Education & Teaching International,* 40(2), 112-122.

Yorke, M. (1998) The management of assessment in higher education. *Assessment & Evaluation in Higher Education,* 23(2), 101-116.

Yorke, M. (2001a) Formative assessment and its relevance to retention. *Higher Education Research & Development,* 20(2), 115-126.

Yorke, M. (2001b) Turn first-semester assessments into richer learning experiences. *Innovations in Education & Teaching International,* 38(3), 277-278.

Yorke, M. (2003) Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education,* 45(2), 477-501.

Yorke, M. & Knight, P. (2006) Curricula for economic and social gain. *Higher Education,* 51(4), 565-588.

Appendix A

This is a proposal for a system in which:

(1) Assessment would predominantly be formative, that is to say that it would be low-stakes, designed to give learners useful feedback on how to improve performance against programme-wide criteria. It would be embedded in the learning activities. Student participation in formative assessment would be a requirement for progress through the programme.

(2) The priority would be giving feedback that was useful, speedy and defensible. Reliability would come second to plausibility of judgement, because if a learner felt that a judgement was wrong, it would be important *in the interests of learning* for there to be open dialogue about that. This could help to reduce the incidence of the undesirable 'final language' of assessment and generally to reduce the negative emotions associated with the assessment of learning (Boud, 1995a).

(3) Each programme goal should be complemented by performance criteria written for different levels of achievement. There would also be course-specific criteria and they should mesh with the overarching programme criteria, although some course criteria would be specific to that course alone. Portfolios (see below) can be useful for students who wish to add their own programme-wide criteria to the official set.

(4) Students should have the programme criteria from the first, regularly use them, share them, and practise applying them, especially through dummy assessment exercises in which they grade anonymous samples of work from previous year groups.

(5) Peer- and self-assessment should be embedded in programmes. Both save faculty time and both help learners to become familiar with programme criteria, hence with programme values and goals. As for self-assessment, this is at the heart of the widely-commended learning practices of reflection and metacognitive thinking (Cowan, 1998; Harvey & Knight, 1996) and the ability to evaluate oneself is held to be intrinsically worthwhile (Boud, 1995b). Its high use value is matched by high exchange value in the employment market (Harvey *et al.*, 1996). Peer-assessment is a good preparation for self-assessment. There is also psychological evidence that the quality of discussion with peers is different from dialogue with faculty. It can enhance learning, since students often defend their positions more fully with peers. Feedback is in some respects better because it is usually discussed more and provokes more thought.

(6) Time for individual reflection (perhaps through a learning journal) *and* for collective stock-taking must be designed into programmes on the principle that reflection is not only a solitary activity (Cowan, 1998).

(7) Information and communication technology can support on-demand self-assessment that can provide feedback and even coaching on points of difficulty. It can also support both real-time and asynchronous peer-discussions, hence it can support peer-assessment.

(8) Authentic assessments become possible. The bugbear of authentic assessments has been getting reliability levels that are high enough for high-stakes purposes (Birenbaum, 1996;

Messick, 1999). When assessments are low-stakes and the main intention is to promote learning dialogues that inform future work, reliability is not the issue.

(9) Students will wish to claim achievements that summative tests do not, or could not warrant. Programme-wide goals will include the development of skills and qualities that faculty cannot reliably assess because it would be too expensive, ethically dubious, or involve situations where their very presence would change what happened. Much rather unproductive thought has been given to ways of getting reliable assessments of such goals. It is simpler to cut the knot by saying that it is practically impossible for faculty to provide useful and reliable judgements in some areas. Their responsibility is to ensure that programmes contain plenty of learning experiences that should stimulate development in those areas. Students should

know that their transcripts will say that these are important areas which have been repeatedly visited in the degree programme and in which graduates should be able to demonstrate achievement. However, faculty will not make summative judgements of achievement evidence in these areas. They will tell students to collect evidence of achievements in these areas and include it in their learning portfolios (Knight, 1995; Wright & Knight, 1999); to benchmark that evidence against programme criteria statements; to test out the plausibility of their judgements through formative conversations with peers (and with faculty as a part of the programme's student advisement arrangements); and to make claims to achievement to employers and graduate schools that they can confidently substantiate in letters of application and interviews.

Dalhousie University in Nova Scotia has an interesting credit-bearing suite of courses that prepare students to create learning portfolios that help them to identify achievements and directions for development against programme-wide and self-set goals. Evidence can come from programme learning activities; from other university activities such as work in societies and voluntary groups; from activities in the home community; and from part-time employment. In this way, and with academic support, students make rounded self-assessments of achievements that faculty could not evaluate reliably or ethically and certainly not with such a breadth of evidence.

(10) Course reviews and inquiries about programme quality should ask about the learning and formative assessment processes (they are perhaps the same thing) as well as explore the security of the high stakes, summative assessments. As suggested in Harvey and Knight (1996), these audit reports are an important way of demonstrating the quality of a department's work.