Chapter 1

Scaling up assessment for learning: Progress and prospects

David Carless

Abstract A definition of assessment for learning (AfL) is provided. From a synthesis of relevant literature, I outline four main AfL strategies: productive assessment task design; effective feedback processes; the development of student understanding of quality; and activities where students make judgments. I explore the notion of scaling up in relation to spread, depth, sustainability and shifts in ownership. Then I present a rationale for the scaling up of AfL following from dissatisfaction with current practices and persuasive research evidence on practices congruent with AfL. I relate the notion of scaling up to the geographical spread of AfL research activity; its somewhat modest impact on university assessment policies; and in relation to the expansion of feedback research. I then consider what conditions might facilitate deeper and broader implementation of AfL, including the role of quality assurance; the importance of leadership and incentives; the development of assessment literacy through professional development activities; and the potential of technology to act as a lever for enabling AfL strategies.

Introduction

Assessment for learning (AfL) is now reasonably well-entrenched as part of higher education (HE) pedagogy. It is well-recognized that assessment is a crucial driver of student learning and that well-implemented assessment processes provide positive prospects for meaningful learning, whereas flawed assessment risks leading student learning in unproductive directions. There has been a wide range of research activity and projects in HE influenced explicitly or implicitly by AfL principles over the last twenty years or so. There is also a rapidly expanding related literature, including various book length treatments (e.g. Carless, 2015; Knight, 1995; Sambell, McDowell & Montgomery, 2013). This range of evidence and reports of practice provide tentative indication that AfL has reached a stage of maturity. In the terminology of educational change, it seems to have become institutionalized (Fullan, 2001) in that the practices are embedded within the pedagogy of a wide number of teachers in HE.

It is, however, difficult to gauge precisely the extent to which interest in AfL has led to widespread implementation at course levels (Boud, 2014). Despite the arguments

for AfL, there remain powerful imperatives surrounding summative assessment and grading which risk overpowering learning-oriented approaches to assessment. These include: fairness and reliability of grading; grade inflation and honours classification; and student malpractices, such as plagiarism or other forms of cheating. Middle and senior managers are usually pre-occupied with quality assurance aspects of assessment, including preventing and managing malpractice rather than encouraging diverse or innovative approaches to assessment (Meyer et al., 2010).

The main aims of this opening chapter are to make a case for scaling up AfL; discuss the extent of implementation of AfL over time and across geographical locations; and frame the collection by charting some key issues in relation to the potentials and challenges for scaling up of AfL. I develop the arguments in the following stages. First, I define what AfL is and synthesize its main implementation strategies. Next, I propose a framework for scaling up and propose key rationales for the scaling up of AfL. I analyse the breadth and depth of AfL implementation through a discussion of AfL research and development in different contexts; its modest but increasing impact on university assessment policies; and in relation to the key issue of feedback. I conclude with a discussion of drivers and factors impinging on the scaling up of AfL and analyze some of the barriers arising.

Afl and Its Main Strategies

At the outset it is important to define what AfL is. There are various terminologies associated with approaches to assessment focused on enhancing student learning: formative assessment; assessment for learning; assessment as learning; and learning-oriented assessment. The term AfL came into common parlance in the early 2000s to emphasize the purpose for which assessment is carried out in contrast to formative and summative assessment which relate to the functions which are served (Wiliam, 2011). In the HE literature, AfL is often not defined explicitly. Accordingly, I adopt the following definition from the literature related to schooling:

> Assessment for learning is any assessment for which the first priority in its design and practice is to serve the purpose of promoting students' learning (Black et al., 2004, p. 10)

In relation to the school sector, the King's College London group led by Black and Wiliam did much to promote and encourage scaling up of AfL practices and I turn now to a discussion of the main AfL strategies for schooling and in HE.

AfL strategies in relation to schooling seem somewhat more clearly defined and agreed upon than those in HE. The following list of five key strategies (Wiliam & Thompson, 2008) is relatively authoritative:

1. Clarifying learning intentions and success criteria

2. Engineering effective questioning and classroom discussions
3. Providing feedback that moves learners forward
4. Activating students as owners of learning
5. Activating students as instructional resources for one another.

The AfL in HE literature carries some resonance with these strategies. Only the second of these is under-explored with a need for further investigation of apprenticing undergraduates into academic discourse through sensitive challenge and induction into academic practices (Black & McCormick, 2010). The relevant HE literature to date seems to lack a definitive statement of key AfL strategies and in order to trace the development of the ideas, I first discuss three key perspectives to help me work towards a synthesis.

In a review of conditions under which assessment supports student learning, Gibbs (2006) elaborates a number of issues in relation to the design of assessment and the development of effective feedback processes. He suggests that assessment tasks should capture student time and effort; distribute this effort evenly over the duration of a course; and engage students in productive learning activity. He considers a number of issues in relation to feedback, including its frequency and timeliness; linkages with assessment criteria; and the impact of feedback on student future learning (Gibbs, 2006).

In a vision of assessment reform, Boud et al., (2010) make a number of points relevant to the current discussion: assessment should engage students in learning that is productive; feedback needs to be used actively to improve student learning; students and teachers should become responsible partners in learning and assessment; and AfL should be placed at the centre of course design.

Sambell et al., (2013) suggest six features of AfL: appropriate balance of summative and formative assessment; authentic complex assessment tasks; self-evaluation activities; rich in informal feedback; rich in formal feedback; and offering confidence-building opportunities and practice.

Synthesizing these works and other relevant literature, Table 1 below summarizes on the left what I see as the main AfL strategies and on the right hand side of the table suggests some means of operationalizing them. These examples of implementation processes are illustrative and not intended to be exhaustive. Technology can act as an enabler of AfL and the fourth example in each category is an online or technology-related strategy.

The first strategy is productive assessment task design: the development of tasks which carry potential to stimulate meaningful learning processes amongst students. This includes designing tasks which encourage students to sustain deep approaches to learning aligned with the learning outcomes sought. This kind of assessment may often mirror real-life elements of the discipline. For example, Glofcheski (this volume) discusses assessment in Law focused on authentic assessment which facilitates a wide range of learning outcomes relevant to future professional life.

The second strategy is represented by the development of effective processes as a central factor in curriculum and assessment planning. A trend in recent work is to examine how feedback designs can promote student uptake of feedback (e.g. Boud & Molloy, 2013). This can involve, for example, the integration of guidance and feedback; and emphasis on students seeking, generating and using feedback. Moscrop and Beaumont (this volume) illustrate the potential of technology, such as a Learning Coach via an Intelligent Tutoring System to enhance feedback dialogues and scaffold student self-regulated learning.

The third strategy relates to student understanding of the nature of quality work and its relationship with transparent criteria or rubrics. A key role of the teacher is to support students in developing capacities to discern quality and make sound evaluative judgments (Sadler, 2010). Dialogue around exemplars, for example, contributes to the development of student expertise in making judgments. Students' enthusiasm for exemplars is a key sub-theme in the analysis of students' experiences of assessment (Carless, this volume).

The fourth strategy follows from the third in that it focuses on making judgments about quality in relation to the work of a peer or one's own work in progress. Giving peer feedback is often even more beneficial than receiving comments because it is more cognitively-engaging: involving higher-order processes, such as diagnosing problems and suggesting solutions (Nicol, Thomson & Breslin, 2014). Peer review processes also help students to calibrate their own judgments and enhance their own self-evaluative capacities.

Table 1. Synthesis of main AfL strategies and processes

AfL strategies	Illustrative implementation processes
Productive assessment task design	Alignment with intended learning outcomes Authentic assessment Integrated and coherent assessment Collaborative writing through wikis

Effective feedback processes	Integrated guidance and feedback Students generating and seeking feedback Closing feedback loops Technology-enabled feedback dialogues
Developing student understanding of the nature of quality	Students generating and/or decoding criteria Applying criteria Analyzing and discussing exemplars Online dialogue about exemplars
Students practising making judgments	Providing peer feedback Receiving peer feedback Self-monitoring work in progress Online facilitation of peer interaction

AfL involves partnership between teachers and learners. Assessment task design is largely in the hands of the teacher but is interpreted, and responded to, by students. Effective feedback processes can be facilitated by teachers but it is only students who can act on feedback. Understanding quality and making judgments also place the student at the centre of their learning with the teacher playing an important guiding and facilitating role. The central role of students in AfL is an undercurrent throughout the volume and a particular focus of the research reported in the chapters by Carless and Jessop.

Scaling up Educational Change

In the terminology of educational change, what are key issues in relation to scaling up of AfL? A starting point for scale relates to quantity: the number of teachers and institutions which are carrying out a specific pedagogic strategy or innovation. A more comprehensive conceptualization of scale comprises four interrelated dimensions: spread, depth, sustainability and shifts in ownership (Coburn, 2003). Spread involves implementation of pedagogic innovation at additional sites or in more groups within existing sites. Depth involves refining pedagogic practice in deep and meaningful ways that influence student learning. Depth also needs to impact the beliefs of teachers and their underlying assumptions about pedagogy (Kezar, 2011). Sustainability relates to longevity, requiring policy and infrastructure systems in place

to support continued improvement in pedagogy over time with potential transfers of ownership to encourage continuous refinement and further scaling-up (Coburn, 2003).

In relation to AfL in schools, Wiliam (2007) suggests that teacher communities of practice are a productive strategy for scaling up. He has developed five scaling up principles which carry potential wider relevance (Leahy & Wiliam, 2012, Wiliam, 2007). First, gradualism in that generally teachers take small incremental steps in implementing change. Second, flexibility is required in order to facilitate teacher adjustment to techniques to make them work in their context. Third, there needs to be a degree of choice so as to enable teachers to select which AfL techniques they are going to implement. Fourth, a certain amount of accountability is desirable so that teachers are accountable to the teacher learning community for implementing changes. Fifth, support occurs through the building of trust amongst participants in the learning community (Wiliam, 2007; Leahy & Wiliam, 2012). All of these issues seem to resonate with HE, including the fourth principle of accountability carrying additional quality assurance dimensions discussed later in the chapter.

Why do We Need to Scale up AfL?

Two key elements of a case for scaling up AfL are dissatisfaction with existing assessment practices; and research evidence suggesting the power of well-implemented AfL strategies. I discuss these in turn below.

First, there has been considerable airing over the last twenty years or so of dissatisfaction from both teachers and students about assessment practices. From the staff perspective, assessment is sometimes seen as a pernicious influence on the learning process: tending to direct students towards grades and instrumentalism rather than a wider learning experience; emphasizing summative assessment to the detriment of more formative approaches; failing to encourage the higher order learning outcomes to which university education aspires; and seen as time-consuming and implicated in unwelcome auditing and quality assurance procedures.

From a student perspective, there is plenty of evidence from institutional surveys both in the UK and other parts of the world that assessment is one of the least satisfying aspects or their student experience. Students' concerns include fairness; lack of clarity about what they are expected to achieve; disappointment if marks do not meet their expectations; emotional challenges, such as pressure, anxiety and discouraging experiences; and concerns about feedback processes, particularly their timeliness and usefulness.

Whilst it cannot be assumed that all of these staff and students' concerns are fully justified and reasonable, they are suggestive of considerable misgivings about aspects of how assessment is currently organized and implemented. These challenges are

compounded by relatively low assessment literacy of staff and students (Norton, Norton & Shannon, 2013; Price et al., 2012). A recent paper (Bevitt, 2015) sums up well a number of imperatives for assessment change: to enhance the student experience; to harness technological developments; to encourage AfL; and to respond to the needs of increasingly diverse student populations in the context of massified HE.

Second, there is a range of research evidence which indicates that approaches associated with AfL are powerful means of enhancing student learning. The landmark Black and Wiliam (1998) research synthesis captured attention by accumulating evidence that well-implemented formative assessment improves student performance in schooling and in HE across a variety of contexts and settings.

The influential meta-analysis of meta-analyses (Hattie, 2009) indicates the visible learning attributed to practices congruent with AfL: student self-evaluation and meta-cognitive strategies; formative evaluation and feedback; and collaborative learning through reciprocal teaching. Of the 138 practices reviewed by Hattie, many of the most effective practices resonate with AfL. Self-report grades (ranked no. 1) and meta-cognitive strategies (no. 13) involve students making judgments; providing formative evaluation (no. 3) and feedback (no. 10) are closely aligned with effective feedback processes; and reciprocal teaching (no. 9) shares facets with peer review and peer feedback.

To sum up, dissatisfaction with existing assessment practices and the research evidence in favour of AfL strategies provide a rationale for in-depth, sustainable attempts at encouraging and supporting more widespread implementation of AfL practices.

Breadth and Depth of Influence of AfL

The next section attempts the difficult task of gauging the breadth and depth of implementation of AfL. I review three possible indicators: geographical spread of research activity; influence on university assessment policies; and the expansion of academic attention to the key AfL strategy of effective feedback processes.

Geographical Spread

A key scaling up factor is geographical spread, the extent to which there appears to be deep and sustained AfL activity in multiple significant settings. As it is impossible to gauge how teaching, learning and assessment are implemented around the world, I discuss research and development activity as one of the indicators of scaling up of AfL. There are a number of examples of positive sustained implementation of practices congruent with AfL in selected international settings.

The UK seems to be a leading context for the implementation of AfL concepts. For example, Oxford Brookes University and Northumbria University both achieved prestigious Centre for Excellence in Teaching and Learning funding for sustained good work related to AfL. Colleagues from both of these institutions have been prominent figures in the AfL-related literature since the 1990s: Margaret Price and Chris Rust at Brookes; Liz McDowell and Kay Sambell at Northumbria. Research from the UK is also well-represented in this volume (see chapters by Jessop, Pitt, Moscrop & Beaumont).

Turning to Australia, there appears to be plenty of research activity in relation to the strategies discussed in Table 1. Two eminent scholars, David Boud and Royce Sadler have produced sustained research on AfL in HE over a period of more than 30 years. Boud's contributions include his early work on self-assessment (Boud, 1995); his equally influential championing of sustainable assessment for lifelong learning (Boud, 2000); and his analysis of the implications for assessment of increasing focus on learning outcomes and standards-based approaches (Boud, this volume). Sadler's work includes his seminal paper on formative assessment (Sadler, 1989); and his analysis of feedback in relation to the development of student understanding of quality (Sadler, 2010).

There is also a continental European school of AfL research, a significant strand of it stimulated by sustained work involving Filip Dochy and his collaborators (e.g. Dochy, Segers & Sluijsmans, 1999; Dochy et al., 2007). A parallel European development is an expanding literature related to the key AfL concepts introduced in table 1. Assessment design issues are investigated, for example, in relation to the assessment of professional competencies (van der Vleuten & Schuwirth, 2005) and a framework for quality assessment in competence-based education (Baartman et al., 2007). Analysis of effective feedback processes includes the interactive tutoring feedback model (Narciss, this volume); and feedback in online environments (Alvarez, Espasa & Guasch, 2012). Developing student understanding of the nature of quality can be facilitated by well-designed and well-used rubrics (e.g. Jonsson & Panadero, this volume). Peer assessment is one of the most favoured means of enabling students to practice making judgments (e.g. Strijbos, Narciss & Dunnebier, 2010).

In sum, this brief geographically-based synopsis is suggestive of considerable research and development work congruent with AfL in certain key settings but little known about various other parts of the world. In China, for example, the long history of competitive examinations represents a challenge to AfL, although there are some nascent initiatives to introduce a more formative orientation to assessment at the university level (Chen et al., 2013).

Assessment Policy Documents

A further indicator of how deeply AfL might be embedded within the fabric of HE pedagogy arises from an examination of university assessment policy documents which are generally readily available on university websites. An earlier synopsis of assessment policies in the UK and Australia (Boud, 2007) found that quality assurance aspects of assessment were predominant. A study of assessment policies in New Zealand (Meyer et al., 2010) reinforces this picture, suggesting that discussion of AfL is largely lacking in institutional policy documentation. In order to scrutinize these findings further, I have undertaken some preliminary analysis of assessment policy documents at a number of universities.

At King's College London, the home of formative assessment research in schooling, the assessment policy document focuses particularly on marking frameworks: different models of marking, including procedures for blind double marking; and conversion of marks from study abroad. The King's feedback policy emphasizes timeliness and the return of feedback within four weeks. Potentially more illuminating from an AfL point of view is a parallel King's Education strategy (2013-2016) which highlights student dissatisfaction with assessment and feedback. The document sets out an aim to reduce the burden of assessment through a more considered, flexible AfL regime. It outlines a number of assessment reviews being conducted in various disciplines, including mapping the student assessment journey; reviewing assessment at program levels; ensuring that assessments are calibrated to encourage lessons from one assessment to be applied to the next; and considering new forms of synoptic assessment above and beyond modules.

The University of Melbourne Assessment Procedure document focuses on 17 procedures including, the operation of Boards of Examiners, compliance, penalties, release of results, supplementary assessment and appeals. There is an additional Coursework Assessment policy, suggesting that assessment should be balanced so as to provide diagnostic, timely and meaningful formative feedback, as well as summative judgments. There is some reference to feedback which is viewed as involving comments indicating to students how they have performed against assessment criteria and how they can improve their performance. Discussion of feedback includes warnings against students communicating with examiners and how students may request access to examination scripts.

The University of Bristol assessment policy documentation contains both procedural and AfL elements. The main regulations and code of practice document focuses on progression, awards and the conduct of assessment in a roughly similar way to the King's and Melbourne procedures. More pertinent from an AfL perspective are the institutional principles for assessment and feedback which highlight the promotion of effective student learning. They include a principle that all assessment is for learning and suggestions for: a range of assessment methods; assessment mapping; the

imaginative design of assessment and feedback; research-informed practices; and encouragement for staff to improve their assessment and feedback literacy. Assessment and feedback are viewed as a conversation which provides students with opportunities to engage in continuing dialogues about their learning.

This brief and selective review of assessment policies in three major universities provides tentative support for the positions of Boud (2007) and Meyer et al., (2010) that university assessment policies generally emphasize rules and procedures, and an emphasis on quality assurance aspects of assessment, such as grading and moderation procedures. There is evidence in some of the policy statements, particularly at the University of Bristol, of thoughtful treatment of AfL elements, such as feedback. I turn next to examine how and why feedback has generated considerable recent attention.

An AfL Priority Area: The Case of Feedback

As a case of scaling up of research and development interest in an AfL area, I now analyze how in the space of twenty years feedback processes for students have gone from being a neglected research niche to a relatively high profile topic. In the late 1990s, feedback was an under-researched area (Higgins, Hartley & Skelton, 2002), yet there has been a remarkable growth in articles focused on feedback in the last 15 years or so. For example, in the main journal of our sub-field, *Assessment and Evaluation in Higher Education*, during the ten-year period 1996-2005 there were 11 articles with feedback in the title (3% of the total articles), whereas from 2006-2015 there were 65 (representing 11% of the articles). Of these 76 articles, 38 appeared in the three year period (2013-2015).

Probably the most urgent and persuasive driver for the expansion of feedback research is the consistent finding in National Student Surveys in England and Wales that feedback is perceived by students as one of the least satisfactory elements of their university experience (HEFCE, 2014; Williams & Kane, 2009). Jessop (this volume), for example, reports students' perceptions of episodic and haphazard feedback not connected to the next task or across modules. Student misgivings about feedback are also reported in other jurisdictions: Australia (e.g. ACER, 2010) and Hong Kong (Carless, 2006), so it seems to represent a widespread challenge. In the UK, the 'feedback issue' generated considerable media attention and was firmly on the radar of university senior management (Williams & Kane, 2009). This attention generated a host of initiatives designed to tackle the perceived problems. Many of these, such as focusing on feedback turn-around times (exemplified by the King's College feedback stipulations alluded to above) tend to be seen as 'quick-fixes' rather than more considered scholarly attempts at reforming feedback processes.

A repercussion was that numerous funded projects on feedback were spawned. A well-known example is REAP (Re-engineering Assessment practices) which was well-anchored conceptually in relation to the aspiration to promote self-regulated learning (Nicol & Macfarlane-Dick, 2006); and in relation to AfL strategies e.g. developing student understanding of quality and students making judgments (Nicol, Thomson & Breslin, 2014). A scaling up element of REAP involved linkages to strategic institutional developments. First, a new university policy for assessment and feedback consolidated REAP principles at an institutional level (Nicol & Draper, 2009). This is important because it moves from the potentially ephemeral work of a project to a more long-term legacy. The second institution-wide initiative was a 'feedback is a dialogue campaign' in association with the Student Union. This brought in different stakeholders through separate campaigns for staff and students on feedback principles and practices supported by leaflets and posters of advice. The extent of long-term strategic commitment to REAP ideas from senior management was, however, a moot point. Once funding dries up or key personnel depart, it is often difficult to sustain project legacies representing a barrier to scaling up.

A potentially productive mini-trend in relation to scaling up good practices in feedback is that a number of British universities have now instigated feedback awards to reward and encourage good practices. Some of these feedback awards are student-initiated or developed. The processes of these awards can surface and celebrate good practice (cf. Hounsell & Zou, this volume). Award schemes stimulate attention to a specific topic, providing rewards and incentives which can encourage the scaling up of good practice.

A further repercussion of increased attention to feedback processes is the expansion of previously under-explored research sub-strands, such as students' affective responses to feedback. The complex interplay between emotions and feedback is analyzed in two chapters in this volume. On the basis of phenomenographic research, Pitt (this volume) uncovered a range of emotional reactions to feedback, some of which were maladaptive and hindered action on feedback. Through synthesizing a wide range of literature, Rowe (this volume) brings out some of the complexities of emotions in relation to feedback and illustrates how a deeper understanding of emotions can play a role in the scaling up of AfL practices. The emotional legacy of feedback is also discussed in the chapter by Ajjawi and colleagues.

To sum up, I am suggesting that the scaling up of attention to the AfL element of feedback arose largely from student survey data which indicated dissatisfaction. This caught the attention of various stakeholders, including senior management, middle managers and staff of various levels leading to a wide range of research and development initiatives.

Drivers and Challenges in Scaling up AfL

I now discuss some drivers which might encourage wider and deeper implementation of AfL. I also consider some of the facilitating and inhibiting factors impacting on possible scaling up.

An inference I draw from the case of feedback is that a potential driver for assessment reform is evidence from quality assurance and quality enhancement processes. Although the continuous auditing agenda has its drawbacks, it may bring to light practices which are unpopular with students or do not stand up to quality assurance scrutiny. For example, program reviews, stakeholder feedback or external examiner reports may identify sub-optimal practices, and these can provide opportunities for middle management overseeing teaching and learning to follow up with action plans. Embedding the improvement of assessment and feedback within quality assurance processes is a key feature of the chapter by Jessop (this volume).

Leadership, especially at middle management levels such as Deans, Associate Deans, Heads of Department and Programme leaders, is a potential lever for assessment change. Commitment from leaders to an AfL agenda might support the scaling up of related practices. Staff involvement is rarely sustained without visible support from institutional leaders. Middle management might develop strategies to encourage AfL, including prioritization of resources, rewards and incentives. Congruent with the expansion of teaching award schemes to include feedback awards there could be similar additions of awards for best AfL practice or best assessment innovation.

There is a danger that adjustments arising from quality assurance or the priorities of academic leaders may reflect conservative approaches rather than AfL. It is important for institutions to develop climates where innovation in assessment is encouraged. The role of trust, or at least minimizing distrust, is a central issue in the encouragement of assessment reform (Carless, 2009). Trust would probably be more forthcoming if there were higher levels of staff assessment literacy and I turn to this issue next.

The development of staff assessment literacy carries potential to contribute to the scaling up of AfL. Building on the AfL strategies summarized earlier in table 1, I suggest that teacher AfL literacy involves a sound grasp of principles and practices in assessment task design; effective feedback designs; and developing student capacities in understanding and applying criteria through making judgments. The development of assessment literacy resonates with the scaling up concept of teachers being better able to respond to contextual challenges when they possess deep understandings of pedagogical principles. Assessment literacy would enhance teachers' capacities to adapt AfL practices to the needs of their students in particular institutional and disciplinary settings. Discipline-specific implementation of AfL practices is

well-represented in this volume, including Dentistry (Bridges et al.), Health professions (Ajjawi et al.) and Law (Glofcheski).

Professional development activities are an obvious starting point for enhancing staff assessment literacy. Seminars and sharing sessions can be useful in exemplifying and disseminating good assessment practice which in the hands of enthusiasts may play a role in scaling up. Encouraging good practice is admirable, but what about reducing bad practice in assessment? Perhaps the most promising strategy to tackle this difficult issue is leadership, allied with a judicious balance of pressure and support. As suggested earlier, quality assurance might be used as a lever to tackle and reduce unsophisticated assessment practices. Mentoring and peer review of assessment practice is also worth scaling up.

An alternative or possibly complementary means of developing assessment literacy is through communities of practice in which 'accounts of practice' are surfaced and shared (Hounsell & Zou, this volume). Such activities resonate with the earlier discussion of communities of practice in schools as a means of enhancing ownership of AfL strategies. Supportive professional communities of practice facilitate collegial support and promote sustainability (Coburn, 2003). In such ways, AfL strategies could be shared, developed and refined amongst groups of colleagues.

Another form of a community of practice is represented by program teams. Embedding AfL at program levels is a useful strategy for scaling up. Program-wide initiatives are a site for embedded professional development of university teachers in that they involve a range of colleagues discussing practice in context. For example, the TESTA (Transforming the Experience of Students through Assessment) project methodology promotes a program-enhancement approach to assessment through the careful triangulation of data from the Assessment Experience Questionnaire and focus group interviews (Jessop, this volume).

The use of technology to enable innovative approaches to assessment and feedback represents a further possible driver for the scaling up of AfL practice. For this potential to be fulfilled there may need to be synergies between assessment literacy, technological literacy and professional development. Related issues are taken up by other chapters in this volume. Moscrop and Beaumont illustrate how technology can facilitate dialogic feedback cycles as a means of encouraging student uptake of feedback. The chapter by Dawson and Henderson takes both a critical perspective on technology-enabled AfL and suggests some possibilities for scaling up both practice and the related research base.

The above discussion is suggestive of some avenues for future research and development activity, extending or going beyond some of the discussion in this volume. What are effective ways of developing staff assessment and feedback

literacy? How does staff assessment literacy help to seed student assessment literacy? What forms of leadership and support are most conducive to developing AfL? Under what circumstances is quality assurance a barrier to AfL and when might it support its further development? What AfL practices can be scaled up to operate effectively with large classes and multiple tutors; and how can technology effectively enable these processes?

The volume is arranged in four parts. Part 1, Enabling assessment change, contains this chapter and the contributions of Boud, Hounsell and Zou, and Jessop. Part 2 focuses on AfL strategies and implementation with chapters by Glofcheski, Bridges and colleagues, Jonsson and Panadero, and Carless. Part 3 is entitled Feedback for learning and contains chapters by Ajjawi and colleagues, Pitt, Rowe, and Narciss. The final section, Using technology to facilitate AfL, involves chapters by Moscrop and Beaumont, and Dawson and Henderson. Most chapters explicitly address the scaling up theme, whereas others focus more on the specific AfL issue which their chapter addresses.

References

Alvarez, I., Espasa, A. & Guasch, T. (2012). The value of feedback in improving collaborative writing assignments in an online environment. *Studies in Higher Education, 37*(4), 387-400.

Australian Council for Educational Research (ACER). (2010). *Doing more for learning: Enhancing engagement and outcomes. Australasian student engagement report*. Camberwell, Victoria: Australian Council for Educational Research.

Baartman, L., Bastiaens, T., Kirschner, P. & van der Vleuten, C. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, *2*(2), 114-129.

Bevitt, S. (2015). Assessment innovation and student experience: A new assessment challenge and call for a multi-perspective approach to assessment research. *Assessment & Evaluation in Higher Education, 40*(1), 103-119.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the Black Box: Assessment for Learning in the classroom. *Phi Delta Kappan*, *86*(1), 8-21.

Black, P., & R. McCormick. (2010). Reflections and new directions. *Assessment & Evaluation in Higher Education*, *35*(5), 493-499.

Black, P., & D. Wiliam. (1998). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7-74.

Boud, D. (1995). *Enhancing learning through self-assessment*. London: Kogan Page.

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* (pp.181-197). London: Routledge.

Boud, D. (2014). Shifting views of assessment: From secret teachers' business to sustaining learning. In C. Kreber, C. Anderson, N. Entwistle & J. McArthur (Eds.) *Advances and innovations in university assessment and feedback* (pp. 13-31). Edinburgh: Edinburgh University Press.

Boud, D. & Associates (2010). Assessment 2020: Seven propositions for assessment reform in higher education. Retrieved from https://www.uts.edu.au/sites/default/files/Assessment-2020_propositions_final.pdf

Boud, D. & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, *38*(6), 698-712.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, *31*(2), 219-233.

Carless, D. (2009) Trust, distrust and their impact on assessment reform, *Assessment & Evaluation in Higher Education*, *34*(1), 79-89.

Carless, D. (2015). *Excellence in University Assessment*: *Learning from award-winning practice*. London: Routledge.

Chen, Q., Kettle, M., Klenowski, V. & May, L. (2013). Interpretations of formative assessment in the teaching of English at two Chinese universities: A sociocultural perspective. *Assessment & Evaluation in Higher Education*, *38*(7), 831-846.

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3-12.

Dochy, F., Segers, M., Gijbels, D. & Struyven, K. (2007). Assessment engineering: Breaking down barriers between teaching and learning, and assessment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp.87-100). London: Routledge.

Dochy, F., Segers, M. & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, *24*(3), 331-350.

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.

Gibbs, G. (2006). How assessment frames student learning. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp.23-36). London: Routledge.

Hattie, J. (2009). *Visible learning*. London: Routledge.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, *27*(1), 53-64.

Higher Education Funding Council for England (HEFCE). (2014). National Student Survey. Retrieved from http://www.hefce.ac.uk/whatwedo/lt/publicinfo/nss/

Kezar, A. (2011). What is the best way to achieve broader reach of improved practices in education? *Innovative Higher Education*, *36*(4), 235-247.

Knight, P. (1995). *Assessment for learning in higher education*. London: Kogan Page.

Leahy, S., & D. Wiliam. (2012). From teachers to schools: Scaling up professional development for formative assessment. In J. Gardner (Ed.), *Assessment and learning* (2nd ed., pp. 49-71). Thousand Oaks: SAGE.

Meyer, L., Davidson, S., McKenzie, L., Rees, M., Fletcher, R., & Johnston, P. (2010). An investigation of tertiary assessment policy and practice: Alignment and contradictions. *Higher Education Quarterly*, *64*(3), 331-350.

Nicol, D. & Draper, S. (2009). A blueprint for transformational organizational change in higher education: REAP as a case study. In T. Mayes, D. Morrison, H. Mellar, P. Bullen & M. Oliver (Eds.), *Transforming higher education through technology-enhanced learning* (pp.191-207). York: Higher Education Academy.

Nicol, D. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218.

Nicol, D., Thomson, A. & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, *39*(1), 102-122.

Norton, L., Norton, B. & Shannon, L. (2013). Revitalising assessment design: What is holding new lecturers back? *Higher Education*, *66*(2), 233-251.

Price, M., Rust, C., O'Donovan, B., and K. Handley. (2012). *Assessment literacy: The foundation for improving student learning.* Oxford: Oxford Centre for Staff and Learning Development.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119-144.

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

Sambell, K., McDowell, L., & C. Montgomery. (2013). *Assessment for learning in higher education.* London: Routledge.

Strijbos, J., Narciss, S. & Dunnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction*, *20*(4), 291-303.

Van der Vleuten, C. & Schuwirth, L. (2005). Assessing professional competence: From methods to programmes. *Medical Education*, *39*(3), 309-317.

Wiliam, D. (2007). Content then process: Teacher learning communities in the service of formative assessment. In D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 183-206). Indiana: Solution Tree.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3-14.

Wiliam, D., & M. Thompson, (2008). Integrating assessment with learning: What will it take to make it work? In C. Dwyer (Ed), *The future of assessment: Shaping teaching and learning* (pp. 53-84). New York: Lawrence Erlbaum Associates.

Williams, J. & Kane, D. (2009). Assessment and feedback: Institutional experiences of student feedback, 1996-2007. *Higher Education Quarterly*, *63*(3), 264-286.