Putting Assessment for Learning (AfL) into Practice in a Higher Education EFL Context

Edmund White

Putting Assessment for Learning (AfL) into Practice in a Higher Education EFL Context

Contents

Synopsis

In the past twenty years or so, assessment in education has become increasingly viewed as a means of guiding and improving student learning, rather than simply measuring it. The assessment for learning (AfL) movement, arising primarily from mainstream primary and secondary education in the UK, strongly advocates this formative, learning-centered view of assessment. At the level of classroom practice, AfL is comprised of a number of key procedures including: the use of effective feedback, self-and peer assessment, questioning and classroom dialogue, and the formative use of summative tests. Based on a series of classroom research projects, this thesis examines these AfL procedures in a higher education EFL context with adult students at *Tokyo Woman's Christian University*. Providing the impetus, the theoretical framework, and serving as a crucial resource for this research, has been a book entitled, *Assessment for Learning: Putting it into Practice* (Black, Harrison, Lee, Marshall & Wiliam, 2003).

After an introductory chapter in which AfL theory and fundamentals are presented, the three body chapters of this thesis are comprised of a series of five reports of AfL procedures being used in various classroom and teaching contexts. Chapter 1 describes a research project centered on the issue of feedback on the first draft of student essays, and begins with a teacher self-assessment of written feedback. This is followed by a report of the impact of the feedback on student essays, and how students felt about the process.

Chapter 2 focuses on the student in the assessor role, and begins in Part A with a report focused on self-assessment of class participation in a freshman English class. This is followed in Part B by examining the issue of peer assessment in a public speaking course.

Chapter 3 first investigates the issue of questioning in classroom dialogue, in particular, student-generated questioning. The final research project report then examines the issue of using summative assessment for formative purposes in promoting student learning. Following Chapter 3, a conclusion draws the five AfL reports together for a more holistic view of the process of putting these AfL procedures into practice in a higher education EFL context with adult learners.

The goal of assessment has to be, above all, to support the improvement of learning and teaching.

(Frederiksen & Collins, 1989, p. 32)

Any attempt to turn ideas into practice will be a learning experience . . .

(Black, Harrison, Lee, Marshall & Wiliam, 2003, p. 118)

Introduction

Learning, assessment, and higher education

How people learn is undoubtedly a complex business. However, it seems that a core component of learning involves the notion of *change* on the part of the learner – whether in developing knowledge and understanding, or skills and abilities in a particular area or activity. Nichols (2004, p. 54) writes that learning is defined by "changes in knowledge, understanding, skills and attitudes brought about by experience and reflection on that experience, whether that experience is structured or not". Explaining that it is more than just a process of ingesting information, Maki (2004) defines learning as "a multidimensional model of making meaning" (p. 1). In a classroom context then, the educators' primary role may be described as 'a facilitator of student meaning-making'. This simple description belies the complexity and challenge of effective teaching, the kind which promotes and maximizes student learning. Assessment has come to be seen as a crucial bridge in aiding this teaching and learning process.

For professionals at all educational levels, the primary goal of promoting and deepening student learning would be an idea that few would take issue with, despite the wide-ranging, often conflicting, views of the most effective ways of achieving this goal. In the past twenty years or so, with the publication of such works as Sadler's (1989) influential "Formative assessment and the design of instructional systems", there has been a continuous and developing recognition of the vital role assessment can play in promoting, rather than simply measuring, student learning. While research and publications have focused largely on primary and secondary schooling in mainstream education, the potential beneficial role of assessment in higher education (HE) has also become more recognized in recent years. For example, in writing about tertiary assessment, Bryan and Clegg (2007) assert that assessment and learning should be seen as working in tandem, with each contributing to the other.

According to Atkins (1995, p. 25) four overlapping purposes for higher education may be distinguished:

1. To provide a general educational experience of intrinsic worth in its own right.
2. To prepare students for knowledge creation, application, and dissemination.
3. To prepare students for a specific profession or occupation.
4. To prepare students for general employment.

Supplementing these purposes, the role of tertiary education in establishing a framework for

continued and lifelong learning has also been recognized (for example, Boud, 2000). According to Boud and Falchikov (2007), the provision of a foundation for lifelong learning beyond the academy in work and other social settings is the very "raison d'être of a higher education" (p. 399).

The impact of assessment in working toward these higher education purposes has become increasingly appreciated. For university instructors, classroom assessment plays a key role in and significant influence on both their teaching and student learning (Cheng, Rodgers & Wang, 2008). The role, influence and beneficial potential of assessment in higher education are well captured by McInnis and Devlin (2002) below in Box 1.

Box 1. *Influence of HE assessment*

Assessment is a central element in the overall quality of teaching and learning in higher education. Well-designed assessment sets clear expectations, establishes a reasonable workload (one that does not push students into rote reproductive approaches to study), and provides opportunities for students to self-monitor, rehearse, practice and receive feedback. . . . Carefully designed assessment contributes directly to the way students approach their study and therefore contributes indirectly, but powerfully, to the quality of their learning.

(McInnis and Devlin, 2002, p. 7)

As will be seen, a number of the issues raised here by McInnis and Devlin (2002), and their general argument that effective assessment promotes more effective student learning and teaching practice, will be evident throughout this thesis. The impact of classroom-based assessment on student learning is the central focus under consideration here. A higher education institution in Japan is the setting for this assessment and learning research report, and while describing a variety of classroom practices and processes, "assessment is our focus but learning is the goal" (Gardner, 2006, p. 2).

In writing about the issue of originality in HE doctoral research, Dunleavy (2003) points out that it rarely entails coming up with an entirely new way of looking at things. Instead, originality in modern social sciences and humanities often involves,

 . . . encountering an established idea or viewpoint or method in one part of your discipline (or neighboring discipline) and then taking that idea for a walk and putting it down somewhere else, applying it in a different context or for a different purpose. (Dunleavy, 2003, p. 40)

This thesis involves taking assessment for learning (AfL) theory and practice from mainstream school education for a walk and putting them down in Japan at a tertiary setting – a woman's university in Tokyo - with students for whom English is a foreign language. Yet, while the context does indeed differ, the purpose remains the same – using assessment to promote and maximize student learning.

This introduction will delineate in more detail the nature of AfL in practice, particularly with reference to tertiary education, and provide an overview of the research conducted at the Japanese university. But first, it is appropriate to clarify the fundamental ideas, practices and theory associated with AfL, or formative assessment. This will provide an underlying framework for the reports and discussions of classroom assessment procedures to follow. It should be noted that this introductory section will simply present some of the essential ideas associated with AfL, and that subsequent chapters, comprised of five research reports, will each offer further details and connections to the now extensive AfL-related literature.

Part A: AfL fundamentals and theoretical background

Fundamental principles and practices

More than ten years after the publication of Black and Wiliam's (1998a) seminal review of formative assessment research, AfL may be considered as 'an established idea or viewpoint' in classroom-based assessment literature for primary and secondary education, particularly in the UK. AfL theory and practice have become widely disseminated and appreciated, if not implemented. A grouping of educators and researchers advising on assessment policy and practice in the UK, the Assessment Reform Group (ARG), is perhaps the key information source for AfL research and literature. The ARG is comprised of some of the most important AfL writers and researchers, and its Internet website provides a rich fount of fundamental summaries of what AfL entails.

In 1999, the ARG published a follow-up to the critical research review by Black and Wiliam (1998a), and noted that while assessment is one of the most powerful educational tools for promoting effective learning, it must be properly used to have this effect. They explain that the research shows that assessment can be used to improve learning depending on five key factors, shown in Table 1 below.

Table 1. *Improving learning through assessment (ARG, 1999, p. 4-5)*

Five key factors for using assessment to improve learning
1. Providing effective feedback to students
2. The active involvement of students in their own learning
3. Adjusting teaching to take account of assessment results
4. Recognizing the profound influence assessment has on student motivation and self-esteem (both crucial influences on learning)
5. The need for students to assess themselves and understand how to improve

According to the ARG (1999), these features are essential for effective day-to-day learning in the classroom, as well as being key components for successful lifelong learning. However, this report also noted that in reality, classroom practice often fails to meet these key factors for both students and their teachers.

In a 2002 publication, the ARG provides the most commonly referred to definition of AfL, and ten founding principles which have become the qualities collectively attributed to this assessment concept (Gardner, 2006). The definition, and founding principles are provided below in Table 2.

Table 2. *AfL definition and ten founding principles (ARG, 2002)*

Assessment for learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.	Assessment for Learning
	- is part of effective planning
	- focuses on how children learn
	- is central to classroom practice
	- is a key professional skill
	- is sensitive and constructive
	- fosters motivation
	- promotes understanding of goals and criteria
	- helps learners know how to improve
	- develops the capacity for self-assessment
	- recognizes all educational achievement

A noticeable theme in both the definition and the ten principles is the primary focus on *the learner* and *learning*, rather than on the teaching or subject matter, as part of the assessment process. Another key idea is determining where students are in their learning, and ways to move forward. One further relevant point here, in the context of this thesis, is the second principle listed above and use of the term 'children'. This is a common reference label for students in the AfL literature, and the ideas and practices are commonly associated with younger learners, rather than adults.

Terminological usage (assessment *of*, *for*, and *as* learning, formative assessment)

When considering or constructing assessment tools and procedures, a critical factor is the purpose for their use. A key distinction, and common reference point in the literature, is between assessment *for* learning, or formative assessment and assessment *of* learning, or summative assessment. James (2006) explains the difference as follows:

A distinction between formative and summative (summing-up) purposes has been familiar since the 1960s, although the meaning of these two terms has not been well understood. A more transparent distinction, meaning roughly the same thing, is between assessment *of* learning, for grading and reporting, and assessment *for* learning, where the explicit purpose is to use assessment as part of teaching to promote pupils' learning. (p. 8)

Stiggins (2002) similarly observes that the crucial distinction between assessment *of* and f*or* learning is "between assessment to determine the status of learning and assessment to promote greater learning" (p. 761). This distinction in assessment purposes is an important one, but in reality, and as will be seen in this thesis, summative and formative purposes are often intertwined.

Some educators, particularly Earl (2003), differentiate between teacher and student roles within AfL, and have introduced the further distinction of 'assessment *as* learning'. This term essentially refers to the role played by the student, while AfL focuses more on the teachers' role in the classroom. Following Glassen (2009), this thesis uses the broader term 'assessment *for* learning' as most of the research literature uses the term in this manner, it is less confusing, and "it confirms that the role of assessment is an integrated one, involving both teachers and students in a mutually responsible, symbiotic and potentially productive relationship" (p. 5).

A point should be made here also about the distinction between the terms 'assessment for learning' and 'formative assessment'. Various writers have linked these two terms (Marsh, 2007), and they are commonly assumed to carry the same meaning. The potential problem, as noted by Black, Lee, Marshall, and Wiliam (2003), is that 'formative assessment' may be interpreted in different ways by teachers and can often mean simply that assessment is frequently carried out and planned at the same time as teaching. The assessment in question may be formative for the teacher, but not for the students (Black et al., 2003). Formative assessment *for the learner* should be the point of emphasis here.

While noting that 'assessment for learning' and 'formative assessment' may convey meanings which are slightly different (Hodgen & Marshall, 2005), they are used interchangeably in this thesis because, as Gardner (2006) notes, "in the final analysis, there is little of substance to distinguish the two terms" (p. 2). Indeed, in the body of AfL literature the two terms are often interchangeably used (for example, Wiliam, Lee, Harrison & Black, 2004; Hodgen & Marshall, 2005; James & Pedder, 2006a; and Harlen, 2007). Wiliam (2008) captures the similar essence of the two terms when writing, "Just as we use the term *formative* to describe the experiences that shape us as we grow up, a formative assessment is one that shapes learning" (p. 9). Shaping, promoting and deepening student learning are also fundamental elements encapsulated in the term 'assessment *for* learning'.

AfL in practice

The above terminological clarification leads us to a final point to be made in this section about how AfL looks, or should look, in practice. Tables 1 and 2 above from the Assessment Reform Group (1999, 2002) present the key principles associated with AfL, but how these translate into classroom practice are not explicated. Because formative assessment may be misunderstood, or weakly implemented by teachers (Black & Wiliam, 1998a) the ARG (1999) insists that it is important to distinguish AfL from other inadequate interpretations of formative classroom assessment. Such practices may include what becomes essentially summative assessment in nature; activities like simply adding procedures or tests to existing work, or on-going assessment which only feeds marks or grades back to students. In such cases, the problem is that *student learning* and its promotion are not fore-fronted. Black et al. (2003) are also explicit about avoiding any misunderstanding when they assert,

. . . It is important to emphasize the critical criterion – formative assessment is a process, one in which information about learning is *evoked* and then *used* to modify the teaching and

learning activities in which teachers and students are engaged. (p. 122, original emphasis)

In order to make good AfL practice explicit and show teachers how it is related to effective teaching and learning, the ARG (1999) identify seven "key characteristics of AfL practice" to correspond with the principles delineated in the tables above. The practical classroom applications of AfL for teachers and students are laid out here in Table 3.

Table 3. *Fundamental characteristics of AfL in practice (ARG, 1999)*

Seven key characteristics of AfL
1. It is embedded in a view of teaching and learning of which it is an essential part.
2. It involves sharing learning goals with pupils.
3. It aims to help pupils to know and to recognize the standards they are aiming for.
4. It involves pupils in self-assessment.
5. It provides feedback which leads to pupils recognizing their next steps and how to take them.
6. It is underpinned by confidence that every student can improve.
7. It involves both teacher and pupils reviewing and reflecting on assessment data.

These fundamental characteristics of putting AfL into practice are important reminders for educators of the essentials that need to be focused on. This ARG (1999) characterization of the practices AfL principles should exemplify is an important point of reference, especially considering the "plethora of products and services that describe themselves as 'formative assessment' " (Wiliam, 2007, p. 1), but in reality do not promote student learning. The seven characteristics also provide excellent criterion for evaluating the AfL practices and procedures implemented in the five research projects reported on in this thesis. They will be used in this manner in the conclusion of this thesis as part of an overview of the entire research project, and to assess the individual five studies according to recommended AfL practice.

Recent descriptions of AfL

While there has been some confusion and misunderstanding of the meaning of 'assessment for learning' or 'formative assessment', there has been some clarification of what these terms encompass in recent years (James, Black, McCormick & Pedder, 2007). We end this section and its presentation of fundamental ideas with three relatively recent descriptions of AfL. These explanations insightfully and descriptively gloss the central assessment/education concept of this thesis.

Recent descriptions of AfL:

♦ Assessment *for* learning: the use of the formative assessment process and its results as an instructional intervention designed to increase - not merely to monitor and grade - pupil learning. Research evidence gathered in hundreds of studies conducted around the world over the last decade . . . shows that the consistent application of principles of assessment for learning can give rise to unprecedented gains in pupil achievement, especially for perennial low achievers. (Stiggins, 2007, p. 17)

♦ The ultimate goal of AfL is therefore to involve pupils in their own assessment so that they can reflect on where they are in their own learning, understand where they need to go next and work out what steps to take to get there. The research literature sometimes refers to this as the processes of self-monitoring and self-regulation. In other words, pupils need to understand both the desired *outcomes* of their learning and the *processes* of learning by which these outcomes are achieved, and they need to act on this understanding. (James et al., 2007, p. 7, original emphasis)

♦ Assessment for Learning (AfL) aims to inform and improve student learning within the regular flow of teaching and learning through students becoming active meaning makers and thoughtful judges of their own learning. (Willis, 2007, p. 52)

As we shall see, putting AfL into practice can also result in *teachers* becoming 'active meaning makers' and more 'thoughtful judges' in the classroom.

In short, AfL is *for* students, *for* learning, and also *for* more effective teaching.

Theoretical background

As the Stiggins (2007) description of AfL above acknowledges, empirical research evidence

for the effectiveness of formative assessment practices is both widespread and extensive (documented in research reviews by Natriello, 1987; Crooks, 1988; Kluger & DeNisi, 1996; Black & Wiliam, 1998a; Nyquist, 2003; and Brookhart, 2005). Theoretical support for AfL principles and practices has also been established, for example from contemporary psychology and the contention that understanding is created by the learner (James, 2006). According to Biggs (2003), this learning theory termed *constructivism* enjoys a long history in the field of cognitive psychology, having the central idea that "what the learner has to *do* to create knowledge is the important thing" (p. 12, original emphasis). As we have noted, implicit in the concept of learning are notions of change and development of understanding or ability. Regarding learning in an educational context, Harlen (2007b) explains:

Although learning processes in detail vary among subjects, underpinning learning in all subjects is a view of learning as the progressive development of understanding, in which new experience is linked to existing knowledge. (p. 117)

In this constructivist view of learning, the development of understanding involves a process of construction and reconstruction of knowledge by the student (Harlen & James, 1997). It is a view of learning that provides a basis for the active participation of the learner in classrooms in which formative assessment is practiced.

Teaching is based on certain assumptions of how people learn, and the three main perspectives on learning identified by learning theory writers are *behaviorist*, *constructivist* and *socio-cultural* (James et al., 2007). Table 4 below presents a summary of the main ideas connected with each of these learning theories. Two variations of constructivist theory are identified: cognitive constructivism and social constructivism.

Table 4. *Summary of three main learning theories (based on James, et al., 2007, pp. 16-20)*

Behaviorist	Constructivist	Socio-cultural
Assumptions about learning originate in behaviorist psychology which sees learning as a conditioned response to a stimulus. Behaviorism is most concerned with behavior, not the mental life going on in a persons' head.	**Cognitive constructivist** approaches focus attention on the mental models that a learner uses when responding to new information or problems. Learning always involves analyzing and transforming any new information. The reception of new knowledge depends on existing knowledge and understanding.	People learn through participating in 'communities of practice', like apprentices. Through membership and activity they come to understand what to pay attention to and what counts as quality in a particular group.
For teaching and assessment: 1) rewards (or withholding them) are powerful ways to establish desired behaviors, 2) a complex skill can be taught by breaking it down and teaching and testing the pieces separately, 3) it is best to learn facts and basic skills first, understanding will come later.	Self-awareness and self-regulation (meta-cognition) are viewed as key elements for learning. Students need to understand what it means to learn and they need to monitor how they go about planning, monitoring and revising. They need to reflect upon their learning and to learn to determine for themselves whether they understand or can do something.	Learning through participation in communities of practice is important. Collective expertise can be more productive than individuals working alone. Group work is essential for learning.
Behaviorist theory has lost favor in recent years, but many practices associated with it are still widespread.	**Social constructivist** approaches extend these cognitive ideas. Contends that learning proceeds by interaction between the teacher and learners in a social context, mediated by the social norms that value the search for understanding.	The role of experts guiding novices is an important concept.

Constructivist learning theory is highlighted and centered in Table 4 to reflect the fact that it provides a central theoretical base for AfL. James et al. (2007) affirm that,

Most approaches to assessment for learning have been developed within a cognitive constructivist framework for understanding learning, although, Black and Wiliam (2006) have begun to develop a theory of formative assessment drawing on socio-cultural perspectives. (p. 18)

Indeed, in the past ten years the theoretical basis of formative assessment has undergone a number of formulations (Black & Wiliam, 2009) and shifting points of emphasis, including

more attention being paid to the social constructivist and socio-cultural theories noted in Table 4 above. Yet, the cognitive constructivist paradigm remains a central one for AfL, and for this thesis, as students are seen to be actively involved in constructing understanding and developing skills during the learning process.

As noted in Table 4, while still influential in practice, behaviorist learning theory has fallen out of favor in recent years. Ellery (2008) observes in Box 2 below that, at least in terms of effective classroom-based assessment, the sands have been shifting away from assessment practices and frameworks associated with behaviorist educational theories and towards more constructivist approaches.

Box 2. *Shifting perspectives on assessment: from behaviorism to constructivism*

The academic value of such an approach [traditional behaviorist assessment methods] has been increasingly questioned and in recent years there has been a shift towards a constructivist, student-centered approach that integrates assessment with learning. This approach assumes knowledge is not a fixed, identifiable entity to be absorbed by the learner but instead is constructed by students based on their own understanding, which is influenced by their background, perspectives and experiences. As a result this type of assessment tends to be more flexible, integrative, contextualized, process- oriented, criteria-referenced and formative. This 'assessment for learning' approach encourages student independence and self-evaluation and can lead to active and deeper learning. (Ellery, 2008, p. 421)

This constructivist paradigm underlies this investigation of putting AfL ideas into practice in a higher educational setting. A desideratum for the university classes discussed in this thesis was for students to participate in knowledge and meaning construction as they actively engaged in the learning process; which is the critical emphasis of constructivism (Ellery, 2008). In the classroom, Murphy (1994) describes how learning, teaching and assessment may be intertwined within a constructivist approach:

… with a constructivist framework, students are seen as active collaborators in the building of knowledge. Learning takes place through interaction, existing in the transaction between student and student, student and text, student and teacher. Viewed from a constructivist perspective then, assessment procedures are inevitably a part of the dialectic of teaching and learning, part of the process which defines what knowledge is, what is learned, and how students learn. Assessments which reflect this perspective provide a means for engaging students in self-reflection and for acknowledging their role as collaborators in the learning process. In sum, a constructivist perspective acknowledges the reciprocity and interdependency of assessment and curriculum. (p. 190)

The emphasis here is on effective assessment promoting student engagement with peers and teacher in the collaborative construction of understanding and meaning, and promoting student reflection on this process.

This connection with constructivist learning principles and formative assessment is commonly made and referred to in the assessment literature (for example, Sheppard, 2000; Black, 2001; Marshall & Drummond, 2006; Carless, 2007; MacMillian, 2007).

To briefly summarize Part A, AfL principles and practices are now well established; as are arguments supporting this formative assessment model, both empirically and theoretically.

Part B: Literature overview – AfL in classroom practice

As noted earlier, each of the studies reported in this thesis have reviews of the pertinent AfL literature related to the particular chapter focus in question (for example, regarding feedback or self-assessment). Prior to delineating and previewing the research conducted and presented in those chapters, this brief literature overview takes a broader perspective of the spread of AfL ideas and practice, and considers this model of formative assessment as related to subject differences, tertiary settings and to an EFL context. A rationale for the research engaged in and reported on here is also provided.

Dissemination of AfL ideas and practices

For AfL, the essential orientation point and key text is the formative assessment research review commissioned by the UK's Assessment Reform Group (ARG) and published by Paul Black and Dylan Wiliam in 1998. Black and Wiliam clearly showed the existence of strong research evidence that formative assessment can promote learning and raise levels of student achievement. In the decade following this seminal and highly influential work, AfL has become an established element in official education policy throughout the UK (Daugherty & Ecclestone, 2006), as well as receiving much attention internationally. In Europe, the Organization for Economic Co-operation and Development (OECD) has published formative assessment research using case studies from eight countries: Canada, Denmark, England, Finland, Italy, New Zealand, Australia and Italy (OECD, 2005). This OECD report recognized that formative assessment is a powerful teaching and learning tool in the new

century: "Teachers using formative assessment approaches guide students toward development of their own learning to learn skills that are increasingly necessary as knowledge is quickly outdated in the information society" (OECD, 2005, p. 22). It should be noted that this extensive case study publication focused on lower secondary education contexts.

Lee (2007) also observes that the classroom potential for AfL has become increasingly recognized, and impacted classroom assessment practices in places such as Hong Kong and Australia, in addition to the UK. Leung & Lewkovicz (2006) report that ARG's ten principles of AfL (listed earlier in Table 2) have received strong professional endorsement, and been incorporated in a number of policy statements in a variety of diverse international jurisdictions, including those countries previously mentioned. At the mid-point of this decade, and reflecting on the years since the research review by Black and Wiliam (1998), Black summarized the situation as follows:

Formative assessment has been a growth industry in the last seven years, fuelled by evidence of two kinds – from research showing that formative practices improve pupils' achievement, and from practice showing that teachers can transform ideas from the research into productive practices. While its impact has been most extensive in the UK, there have been significant innovations in other countries. (Black, 2005, p. 133)

Later, in 2009, Black noted that a good deal of research attention had recently been focused on formative assessment, and this included second language education. He also commented on the lively discussion among language educators and applied linguistics regarding the relationship between assessment and learning (Black, 2009).

This decade has seen a considerable amount of work in terms of research and professional conferences in the Asia-Pacific region where interest in AfL has expanded in recent years (Klenowski, 2009). To take just one example, since 2006 an annual International Conference on Assessment for Learning has been held, most recently in New Zealand in 2009. AfL has become part of education policy in such places as Singapore, Malaysia and Hong Kong, and the subject of published research in Korea (e.g. Butler, 2009). Primary and secondary education has been the main contextual setting for this development of AfL theory and practice.

However there has been some limited development of formative assessment practice in higher education contexts in Asia. For example, in 2005, tertiary educators from the region met for *ELTA 2005: The First International Conference on Enhancing Teaching and Learning through Assessment*. A range of papers were presented here, and this was later

followed up by a related book describing the alternative use of assessment with illustrations of selected cases from different universities (Frankland, 2007). Content is primarily related to courses such as business, computing, graphic design and engineering rather than the teaching and learning of English as a foreign language.

Despite such developments, and while AfL ideas and practice have indeed become widely disseminated and incorporated in policy statements and classroom practice, its implementation, despite being significant, still remains limited. In particular, high-stakes state-mandated testing which exists in such places as most American states, and in England, makes it difficult to effectively implement formative assessment practices (Wiliam, Lee, Harrison & Black, 2004). Cumming (2009) points out that while AfL principles have been well established, there are scarce examples of their programmatic applications.

Carless (2007) makes this cautionary note also, disclosing that while there are indeed 'pockets of successful implementation in schools' in such places as England, Australia and New Zealand, large-scale implementation is considerably challenging and has not occurred (p. 173). Heavy workloads and large class sizes present barriers for teachers wishing to implement formative assessment practices in their classrooms, and may lead teachers to believe that while AfL is theoretically sound, it is rather impractical, time-consuming and incompatible with schooling demands (Carless, 2007). Making a further observation related to non-western cultural contexts, Carless (2007) writes: "Prospects for the implementation of formative assessment are even more daunting in various international contexts where transmissive teaching and summative assessment have characteristically dominated" (p. 173). As will be noted, transmissive teaching and the domination of summative assessment are characteristics of the Japanese educational context.

Yet, it must be acknowledged that AfL has had an impressive impact in the relatively few years since the work of ARG educators and researchers, and others, have promoted and disseminated its ideas and practices. James et al. (2006) write, "the huge interest in assessment for learning in recent years is because it shifts assessment practices to serve formative purposes; to improve and not just measure learning" (p. 1). Yet despite the extensive interest and body of literature already created, there is much to learn and explore, as Black (2005) affirmed: "There is a new field here [formative assessment], which will continue to be rich in possibilities for fruitful interactions between research and practice, to the benefit of both" (p. 135). One possible avenue for fruitful research and interaction with practice, one little explored thus far, is implementing AfL in a higher education context and

with EFL-based subjects. As such, the research described in this thesis is focused on the 'programmatic applications' of AfL principles (Cumming, 2009) and their impacts on student learning, rather than engaging in the various and ongoing theoretical debates that exist around this assessment model.

Subject differences and AfL practice

While AfL in classroom practice, and related research and writing, mostly occurs in primary and secondary schooling, it is also concentrated on certain subject areas. According to Harlen and Winter (2004),

> There is now a considerable body of literature on the type of classroom assessment that benefits learning, much of it carried out in relation to mathematics, science and English (*as a first as opposed to a foreign or additional language*) . . .. (p. 391, emphasis added)

Marshall (2007) posits a rationale for the reasons why much of AfL literature arises from work in math and science subjects.

> Much of the literature on formative assessment, or assessment for learning (AfL), has arisen out of the work of math and science. This is possibly because the constructivist view of learning in these disciplines lends itself to very clear paths of progression onto which formative assessment can be readily mapped. Work of a similar nature does not exist in English and the humanities or in the social sciences, where progression is a much messier business. It is hard, for example, in English to be precise about the developmental trajectory of the imagination. (p. 136)

The ARG (2002) recognizes that the way AfL principles manifest themselves in different subjects may differ. However, Black et al. (2003) explain that while different formative assessment techniques may have more or less use in different subjects, all of the broad strategies (e.g. self-assessment, feedback, sharing criteria with learners) are applicable across the subject spectrum. It should be noted here that this claim is a subject of ongoing debate in the related literature, with some writers questioning the application of these formative techniques across all subjects (e.g. Rea-Dickens, 2001; Davison and Leung, 2009). Based on a consideration of the research conducted with teachers in all subject areas in the UK, Black and Jones (2006) observe that "formative assessment has *generic* features, which will apply to learning across all stages and school subjects, and features which are *specific* – to primary teachers and to individual schools subjects" (p. 4, original emphasis). It is these generic

features (such as self-assessment or effective feedback) and their practical application to an EFL context in a Japanese tertiary setting that are the focus of the research projects reported on in this thesis.

One difference noted by Black et al. (2003) is concerned not with a particular school subject itself, but rather how it is interpreted by the teacher in the classroom. Black et al. (2003) contend that when learning goals are very specific, teachers will tightly regulate students' work. On the other hand, regulation of students work will be looser when the goals are less well defined. As we shall see, this notion of learning goal specificity and teacher regulation of student work has implications for the EFL courses described in this thesis, where "subject areas may not have a clear linear or hierarchical structure" (James et al., 2007, p.). Wiliam and Leahy (2007) are more specific about the teachers' classroom regulation of learning in different subjects:

In most teaching of mathematics and sciences, the regulation of learning will be relatively tight. . . . In contrast, in the teaching of language arts and social studies, the regulation will be much looser. Rather than a single goal, there is likely to be a broad *horizon* of appropriate goals, all of which are acceptable . . . . (p. 35, original emphasis)

Thus, putting AfL into practice may be affected by the learning goals being focused on, and the classroom subject being taught. However, Hodgen and Marshall (2005) report that, rather than subject specific approaches, AfL research has focused largely on generic pedagogic approaches and strategies that are applicable to all teaching and learning in all subject areas.

There is some limited amount of literature available regarding language teaching and AfL (for example, Ofsted, 2003; Black & Jones, 2006), particularly coming from UK schools and the teaching of what is termed Modern Foreign Languages (MFL) (e.g., French, German, and Spanish). In a report about good assessment practices in MFL, the Office for Standards in Education (Ofsted) in the UK remind teachers that "Assessment in modern foreign languages (MFL) is most effective when it is seen as an integral part of teaching and learning" (Ofsted, 2003, p. 1). This is a common underlying theme of AfL practice: using assessment to inform teaching and learning.

The Black and Jones (2006) article on the learning and teaching of foreign languages and formative assessment is one of the few available resources for modelling how to put AfL into practice in such courses. Aside from the generic/specific formative assessment features distinction mentioned earlier, two further points in particular are worthy of note from Black

and Jones (2006, p. 5). They emphasize:

- [implementing] a formative assessment framework where language learning is demystified through the sharing of learning intentions and success criteria, and that
- the ultimate aim of language learning is for learners to make progress in their foreign language competence and to monitor their own progress as independent language learners.

While Black and Jones (2006) are referring to primary/secondary language learners in a UK curriculum with government set standards (and students and teachers typically sharing the same first language of English), this 'ultimate aim' for both learner progression of language competence and progress monitoring, is universally applicable; including the context of the research reported here. These crucial language learning aims (progress in language competence, and monitoring of this progress) are also seen in this thesis as being embedded in the broader aims of language teaching and assessing, with the "ultimate goal" being "improved learning through effective assessment practices" (ARG, 2008, p.18).

Research rationale: AfL and higher education

As with the OECD (2005) research report referred to above, the bulk of AfL research carried out has been in primary and secondary schooling (Yorke, 2003; Murphy, 2006). Indeed, while formative assessment principles may be generic and universally applicable, the AfL literature is heavily based on, and geared towards, the education of younger learners, largely from developed Anglophone cultures (Stobart, 2006). Murphy (2006) observes that the body of knowledge is growing regarding the application of formative assessment in school classrooms, "*but in higher education we are still at an early stage of understanding how effectively this approach to assessment can be developed within different higher education institutions and in different areas of the curriculum*" (p. 42, emphasis added). Reflecting the point made earlier about behaviorist learning theory still retaining influence, and despite the shift towards constructivist theories of assessment noted earlier (Ellery, 2008), Biggs (2003) writes that in HE contexts "the view of university teaching as transmitting information is so widely accepted that delivery and assessment systems the world over are based on it" (p. 12). Pryor and Crossouard (2005) concur, and claim: "Within universities, especially the more prestigious ones, traditional forms of assessment have been largely taken for granted and

developments in formative assessment ignored" (p. 1).

Willis (2006) observes that a lot is known about "the complexity of AfL as a classroom pedagogy" (p. 57), but she also acknowledges that there is still much more to understand and calls for more studies of AfL practice in varying contexts. Little has been written about the application of AfL practice in language learning settings, particularly in higher education. Cumming (2004) writes of surprise at the small number of systematic analyses of assessment practices in language classrooms. Brindley (2007) reports, "To date, relatively few studies have been undertaken in language learning contexts that have investigated classroom assessment" (p.1), and further observes that "relatively little information is available" with regard to the assessment practices of tertiary English language teachers (p. 4).

In one of the few journal discussions of formative assessment in HE, Yorke (2003) lists some of the higher educational pressures which threaten formative assessment usage. Among others, these include: an increasing concern with standards of attainment and resulting emphasis on summative assessment; as well as the increase in staff/student ratios which consequently decrease attention given to individual students. While acknowledging these pressures, Yorke (2003) proposes that, "A major challenge for higher education is to respond to the main 'message' of Black and Wiliams' (1998) review [that] formative assessment is, after all, a key tenet of good teaching. " (p. 482). The research conducted and reported on here is one response to Yorke's (2003) HE formative assessment challenge.

Finally, in one of their more recent discussions of AfL in the classroom, Black and Wiliam (2006) propose areas of further exploration and research. These include, "the need to extend work of this nature to other groups, notably pupils in infant and junior school and *students in post-16, tertiary, and non-statutory assessment settings*" (p. 24, emphasis added). This thesis extends AfL research to other groups and contexts, and is the only work of this kind implementing the full range of key formative assessment procedures into classroom practice in a tertiary setting.

Part C: Research focus

As noted at the beginning of this introductory chapter, and reflecting a point about research foci made by Dunleavy (2003), this thesis involves taking established ideas (AfL) from one part of a discipline (mainstream education in primary/secondary schools) for a walk and putting them down somewhere else, applying them to a different context (a higher education