The evolution of the ejournal
Disseminate your research immediately through a new online, open access, peer reviewed journal

Online Educational Research Journal

This paper was downloaded from

The Online Educational Research Journal

(OERJ)

www.oerj.org

OERJ is an entirely internet-based educational research journal. It is available to anyone who can access the web and all articles can be read and downloaded online. Anybody can submit articles as well as comment on and rate articles. Submissions are published immediately provided certain rules are followed.

Assessing Assessment Theories

Abstract
Assessment is a ubiquitous process which dominates our lives in many insidious and covert forms in addition to the obvious ways: it is perhaps the central, as well as the most onerous and time-consuming aspect of many educationalists' work. It has dominated learning and teaching, although the study and research on assessment itself, particularly assessment theory, is often relegated to the realm of specialists. This paper wishes to persuade educationalists that engaging with assessment theory is not only essential for our understanding assessment, but also for co-ordinating and carrying out effective and equitable learning and teaching. By resolving central problems in assessment we can begin to realise its true place at the heart of supporting learning and teaching. Theory enables us to rationalise our arguments for practice and empirical research. As the key to providing a complete picture and understanding it is the cement which holds the house together. This paper wishes to provide an impetus for opening the discussion on issues in assessment which are manifest as dichotomies: formative versus summative; functions versus processes; formal versus informal assessment; formative assessment versus "Assessment for Learning". These will be examined with reference to central assessment discourses and theories.

Key words: assessment, theory, formative, summative, Assessment for Learning

1. Introduction

Practice and reflections on practice are insufficient to ensure optimum understanding of what we do. Although empirical research can help to clarify and support this (Lakoff and Johnson, 1999), it is theory which remains the single most encompassing aspect which will provide coherence and cohesion to our practices, arguments and research.

In the past 40 years great changes have taken place in learning and teaching, and a strange separation appears with assessment. Whereas the former has developed pedagogies according to learner and learning-centred rationales, assessment has not followed these logical developments and remained essentially teacher-centred. There are a number of reasons why this might be so. One critical reason is posited that the logical inclusion of learners in assessment has seemed a step too far in relinquishing the locus of power from tutors and transferring it to learners. Others are linked directly to assessment, which is the focus of this paper. Perhaps the necessary theoretical discussions which could bring learning, teaching and assessment together coherently have either not been considered available or not been convincingly engaging to the academic community. This could be due to a number of factors: theories of assessment and much of the work on assessment is often considered a specialist area; similarly, the literature on assessment is relatively limited, often confined to the ubiquitous chapter in learning and teaching books. Finally, the theories of assessment are little discussed and find little harmony within and across education communities (Taras 2009).

This paper wishes to set in motion a discussion about critical assessment issues which have been identified. It has two aims: firstly, to explore and evaluate current discussions of terminologies of assessment in order to better situate the relationships between concepts and thus the theories; and secondly, to discuss implications of these theories and terminologies for equitable and sustainable practice and research. It explores some of the most influential literature on assessment terminologies and theories to map ideas and premises which have been presented to the academic community. It examines definitions of formative and summative assessment and situates them within the wide understandings of roles and functions, and processes of assessment.

A number of unnecessary and unhelpful dichotomies are highlighted and the impacts that these dichotomies have on potential understandings of assessment. These are formative versus summative; functions versus processes; formal versus informal assessment; formative versus "Assessment for Learning". Without theory it is difficult to have coherent growth, development or understanding of what we do, how we do it and importantly, why we do it. "Assessment for Learning" (AfL) is used as an example to illustrate how limiting the potential for development can be if there is no solid theoretical foundation.

2. Theoretical Parameters

2.1. A note on the concepts of Theory and Practice

Practice is what we do in order to achieve an intended result or outcome. Empirical research collects data to try to ascertain if our practice and what we do a) actually does what we think it does and b) have the intended results which we would wish for. Theory provides the coherence and the logic for the practice; it provides the basic frame for rationalising empirical research and also, it stands alone as a logical and coherent rationalisation. Therefore, it links practice and empirical research into a coherent narrative. Without it practice would be a series of activities, without it empirical research would be simply a collection of data. Theory can also make claims of "truth" which not only stands alone but can be valid in different contexts to the same degree and in the same way. Therefore importantly, theory is generalisable on a conceptual level. This means that "theory" can be presented as the glue which (rationally and coherently) holds together practice.

2.2. Definition of Summative Assessment

Summative Assessment (SA) is normally envisaged as a formal assessment with shared criteria, outcomes and standards. It is a summation of a unit, course or programme of learning which is seen to take place towards the end and which is usually graded and part of an accredited unit. Issues of reliability and validity are linked to its formal nature (Black & Wiliam, 1998; Sadler 1989; Scriven 1967).

2.3. Definition of Formative Assessment

Formative Assessment (FA) is understood as assessment for feedback which provides support for learning. The value of FA is linked to it being provided in good time for learners to adapt their thinking and work: FA is believed to provide feedback which is separated from the emotional and sensorial aspects of a graded judgement. Increasingly, the understandings of FA are linked to informal drafts of work which are not linked to reliability and validity issues. In the AfL context it is often an ad hoc part of the classroom process (Black & Wiliam, 1998; Black & Wiliam, 2006).

These definitions are part of the general beliefs of SA and FA. When appropriate, these definitions will be refined and explained within my own personal beliefs and understandings which have emerged from evaluated processes of assessment and the literature.

2.4. Definition of Assessment (or Evaluation)

Scriven's seminal article (1967) entitled "The Methodology of Evaluation" discusses principles and processes of evaluation which need to be understood no matter what the context. This is clarified when he describes evaluation as a methodological activity and makes no differentiation in the process whether it involves a coffee machine or teaching machine. He is dealing in principles which are considered appropriate to any context. This bases the process within a theory as noted at the start of this paper.

In his article, Scriven uses the term "evaluation" where this paper will use "assessment" as a generic term to encompass both "evaluation" (generally referring to institutional and curriculum level) and "assessment". In education, the equivalent term in a micro context is "assessment" which is generally used when focusing on student work and processes. This current paper uses the term "assessment" as a generic term for any evaluation, assessment, judgement or decision.

Scriven describes the process of "evaluation" or "assessment" as requiring the gathering of data, establishing weightings, selecting goals and criteria in order to compare performances and justify each of these. In other words, to make a judgement we must decide on what elements are important, why these are important, how each element is important in relation to the others and finally, provide a justification of all the choices made. Assessment, according to this definition, represents a choice which may require an ethical justification. It is not just an opinion which can be laden with our own prejudices and feelings, but a process which at least attempts to provide accountability for the action and decision.

This is the working definition which is used in the paper and it provides the justifications which are critical in ethical assessment.

The results of an assessment and what we do with these are directly linked to the parameters which have been decided before the assessment took place. That is, the criteria have indicated what the important points of focus, the outcome(s) signal the purpose(s) or function(s) of the assessment and the standards demonstrate and provide guidelines to the level required of the work. Any deviation or change in these aspects either during or after the assessment should be signalled, explained and justified: this is necessary in order to have an ethical, *transparent* assessment.

Therefore, inherent in any definition of assessment is

1. the process
2. the specifications of the elements
3. the relation of all the elements to each other
4. the salient points, outcomes, purposes, functions, levels.

In other words, almost everything relating to the assessment is present. If these parameters of the assessment process are presented with the assessment, then even if the results are examined 10, 20 or 100 years later, it is possible to understand the process of the assessment and question any disagreement with any aspect. Therefore, to summarise, the definition will or should provide us with all the information pertaining to the assessment except for the contextualisation.

The contextualisation and contextual understandings are part of what are often referred to as belonging to or being part of a community of practice. There must be a flexibility in interpretation of the parameters and outcomes as long as there is an approximate alignment of contextual understandings, that is, educational, social, political, cultural and other contributing factors.

The required flexibility is to permit creativity and originality even when there are published or set criteria and outcomes. There is less flexibility in interpreting standards as these provide the basis of equitable comparisons for the judgements or assessments.

2.5. Roles or functions of assessment

In addition to clarifying the process of assessment, Scriven (1967) discusses the contentious aspects of the roles or functions of assessment. The question which relates to this is why do we assess? The plethora of available responses belies the simplicity of the question. It can be argued that assessment is one of the basic skills for survival: judgements enable an understanding of our environment, what we do, how we do it and how we can change and/or improve what we do.

Because of the ubiquitous nature of assessment, the relative simplicity of the process and the universal commonality of the needs to assess makes us all specialists and experts. However, developments in the past 40 years have provided contradictory and conflicting discourses which have confused the relative simplicity of assessment. As educationalists, we rationalise what we read and accommodate conflicting theories. However, this does not always result in felicitous understandings and can impact negatively on what we believe and what we do. Discussing contradictions can help to resolve these conflicts and this paper wishes to be part of this discussion.

One reason why the roles and functions are important to the discussions of theories of assessment is that they may displace our focus from the results of assessment which ensue from the processes of assessment, to the more tenuous and less controllable functions and uses. We need to know and understand judgements of our work but also how, when, why and contextual aspects can lead to assessment damaging perceptions of self and worth. However, focusing on these aspects clouds the primary necessity of assessment. The fact that feedback is often used to manipulate people and emotions or even as a power instrument does not negate its centrality. Can we control how this is used and how it is perceived? I would suggest only to a limited extent. This is the crux of the discourses around functions of assessment (Black & Wiliam, 2006; 2009; Scriven 1967; Taras 2005; 2009; 2010).

It can be argued that assessment is crucial to all education and to all learning. What we do with it is what is problematic. Functions are about what we do with assessment and how we do it. Intrinsically it should be a neutral process which is ubiquitous and indispensable to every aspect of life from learning to walk to adapting to new circumstances. This means that to a great degree, if we wish to ensure that assessments are ethical, then these need to be clearly contextualised and recorded.

When Scriven refers to the roles and goals of assessment, he is referring to how and why assessment can be used. But it is less about these functions which he is concerned about than the fact that we may be distracted from the process. Dilution of this process he claims makes it difficult for the process to provide the answers to the questions we require.

Furthermore, the fears associated with assessment have further glossed over the results which express "merit, worth, value" (Scriven, 1967, p42). When people criticise and condemn assessment they are in fact attempting to attenuate the negative perceptions, consequences or reactions which people may have. The assessment itself is a neutral process which weighs the evidence in order to provide an estimated judgement of the work. If this is not done, how can we know or judge what is good, what is weak and importantly, how to improve. The assessment provides feedback which is a key to improvement. The above discourse provides an approximate summary along with some implications which ensue from Scriven's seminal article.

What is particularly pertinent about the understandings of processes and functions is that it permits us to make the critical difference between what we do, how, why and when, and what we do with the results. Focusing on the results (which ironically is the summation of assessment), detracts from ensuring and checking that we have carried out the process ethically and precisely. The crux between process and functions is surprisingly little discussed and debated in the academic community (Taras 2009; 2010; Bennett 2011; Good, 2011). The differentiation between the two is one of or perhaps even the most important concept which drives assessment today. And yet, this concept is not even considered important enough to be debated in academic communities. It is generally accepted as fact that the distinction between SA and FA is based on functions without really understanding the consequences of this belief. The functions of assessment are often considered to distinguish between summative and formative assessment (Berry & Adamson 2011; Black & Wiliam, 1998; Black & Wiliam, 2006; Gardner, 2006 – all chapters). How we understand these terms

is central to assessment. It seems strange that we are so concerned about what we do with the results i.e. the functions, when we are cavalier about how these results were obtained i.e. the process: this mitigates against ensuring ethical and equitable practices. This is one of the first major problems for assessment from the literature.

2.6. Sacrificing process for functions

Scriven warned against losing sight of the processes of assessment as opposed to the functions: his original distinction between SA and FA were based on process. The change in the distinction between the two to that of functions originates in the work of Bloom et al., (1971). In the context of mastery learning, they used FA in their cyclical bite-sized aspects of learning to provide feedback and therefore support learning.

To begin with, functions cannot be controlled, limited or necessarily adhered to within any time constraints: that is, even if we decide that an assessment should be created, undertaken and graded with a pre-determined function there can be no guarantees of how this will be used, who will use it and when in the future the results of this assessment might be adopted retrospectively using different functions than those intended. Although we cannot control the functions of assessment we can control the processes of assessment and the parameters of the process, that is, the criteria, outcomes and standards which form the basis of the assessment. These are inseparably interlinked to the process and ensure that assessment can be not only controlled but justified and explained.

Furthermore, apart from focusing on functions, which are social, educational, political and often vague, there is another aspect of the SA and FA dichotomy based on functions which is problematic: focusing on functions has led to equating SA with formal exams and tests and equating FA with informal classroom processes. This is perhaps the second major distinction which has an enormous impact on our assessment practices and will be explored in the following section.

2.7. Formal assessment versus informal assessment

As noted, another problem which appears in the literature within the dichotomy of SA and FA is that SA has become linked with formal exams or tests and FA is equated to ad hoc classroom/work and feedback (Berry & Adamson 2011; Black et al., 2003; Black & Wiliam, 2006; Wiliam, 2007). This is problematic because SA work which is generally the focus of much hard work and time, and importantly, is the focus of validity and reliability issues, can be marginalised because feedback and the use of feedback for learning has become linked to FA. Furthermore, informal, ad hoc assessments are linked to learning as if issues of reliability and validity are not important for accuracy of feedback and for learning (Bennett 2011).

In this sense, recordings of classroom assessments which have no verifiable trail (paper or otherwise), are the most dangerous because they are not subject to reliability and validity constraints. For example, in external national exams, although all human processes and products are subject to error, there is generally a great deal of work done by various teams to ensure both reliability and validity. Generally speaking, we have to assume that despite human failing, there is a neutrality of intention in the results and that these are meant to reflect indirectly aspects of learning and achievement of the exam candidates.

However, in the case of classroom assessments, which are produced and recorded by teachers as a regular indication of student progress, they can and often do find their way to generalisations of student progress and ability. These can take the form of reports which could be seen by parents, head teachers and future class teachers. These have very little attention or scrutiny of issues such as reliability and validity because they are considered "informal". The consequences of assessment have been extensively documented and many are detrimental to individual self-esteem, perceptions of worth, future prospects and careers

(Berry & Adamson 2011; Broadfoot 2010; Stobart 2008). These consequences are generally attributed to SA, that is, exams or tests that "count". It is arguable that all assessments count and that all have a personal impact on individuals, particularly informal, regular classroom assessment, whether recorded or not (Bennett 2011). One reason among many is that these are the bread and butter of learning and are generally shared with the immediate learning community: an individual's identities, perceptions, self- and peer-beliefs often result from these interactions and assessments. Recognising the potentially detrimental impacts of classroom assessments does not in any way reduce the impact or importance of official external exams.

Functions of assessment are its social, political, educational and other uses which can be made of the results of the process. Isolating and maintaining one single use or functions from the plethora available is almost impossible, whether this is at the time the results are obtained or at some later date. We can control the process, that is, how we do something, but what we do or can do with the results is beyond our control.

Much of the AfL literature based it premise of the SA and FA distinction on functions (Berry & Adamson 2011; Gardner, 2006 – all chapters; Stobart, 2008). Recently, counter-arguments must have made an impression as they are being distinguished by the functions which they actually serve as opposed to the functions which had been decided prior to the assessment (Black & Wiliam, 2009; Wiliam, 2007). This argument also has limited mileage because of the difficulty in imposing limits on functions: also, they may be attributed retrospectively making it even more difficult to control them in the long-term.

Much of the generic aspect of Scriven's work, including the arguments presented above and particularly the transfer of these arguments to "assessment", that is evaluative processes related to the microcosm of the classroom, are often refuted for two main reasons: that these terms are not transferable from the more generic context of curriculum development in which Scriven placed his discourse, and the claim that the terms SA and FA as distinguished by Scriven do not apply to processes but to functions of assessment. Scriven's own words are a testament to which "reality" can be interpreted and is embodied in this paper. In addition, an obvious observation can be made about any work in the public arena which is open to scrutiny and interpretation by all: its uses and functions depend on each individual who engages with it and are not depended on the author's original intentions or wishes.

As noted in this section, two distinctions of SA and FA are problematic: one is the formal versus informal dichotomy and the second is that FA provides valuable feedback and SA does not. A further issue is that SA and FA are distinguishable by having different functions of assessment. An important challenge which remains for users and developers of AfL is to explain how SA and FA relate to each other and also how functions relate to processes of assessment.

3. Situating Assessment for Learning in Theory, Practice and Empirical research

Since AfL has been at the forefront of developments for the past 40 years, it will be used as an example to clarify difficulties which may arise when there is no solid theoretical foundation to support practice.

AfL developed from the desire of a community of academics to minimise the impact of external exams and tests in the classroom and to make classroom learning in schools a more creative and learning focused environment, hence the term "Assessment for Learning".

Black and Wiliam took the lead in spear-heading AfL in and after their 1998 seminal article "Assessment and Classroom Learning". This review article differed from others before it (Biggs, 1998) in that it separated FA from SA and focused on the former. How focus on FA should best be done was decided by examining the empirical research which had demonstrated proven results in supporting learning. This research was not without theoretical

support, therefore the best practices could be collated to best effect. This was the rationale behind the AfL interventions which were selected and trialled with teachers in the UK (Black et al., 2003). In addition, this also seems to be the rationale behind the AfL interventions which were selected, disseminate and trialled across the globe (Berry & Adamson 2011).

The theoretical aspects of AfL were thus believed to be in part subsumed within the empirical research, but more importantly, since AfL was subsumed within FA, AfL was presumed to have the weight of the theory of FA to support it. This might have been the case if AfL interventions had been clearly integrated and framed within firstly, the assessment process and secondly, within FA processes. However, since explicit links with the processes of assessment were not made because the premise of the difference between SA and FA was according to functions, this did not occur. The assumption became a conceptual leap and AfL was to some degree isolated from both assessment and learning theories (Taras 2005; 2007; 2009; 2010). Subsequent work, particularly by Black and Wiliam (Black & Wiliam, 2006; Black & Wiliam, 2009; Wiliam, 2007; Wiliam, 2009) attempted to address the issue of lack of situatedness of AfL within pedagogic theories as signalled by Perrenoud (1998). However, this did not address an even more important aspect of AfL which is to situate it within assessment theories and processes. Each intervention remained a self-contained unit. Stobart recognises the problems inherent in this and that AfL has resulted in being seen as "a series of classroom 'handy hints', rather than a theory-driven approach to teaching and learning" (ibid, p149). He further notes that

> "This does not mean that there is no theoretical underpinnings; simply that it has not been organised, and may not need to be, into a stand-alone theory" (ibid, p145)

After 40 years of AfL, and despite world-wide research on classroom practice it seems surprising that Stobart believes that "…our understandings are still at an early stage" (Stobart, 2008 p145).

In retrospect, it seems a glaring omission not to situate AfL within assessment theories and processes especially since three of the four AfL interventions which have been used world-wide are in fact aspects of feedback which come directly from SA work, that is feedback through marking, peer and self-assessment and the formative use of summative tests. This is perhaps the ultimate irony of the work around AfL: that the turning from SA which first inspired its development, has become its nemesis which has plagued it on a number of levels. In fact, as will be demonstrated in this paper, the very exclusion of SA has been the most detrimental aspect of AfL.

3.1. Where is the theory in AfL?

As noted, AfL developed out of the laudable aims of academics to refocus classroom practices in order to support learning and to move away from teaching to exams. The presentation of the AfL interventions and how they were developed and brought into the initial classroom trials although interesting cannot be dealt with in detail here as it is beyond the scope of this paper (see Black et al., 2003; Taras 2009 for further discussions). However, one salient aspect which emerged and will be used as an example, is that the understanding of criteria could not be isolated from the assessment contexts which was the original decision. Therefore, explicit criteria discussions became necessary and integrated as part of all of the interventions. This would seem to have been a perfect opportunity for the academics involved to question why their initial assumptions concerning being able to isolate criteria from specific assessments had had to be re-adjusted and importantly, how this could be rationalised within assessment theories. None of this seems to have been done in any of the literature associated with AfL.

Since criteria (whether explicit or implicit) are one aspect of a number of parameters necessary to carry out assessment, it is logical that they should be integrated into every

'assessment' exercise. The fact that an exploration of a generic concept of criteria is not sufficient to transfer into specific examples of assessment is an important discovery in itself. This would seem to signal that learners require an explicit focus and discussion to explore individual assessments: this logically leads to the importance of explicit, clear parameters and mitigates against implicit (and often linked to informal) assessments as a short-cut. Short-cuts which are evident in informal, ad hoc assessments would seem to show that this is not helpful to learners and learning.

It can be argued that clarification of how theory (and what theory) can coordinate and rationalise processes and thinking was missing from the developments of AfL interventions. Separating FA from SA has further distanced the AfL interventions from theory because it removed it from the scrutiny of an explicit assessment process. The question "where does feedback come from?" was never asked. The original assessment step (of a summation of opinion at any given stage) was excluded because of potential links to SA. Therefore, the implicit first step of assessing remained implicit and the process of assessment became implicit and reinterpreted as classroom pedagogic processes.

Within the AfL discourses, the scrutiny of practice and the collection of empirical data has been consistently prioritised. It has been demonstrated that education departments at university in addition to teacher researchers are under pressure to collect evidence to support practice to the detriment of theory (Tight 2004). However, as argued earlier, without theory, we cannot justify why we are doing something. It is also difficult to improve and develop ideas if the "why" of theory is missing.

In addition, the same criticism continues to be made about AfL, that is, that the interventions are often relegated to handy hints and mechanical activities in many classrooms and these have been acknowledged by its most ardent supporters (Stobart 2008). This would seem to be borne out by the fact that, apart from the initial changes of integrating criteria into all the interventions, there appear to have been no alterations to the original interventions. If the interventions have remained the same, the opinions of central promoters of the movement, the discourses and supporting rationales for AfL have changed quite radically, have been strangely illogical and do not lead to resolving any of the above mentioned problems.

The major players in discussing and developing AfL theory seem to agree on the following:

- That AfL is a movement which presents principles for using assessment to support learning in classroom interaction
- SA and FA are differentiated according to functions and that AfL is supported by FA
- AfL and FA became increasingly synonymous and interchangeable as concepts
- AfL and FA became increasingly associated with informal classroom practices
- That AfL was a-theoretically represented though not lacking in theoretical links

Due to the last point, a number of publications ensued which claimed to provide the theoretical framework for AfL, however, most of these did little more than reiterate the beliefs signalled in the previous bullet points (Berry & Adamson 2011; Black & Wiliam, 2006; Black & Wiliam, 2009; Gardner, 2006 – all chapters; Wiliam, 2007; Wiliam, 2009).

For example, one informative rationalisation and explanations in relation to AfL's "theoretical" developments can be found in major leaders in the field, Black and Wiliam. In their 2009 paper ("Developing the theory of formative assessment", whose title echoes that of the 2006 paper in that the only difference is replacing the indefinite by the definite article), they provide an evaluative descriptions of the focus of their previous major works. In the 2009 paper they systematically admit that theory was neither a focus or a priority of their previous work. They clarify that the lack of theory or theoretical support on work had stemmed from the fact that they had systematically prioritised practice and how this practice could be improved by other empirical evidence. This evaluative list reads likes a very strange confessional which has neither been exhorted nor which justifies their previous and

continuous lack of engagement with theory. More surprisingly still, the article repeats the pattern of claiming that they are examining theory when in reality this is not the case.

3.2. Formative Assessment or Assessment for Learning?

The seminal article in 1998 reviewed FA interventions which had been shown to support learning. Four were subsequently chosen and became the basis for the AfL interventions from which the concept of AfL was developed. These were thus a series of practices which would bring innovative learning interactions supported by research into the classroom. The concepts of FA and AfL have gradually become indissociably linked and increasingly synonymous (Broadfoot, 2008 p216; Gardner, 2006 p197; Harlen, 2006 p103; James, 2006 p49; Stobart, 2008 p16; Wiliam, 2007 p1054; Wiliam, 2009 p6, 7).

After the split between FA and SA, this linking of FA and AfL as if they were one and the same has perhaps had the most detrimental impact on both concepts. One reason is that the weight of the theoretical, empirical and practical literature and discourses of FA were adopted into the service of the new, untried and undeveloped AfL.

To confuse the understanding of FA and the links of AfL to theory further, there are two definitions of FA which have surrounded the AfL framework. These fall into two categories: one is based on Sadler's theory of formative assessment (1989) and focuses on product assessment (Black, 2003c p2**;** Black et al., 2003 p15, p121; Wiliam, 2000 p15). The other is based of the understanding of formative assessment as a classroom learning and teaching pedagogy process (Black, 2003a, b, c; Black et al., 2003 p2; Wiliam, 1994; 2000; Wiliam & Black, 1996 p8).

Having two different incompatible definitions for a single term contradicts what was described as theory at the start of the paper. As noted previously, theory provides the rationale at the level of principles. Theory is generalisable at a conceptual level to provide logic and coherence of rationalisation to support practice and research. Contextual differences should not impact on theory as it deals with principles. These two definitions contradict the definition of theory.

Over time, criticisms of this duality (Taras 2009) must have been considered because increasingly discourses have separated AfL from FA. FA has become a teacher led and teacher-centred means of getting feedback from classroom processes and learners in order to adapt or change the teachers' learning and teaching strategies. Therefore, AfL is about learners and their role in the classroom despite using assessment interventions and FA is about teachers and how they improve their teaching (Black & Wiliam, 2009).

This however, becomes even more problematic as FA has a long tradition both within HE and in the compulsory sector. Changes in discourses take place constantly, however, within the academic community, it is accepted practice to place current work within current discourses and contextualise and justify changes. Breaking the rules of academic protocols does nothing to aid sharing of ideas within the established community nor does it help to integrate newcomers (Taras 2013). This is particularly important in the context of education departments and teacher education, where a relatively narrow choice of literature tends to be the norm.

3.3. ***Feedback, the missing link between assessing and learning***

A question remains as to whether AfL is assessing or learning? It is perhaps useful to begin by looking at feedback. Both learning and assessing often use feedback as a starting point. Feedback fuels learning and the quality, appropriateness, timeliness and engagement with it has been acknowledged and demonstrated as important factors which influence its ability to do so. However, more problematic and less explicit and focused is the process which leads to the production of the feedback. The further back from the product of feedback one goes, the

less explicit and clear the process tends to be. This process, whether analysed in the direction of provenance, or of destination, would lead to the basic principles which inform both the process and the product of feedback.

Examining this process in more detail will help to demonstrate the essential links between learning and assessing. Let us begin with product feedback because this is often the main focus. Assuming, for expediency's sake that this feedback is appropriate in quality and timely, by working backwards we will have used pertinent criteria which are appropriate for the context, support learning outcomes, programmes, courses and aims and so are relevant for the context. The criteria are linked to standards and both are context efficient. Therefore, the parameters which select and decide the criteria and standards for the feedback have been in place prior to the production of the feedback even though they may often be implicit rather than explicit and shared with all the participants.

So, what is this process which uses criteria, standards and contextually specific factors of learning outcomes and aims? This process is assessment. Since this is obvious, why is the term assessment (or evaluation, testing or any other term which may be used to describe this process), so rarely used to link feedback with learning. Taras (2007) examines linguistic factors to explain why terms linked to assessment are so emotionally and academically problematic. Stobart (2008) and Broadfoot (2007) focus on social, political and historical factors which have made assessment problematic throughout the ages. Most people would agree that these works demonstrate that assessment cannot be viewed as either a neutral or a context-free process and that although a universal and ubiquitous process it is often marginalised into disparate and isolated realms of denial and resistance to its existence.

Therefore, the question which remains to be asked is why the many AfL developers have felt torn by the choice between assessing and learning? Why, if feedback is the linking element to both and it cannot exist without explicit or implicit assessment, are they so adamant in their later work (particularly Black & Wiliam, 2009; Wiliam, 2007) that AfL is situated in the realm of learning and that their particular paper excludes AfL as product assessment of learners' work? This when both the AfL interventions and the framework of AfL (and/or formative assessment) explicitly include product assessment of learners' work (Black & Wiliam, 2009 p24-5; Wiliam 2007, p1064). This question is further reflected in their reaction to external critiques of their work, to which they seem to have responded if it is linked to learning, but which they seem to have ignored if linked to assessment.

For example, the reproach made by Perrenoud (1998), that in their 1998 seminal paper they had not situated AFL within theories of learning. This they have addressed and referred to in a number of articles, more specifically Black and Wiliam 2006 and 2009. Their concern to respond to this critique makes it even more surprising that they have not acknowledged or responded to Biggs' (1998) critique that they have separated formative assessment (and by extension AfL) from summative assessment. Indeed, much of their work has painstakingly demonstrated to their satisfaction that formative and summative assessments are often irreconcilable and at odds with each other (see papers in Gardner 2006). This despite feedback from teachers' who participated in one dissemination project for AfL refusing to separate summative and formative assessment (Black et al., 2003 p31) and secondly, the fact that they use Sadler's (1989) work to support their discourse on feedback.

Many of the problems and issues arising from the lack of theory to support AfL practices can be explained by the lack of explicit links between learning and assessing and also by isolating feedback from the assessment process. The links between summative, formative, self assessment and feedback are framed within an epistemology which is based upon the processes of assessment. Understanding assessment through a scrutiny of the process demonstrates that any basic assessment process begins with an explicit or implicit setting of assessment parameters which relate to the context. Making the processes and parameters

explicit and situating these within a social-constructivist theory of learning enables more ethical and transparent assessment procedures. It also situates feedback within a social-constructivist epistemology of assessing/learning because it defines feedback from the perspective of learners who are required to understand, engage with and incorporate feedback into the original assessed work. This means that the feedback is used to update and improve the work as opposed to remaining as information (or knowledge of results). Within this learning theory and epistemology, assessing and learning are interactive and interacting aspects of a single process which invariably links the different stages and steps of the process. Also, the distinction SA – FA only has any real meaning, particularly in an educational context, if FA belongs to the learner. Learners are the ones who learn and who ultimately make all the decisions about learning, whether consciously or unconsciously. Anyone and everyone can and does carry out SA which is an assessment or evaluation. It is FA which has become problematic because despite discourses of learner and learner-centredness, learner autonomy and independence and empowering learners, it is still taboo to open assessment to learners. But that is another story for another paper.

4. Conclusion

This paper has explored discussions and discourses of assessment. Theory is the glue or cement which holds together our reasoning, ideas and our practices: without it we can still function and develop, but it can be argued that developments are more coherent and conducive to logical analysis and improvement if theory is there to support practice and give substance to empirical research. By exploring basic definitions and understandings of assessment and how the terms interrelate, it has been possible to map our discourses and pinpoint important issues for discussion within academic communities.

The importance of explicit processes as a basis for all assessments has been signalled. Although functions of assessment tend to dominate our lives, particularly social and political functions which may be prioritised over education ones, by prioritising processes it is possible to ensure we can monitor and maintain equitable, logical assessments which can begin to deliver what they promise. A major problem which has been identified in relation to this concerns the volatile and transient nature of assessment if the focus is on functions and not process. Since functions are impossible to control and sustain, then they can mislead. Logically, the focus should be on what we do and how we ensure that practice is explicit and within a shared understanding in order to produce defensible results: concerns about what we do with these results should be the second stage in the process.

There is another aspect which has been identified as being problematic which is also linked to the dichotomy of SA and FA being based on functions. Focusing on functions has inadvertently led to equating SA with formal tests and equating FA with informal classroom processes. This is problematic because the formal versus informal assessment distinction has led to placing impossibly heavy expectations on FA. FA cannot deliver if instead of providing coherent, valuable and valid feedback it becomes a haphazard, informal, unregulatable process which is a prioritising of timing over reliability and validity issues. FA should not be reduced to an ad hoc classroom process when it has the critical task of supporting learning. The quality of the feedback has been shown to be more important than the quantity in support learning.

A salient, current example to demonstrate how lack of a systematic and strong theoretical foundation can destabilise our innovative practices is AfL. With hindsight, it is easy to see that separating AfL from assessment theories and processes has been detrimental to providing AfL with the credible theoretical support it needs. This is particularly true because three of the four AfL interventions which have been used world-wide are in fact aspects of feedback which come directly from SA work. These are: feedback through marking, peer and self-

assessment and the formative use of summative tests. Therefore two important aspects which problematise AfL are that assessment, and more specifically, SA has been increasingly excluded from its discourses. This leaves the assessment practices in search of a theory when the theory is there all along. It is perhaps the ultimate irony that targeting and attempting to eliminate negative pressures of external assessment from the AfL classroom has led to SA becoming its nemesis by its very exclusion. The very exclusion of SA theory has been one of the most detrimental choices for AfL.

To conclude, despite the problems which have beset assessment theory these past 40 years, many can be resolved within the current literature. Importantly, they can support innovative assessment practices and form the rationale for research. AfL can form part of innovative practices of assessment and be supported by assessment theories if the academic community is willing to discuss the possibility.

References

Bennett, R.E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18 (1), pp.5-25.

Berry, R. & Adamson, B. (Eds.). (2011). Assessment reform in education: policy and practice. Dordrecht: Springer.

Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education: Principles, Policy & Practice*, 5(1), 103-110.

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

Black, P., Harrison, C., Lee, C., Marshall, B., Wiliam, D. (2003). *Assessment for learning. Putting it into practice.* Maidenhead: Open University Press.

Black, P. (2003a). Turning Research Results Into Practice: How does the D Fit into the R&D? Paper presented at AERA Chicago 21 April 2003. Mathematics Education SIG Session 20.061. http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html (Accessed 14 May 2003).

Black, P. (2003b). (with the King's College London Assessment for Learning Group Harrison, C., Lee, C., Marshall, B., Wiliam, D.) The Nature and Value of formative Assessment for Learning, Paper presented at AERA Chicago 22 April 2003, Presidential invited session 34.011. (Accessed 14 May 2003)
http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html.

Black, P. (2003c). (with the King's College London Assessment for Learning Group Harrison, C., Lee, C., Marshall, B., Williams, D.) Formative and summative Assessment: Can They Serve Learning Together? Paper presented at AERA Chicago 23 April 2003. SIG Classroom Assessment Meeting 52.028. (Accessed 14 May 2003)
http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html

Black, P. & Wiliam, D. (2006). Developing a theory of formative assessment, in: Gardner J. (Ed) *Assessment and learning.* 81-100. London: Sage.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.

Broadfoot, P. (2008). Assessment for learners: Assessment literacy and the development of learning power, in: A. Havnes, & L. McDowell (Eds) *Balancing Dilemmas in Assessment and Learning in Contemporary Education.* 213-224. New York/London: Routledge,

Gardner, J. (Ed.). (2006). *Assessment and learnin*g. London: Sage.

Good, Robert. (2011). Formative *Use* of Assessment Information: It's a Process, So Let's Say What We Mean. *Practical Assessment, Research & Evaluation*, 16(3). Available online: http://pareonline.net/getvn.asp?v=16&n=3.

Harlen, Wynne. (2006). On the Relationship Between Assessment for Formative and Summative Purposes, in: Gardner J. (Ed) *Assessment and learning.* 103-118. London: Sage.
James, Mary. (2006). Assessment, Teaching and Theories of Learning, in: Gardner J. (Ed) *Assessment and learning.* 47-60. London, Sage.
Lakoff, G. and Johnson, M. (1999) *Philosophy in the Flesh: the embodied mind and its challenge to western thought*. New York, Basic Books.
Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning, *Assessment in Education: Principles, Policy and Practice*, 5(1), 85-103.
Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*. 18, 145-165.
Scriven, M. (1967). The Methodology of Evaluation**,** in: R. Tyler, R. Gagne, & M. Scriven, (Eds) *Perspectives on Curriculum Evaluation* (AERA Monograph Series – Curriculum Evaluation). 39-83. Chicago, Rand McNally & Co..
Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. New York/London: Routledge.
Taras, M. (2005). Assessment, Summative and Formative – some theoretical reflections, *British Journal of Educational Studies*. 53(3), 466-478.
Taras, M. (2007) Assessment for Learning: understanding theory to improve practice *Journal of Further and Higher Education,* 31(4) 363-371.
Taras M. (2009) Summative Assessment: the Missing Link for Formative Assessment *Journal of Further and Higher Education*, Feb 33(1), 57–69.
Taras, M. (2010). Back to Basics: definitions and processes of assessments. *Revista Práxis Educativa*, 5(2), 123-130.
http://www.revistas2.uepg.br/index.php/praxiseducativa/issue/view/162/showToc
Taras, M. (2013 in press.). Feedback on Feedback: uncrossing wires across sectors**.** In: Merry, S., Price, M., Carless, D. and Taras, M. (Eds.). *Reconceptualising Feedback in Higher Education*, London and New York: Routledge.
Tight, M. (2004). Research into higher education: an a-theoretical community of practice? *Higher Education Research and Development*, 23(4), 395-411.
Wiliam, D. (2000). Integrating Summative and Formative Functions of Assessment
Keynote address to the European Association for Educational Assessment, Prague: Czech Republic, http://www.kcl.ac.uk//depsta/education/hpages/dwliam.html (Accessed 15 May 2003).
Wiliam, D and Black, P, (1996). Meanings and Consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22,(5) 537-48.
Wiliam, D. (2007). Keeping learning on track: classroom assessment and the regulation of learning, in: F. K. Lester, (Ed) *Second handbook of mathematics teaching and learning.* 1053-1058. Greenwich CT: Information Age Publishing.
Wiliam, D. (2009). *Assessment for Learning: Why, What and How.* Institute of Education: University of London.