Curriculum Journal

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713695259

Classroom assessment within the alternative assessment paradigm: revisiting the territory
Michael A. Buhagiar a
a University of Malta,

Online Publication Date: 01 March 2007
To cite this Article: Buhagiar, Michael A. (2007) ‘Classroom assessment within the alternative assessment paradigm: revisiting the territory’, Curriculum Journal, 18:1, 39 - 56
To link to this article: DOI: 10.1080/09585170701292174
URL: http://dx.doi.org/10.1080/09585170701292174

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Classroom assessment within the alternative assessment paradigm: revisiting the territory

Michael A. Buhagiar*

University of Malta

Assessment reform has been on the educational agenda of many countries for at least two decades. In this article—which begins by charting what lies behind the calls for a paradigm shift in assessment and what is being proposed instead—classroom assessment is explored in detail in relation to its links with teacher and student assessment, as well as with formative and summative assessment. The emerging notion that classroom assessment embodies all forms of assessment that take place within the confines of the classroom is subsequently analysed from an ‘assessment for learning’ perspective, which is increasingly being accepted as the way forward if teaching, learning and assessment are to become fully integrated as demanded by the underlying philosophy of the new, alternative paradigm. However, noting the gross difficulties facing the translation of new policies into classroom practices, the article finally suggests what can be done to rectify this certainly frustrating, but also potentially dangerous, situation.

Keywords: Alternative assessment paradigm; Assessment for learning; Assessment reform; Classroom assessment

Rationale for change

Assessment developed historically for the purposes of selection and certification—particularly, selection for further educational opportunities beyond the minimum state provision and for employment (Torrance, 1995). The pressing need to find a mechanism of selection that would be socially acceptable and would identify the ‘best’ candidates led to a premium being put on assessment techniques that appeared to be fair and objective, and had high levels of reliability (Broadfoot, 1995). The ultimate scope—which is still a common perspective—was to select the most ‘talented’ for the most educational investment in the mistaken conviction that these students would later be able to ‘put most back’ into the economy (Torrance, 1995).

*Department of Mathematics, Junior College, University of Malta, Msida MSD 10, Malta.
Email: michael.buhagiar@um.edu.mt
What was practically an economic decision led to the triumph of examinations—an assessment device that, as Nuttall (1975) recalls, traces its origins to an examination that was imposed in the first millennium BC on aspirants competing to join the Chinese Imperial Service in an effort to stamp out nepotism and bribery. Many centuries later, the notion that examinations promote equal opportunities still carries much weight (Gipps & Murphy, 1994). As pointed out later in the article, school practices remain in fact largely conditioned by the traditional vision of assessment characterized by tests, examinations, selection, grades and marks.

However, assessment experts have for the past couple of decades recognized and argued in favour of assessment as an essential component to learning. This theoretical shift away from ‘measuring’ learning and towards assessment that is explicitly designed to promote learning came primarily in response to our growing understanding of learning as a meaning-making process in which, contrary to our prior understanding that knowledge can be passed directly from one head to another, much depends on the learner’s constructions of his or her own experiences. Another contributing factor was the realization of the multiple ways in which the still ubiquitous presence of the traditional forms of assessment—namely, tests and examinations—in contemporary educational systems (see Broadfoot & Black, 2004) affects negatively the teaching–learning environment. This form of assessment is definitely ‘unfriendly’ on a number of counts:

- **Learning unfriendly**: traditional tests and examinations may reveal what students remember about what we think they should remember, but do not help us get to truth, meaning, purpose or utility (Ellis, 2001). They basically represent a narrow (only assesses a very narrow range of academic qualities), sudden death (candidates get credit only for what they do during the brief examination period) and a non-informative (the grade does not define what a candidate knows and can do) form of assessment (Desforges, 1989).

- **Curriculum unfriendly**: examinations emphasize recall of factual knowledge with a heavy reliance on memory and rote learning (Gipps & Stobart, 1993). They consequently can lead teachers to coach their students in a narrow range of test-taking skills rather than encourage them to teach a broader range of higher order competencies and understandings (Gipps & Stobart, 1993; Torrance, 1995). As a result, students develop problem-solving strategies that help them pass examinations without reaching the intended learning goals (Brookhart, 1999).

- **Teacher unfriendly**: the knowledge that test scores are the sole means of describing and judging schools produces in teachers the determination to do what is necessary—which translates itself in a ‘testlike’ teaching programme—to avoid being publicly embarrassed and humiliated (Smith, 1991). Many teachers may in fact even be willing to engage in questionable, educationally indefensible practices to increase their students’ test scores (Popham, 1991).

- **Student unfriendly**: students find themselves not so much described by examinations as constructed by them (Hanson, 2000). An unwelcome by-product of this ‘intrusive’ process is labelling, which can set an unconscious limit on what
students are perceived as being able to do. These teacher expectations, apart from allowing the ‘weaker’ students to achieve a great deal less work than the ‘better’ ones, lead to the realization of the self-fulfilling prophecy in students (Rowntree, 1987). Traditional assessment, instead of motivating all students to work harder and encouraging them to carry on learning, has a demotivating influence on most of them and effectively pushes them out of the system (Broadfoot, 1996).

Although these characteristics belie the traditional notion that examinations promote equal opportunities through the provision of fair, objective, reliable and precise measures of achievement (see Murphy & Torrance, 1988; Gipps & Murphy, 1994; Broadfoot, 1996), the examination system continues to enjoy pervasive support among the general public (Broadfoot, 1996). This persistent support may well have to do, at least in part, with the relative ease with which examinations can be administered and their results understood (see Brown, 1990; Gipps & Murphy, 1994; Broadfoot, 1995).

In spite of this public acceptance the fact remains, however, that the traditional assessment model hinders learning in multiple ways. The very essence of traditional assessment, especially within the confines of the classroom, repudiates the foundations of a quality education for all, that same notion which is increasingly finding its way in educational discourse and policy documents in the understanding that it would ultimately benefit both individuals and society at large.

If we truly believe in inclusion and diversity—which builds on the understanding that everyone is capable of learning and worthy of the best possible investment in his or her education—it becomes unsustainable to continue using an assessment model that has traditionally developed to focus on selection, certification and accountability. In particular, the recognition that assessment is now required to achieve a wider range of purposes than this (which includes supporting teaching and learning, providing information about students, teachers and schools, and driving the curriculum and teaching) has increasingly rendered the traditional model underpinning assessment theory an inadequate framework and has necessitated the development of a new theory to further our understandings of, and practices in, educational assessment (Gipps, 1994). It is to this new reconceptualization of assessment, which I find so educationally promising, that I now turn my attention.

The alternative assessment paradigm

When still in its infancy, Murphy and Torrance crystallized the concerns and hopes of the alternative assessment paradigm:

Our own predominant concern is not for assessments to be psychometrically pure and reliable, but for them to play a constructive role in the educational process and as a part of that role to provide valid information about educational achievement in the fullest sense of the meaning of that phrase. Such achievement is not that which is easily measurable, but that which is desirable in terms of the broad aims of those concerned with what children
gain from the process of education. (Murphy & Torrance, 1988, p. 100; emphasis in original)

Their comments heralded a new emphasis on assessment procedures that provide a useful picture of what students know and can do. It was a way of saying that assessment should reward holistic learning and encourage qualitative understanding (see Hildebrand, 1996). From this perspective, assessment ‘has a constructive focus where the aim is to help rather than sentence the individual; thus it emphasizes the individual’s achievement relative to him or herself rather than to others, or in relation to defined criteria’ (Gipps & Murphy, 1994, p. 261). Implied within this new understanding lies the notion that the purpose of assessment has changed from categorizing students for assignment to pre-determined curricular and instructional programmes to tailoring instructional programmes to learners’ individual needs and to connecting learners and groups of learners in mutually beneficial learning experiences (LaCelle-Peterson, 2000).

This necessitates a focus of attention on each student’s learning, as everyone is believed to have his or her own particular requirements. Rather than unrealistically seeking equality of outcomes and the provision of identical experiences for all, equity in assessment is thus about assessment practices and interpretation of results that are fair and just for everyone (Gipps & Murphy, 1994). The embedded emphasis on the realization of everyone’s potential reflects in turn the current repositioning of assessment in relation to learning—that is, a change from being a ‘measure of learning’ to becoming a ‘support to learning’.

Assessment in the new paradigm is in fact no longer seen as a scientific or objective activity. In recognition of the postmodern notions of multiple realities, subjectivity and knowledge construction, it is perceived instead as ‘an inexact matter and can never be an exact one’ (Harlen, 1994b, p. 139). This shift from the psychometric model of assessment to the educational model draws on the postmodern condition that requires a suspension of belief in the absolute status of ‘scientific’ knowledge (see Gipps, 1993). This means that no matter how much we try to calibrate the ‘measuring’ instrument, we can still never know what is inside a student’s head. Assessment can instead only tell us what the student can do in particular circumstances.

Keeping in mind that domains and constructs are multidimensional and complex, that assessment is not an exact science, and that the interaction of student, task and context is sufficiently complex, we still cannot know what the student can do in other circumstances. This renders the generalizability of assessment to other tasks and contexts limited, if not dubious. As a consequence of the decline of external normative constraints, there is now more need for individuals to discipline themselves through internal mechanisms and a corresponding search for ways in which this may be achieved (Winch & Gingell, 1999). This calls for what Ellis terms ‘reflective assessment’ that requires practitioners, including students, to engage in a metacognitive assessment process in order to foster ‘a practical sense of what works, what is good, what has meaning, and ultimately, why’ (Ellis, 2001, p. xv).
The embedded denial within the new paradigm of the existence of a ‘true score’, however, reconceptualizes rather than bans the use of tests and examinations in assessment. Constructivist theories demand, in fact, that tests show what students know and can do, as well as facilitate good learning—what Glaser (1990) calls ‘placing tests in the service of learning’. Within this emerging framework, tests should consequently be ‘ambitious instruments aimed at detecting what mental representations students hold of important ideas and what facility students have in bringing these understandings to bear in solving their problems’ (Shepard, 1991, p. 9).

The retention of tests evidently builds on the fundamental notion that ‘assessment does not stand outside teaching and learning, but stands in dynamic interaction with it’ (Gipps, 1994, p. 15). It is all about fostering a much more comprehensive system that supports the appropriate use of multiple methods of assessment that are needed to cover the full range of achievements targeted—namely, knowledge, thinking, processes, products and dispositions (Gipps, 1994). The understanding is that the multiplicity and variety of assessment practices provide higher quality (as the strengths in one source compensate for the weaknesses in others) and fairer information (as all assessment methods may be said to have a certain amount of bias).

Fundamental to the new paradigm is the growing realization that ‘assessment is more than just a technical activity; it is a human activity that influences and affects many people’ (Airasian, 2000, p. 22). With students being the most vulnerable in this respect, it becomes imperative to always strive to obtain trustworthy information before making important decisions that can influence them. Moreover, once the information is collected, there is a responsibility to protect its privacy, to recognize its decision-making limitations, and never to use it to demean or ridicule a student. These ethical standards form an integral part of our reconceptualization of assessment that, as Broadfoot (1996) points out, is driven above all by the desire to channel the powerful impact of assessment to promote, rather than inhibit, learning.

This understanding explains why quality in assessment now depends mostly on the impact of assessment on learning rather than the accuracy of assessment per se (Broadfoot, 1996). The emerging idea that the indicator of the quality of educational provision is the individual students’ learning outcomes has now made it necessary to begin considering the traditional notions of reliability and validity in relation to the contexts and purposes of assessment. Thus, for instance, within the classroom, where the teacher wants information about students on a regular basis with minimum interruption of normal work,

quality in assessment means an assessment made and interpreted on the spot which provides the type of information required (high validity) and with the greatest degree of reliability possible in the circumstances. (Harlen, 1994a, p. 13)

This distancing from the traditional conceptions of reliability and validity is again evident in Gipps’s (1994) understanding of quality assessment, which she defines in
terms of ‘trustworthiness’. The qualities of a trustworthy assessment include: (a) **credibility**—which comes from regular ongoing assessment in the classroom, and the inclusion of parents in the assessment dialogue; (b) **transferability**—which requires the assessor to specify the context in which a particular achievement is demonstrated, so that others may judge whether this is transferable to other contexts; (c) **dependability**—which makes the assessment process open to scrutiny, and subject to an audit process of quality control; and (d) **authenticity**—which depends on the extent to which the relevant constructs are fairly and adequately covered in the assessment.

**Classroom assessment**

Assessment in education spreads over a number of closely knit scenarios that vary primarily across a continuum of formality:

> At one end of a dimension of formality, the task may be normal classroom work and the process of gathering information would be the teacher reading a pupil’s work or listening to what he or she has to say. At the other end of the dimension of formality, the task may be a written, timed examination which is read and marked according to certain rules and regulations. (Harlen *et al.*, 1992, p. 217)

The scenario that is directly relevant to this article concerns assessment that—irrespective of the level of formality, the assessors and the type of tasks involved—originates inside the classroom as opposed to outside it. In the following sections, apart from exploring this specific form of assessment, which I am calling ‘classroom assessment’, I also examine the implications that the alternative assessment paradigm bears upon it in order to be able then to better understand how the present assessment situation inside classrooms compares against the criteria of the new paradigm.

**Defining classroom assessment**

I like to think of classroom assessment—which I see as an umbrella term that transcends, albeit with changed emphases, the different assessment paradigms—as the taking of the temperature of the teaching–learning environment inside the classroom. For me, it is the link between the assessment process and the place from which it originates (i.e. the classroom) that defines the term rather than ‘the why, when and how the temperature is taken’ or ‘who takes and interprets the temperature to then decide on how to act’.

It follows that although I am using a term that is much used in the American literature (e.g. Phye, 1997a, 1997b; Bright & Joyner, 1998; Bryant & Driscoll, 1998; Brookhart, 1999; Popham, 1999; Airasian, 2000; Anderson, 2003), my understanding of it is slightly, albeit significantly I believe, different from the manner in which it is generally portrayed, even if only by implication.

The recognition of student agency in learning (see Murphy, 1996), which has been a key element in the development of the new assessment paradigm, makes me critical
of the numerous attempts in the literature (e.g. Webb, 1998; Anderson, 2003) to portray, possibly unwittingly, the teacher as the point of reference in classroom assessment at the almost complete exclusion of students. In my view, to accept this would be like saying that classroom assessment is the same as teacher assessment.

While this might have been possible to sustain within the traditional paradigm, my understanding that teacher assessment is ‘Assessment made by teachers of pupils’ attainment, knowledge and understanding’ (Gipps, 1994, p. 123) encourages me to argue instead that although teacher assessment is undeniably an essential component of classroom assessment, it is not comprehensive enough to capture all the learning possibilities within the classroom offered by the alternative assessment paradigm. Indeed, according to this paradigm, it is only when students engage in self- and peer assessment—which I see as complementing teacher assessment inside the classroom—that assessment can be truly said to support their learning as opposed to inhibiting it.

Again, even though the rationales of teacher assessment and assessment by students are both linked to the constructivist model of learning (see Gipps, 1994)—and are meant as such to reflect primarily the learning needs of students—classroom assessment, which in my understanding incorporates both forms of assessment, is still not the same as formative assessment. For assessment is formative when it is ‘used to identify what pupils have learned, what they have not learned and where they are having difficulty. In this way, it supports the teaching–learning process’ (Gipps & Murphy, 1994, p. 260). Formative assessment is, moreover, a low-stakes assessment situation that carries few, if any, long-term consequences for the teacher and students. It follows that whereas formative assessment refers to the function of assessment (i.e. the use of feedback into the teaching–learning process), teacher assessment and student assessment refer to the person who makes the assessment (i.e. the teacher and students respectively) (Gipps, 1994). This means that the classroom assessment process is formative only as long as teachers and students use it formatively.

It has to be said, however, that with teachers the decision to use classroom assessment formatively or otherwise is often not just a matter of personal choice. For although teachers are in a position continually to gather unique information about students that offers great promise for formative assessment (Nuttall, 1993; Calfee & Masuda, 1997), they are often required periodically to collate and report assessment results to third parties. Even when their assessments are not used for external purposes, the school itself is likely to want them to generate assessment information for internal purposes.

The possible subsequent use of teacher-generated information for both managerial (i.e. to select and certify students) and accountability purposes (i.e. to evaluate teachers, schools or age groups at national level) (Gipps & Murphy, 1994) gives classroom assessment a summative dimension that, as Harlen et al. (1992) point out, is primarily concerned with summarizing information about student achievements at particular times. It is thus mostly because assessments made by teachers are not synonymous with formative assessment (see Torrance & Pryor, 1998) that classroom
assessment can also be a high-stakes assessment situation that carries long-term consequences for students, and possibly the teacher and school.

The purposes of classroom assessment

The literature frequently refers to the purposes and uses of assessment (e.g. Rowntree, 1987; Murphy & Torrance, 1988, 1990; Eisner, 1993; Gipps & Stobart, 1993; Black, 1998; Airasian, 2000). This is usually in a form of a list with an accompanying warning that the list in question is by no means exhaustive. Gipps and Stobart (1993) provide one of the more comprehensive lists. They identify six uses of assessment:

1. **Screening**: this refers to the process of testing groups of students, normally at primary level, to identify individuals who are in need of special help.
2. **Diagnosis**: this involves the use of tests to identify children’s strengths and (more usually) weaknesses.
3. **Record-keeping**: test scores and teacher assessments are put into student records to then help in the transfer process from one school level to the next.
4. **Feedback**: results provide feedback about the progress of individual students and the teacher’s success. On the other hand, results of classes can provide information to the school administration about the progress and success across the school, and school results can be used by outsiders to ‘evaluate’ schools and teachers.
5. **Certification**: a student is provided with a qualification that signifies that he or she has reached a certain level of competence or knowledge.
6. **Selection**: students are selected into different institutions for further and higher education. They can also be allocated to different streams or sets within institutions.

I consider the above list to be fairly representative of the literature, except for its lack of reference to the two closely interrelated functions of ‘motivation’ and ‘control’. With motivation the idea is to use assessment in order to encourage the student to learn. Thus, for instance, teachers and parents can and do appeal to the value of examinations in the job market as an incentive for students to behave well and work hard at school (Gipps & Stobart, 1993). But in a system that is based on rewards and punishment, motivating students to work can actually result in controlling them, in getting students to do something they might not otherwise be inclined to do. As the line between coercion and encouragement is indeed fine, I would argue, like Rowntree (1987), that much therefore depends on the intentions and perceptions of the teacher and student, and the relationship between them.

Although Gipps’s and Stobart’s list and my two additions, as is usually the case with other lists, refer to assessment in general, I find that all the included functions apply equally well to classroom assessment as defined in this article. This is particularly so in view of the close coexistence in classroom assessment between the
formative and summative dimensions—a coexistence that is respected and reflected in the above enlisted functions. Gipps and Stobart (1993) themselves acknowledge that their six uses can be classified under ‘professional’ or ‘managerial’ nomenclatures according to whether the assessment helps respectively to enhance the educational process (read ‘formative’) or to manage and monitor the education system (read ‘summative’). Their emphasis on the ‘aftermath of assessment’ is, moreover, in line with my understanding that, rather than the actual assessments, it is the interpretation of the resulting data that may distinguish formative from summative practice (Wiliam & Black, 1996). I do not thus consider the terms ‘formative’ and ‘summative’ to typify assessments, but as descriptions of the use to which assessment information is put (Wiliam, 1998).

Given that the same assessment can be used both formatively and summatively—even if, in general, an assessment would have been designed so as to emphasize one of these two functions (Wiliam, 2001)—makes it necessary to prioritize among the various purposes of assessment according to one’s operating context. My understanding that classroom assessment should be more focused on using the inferences drawn from the emerging information to inform instruction and to monitor day-to-day progress (Joyner, 1998)—which is to say that it should primarily support the learning process—aligns my position with that of Gipps and Murphy (1994) who argue that the main purpose of assessment within the classroom context should be for professional rather than managerial or accountability purposes.

Embedded in this vision lies the notion that, contrary to what invariably happens in the traditional paradigm, assessment should be used to maximize learning services for all students rather than to legitimize minimal learning services for many of them (see LaCelle-Peterson, 2000). The recognition that formative assessment is the type of assessment that really matters inside the classroom (Nuttall, 1993) builds on the realization that certification and selection are artefacts of our social and educational system and that they, as such, are not central to the teaching and learning processes (see Gipps & Stobart, 1993).

Classroom assessment in the wider assessment context

Although my position is consequently that teachers have the duty as educators to focus on the formative function of assessment, I think it would be unyielding for teachers to drop their summative role for as long as educational systems continue to be dominated by assessment for selection and certification purposes. For the summative dimension of teacher assessment can benefit the quality of such assessments (which are likely to continue for the foreseeable future) through the inclusion of skills, competencies and knowledge that cannot be assessed by the more traditional paper-and-pencil approach (see Broadfoot, 1996; Broadfoot & Black, 2004; also National Council of Teachers of Mathematics (NCTM), 2000).

As far as the classroom reality goes, however, this improvement in assessment would have been of little concern were it not for the fact that assessment for selection and certification purposes supports instruction most strongly when classroom work
and teachers’ judgements are both valued and included in it (see NCTM, 1995). I would consequently argue that in so far as it is best that teachers are involved in both formative and summative assessment, the problem remains that of determining how teachers should go about satisfying the different requirements of these two roles.

The literature offers two main views as to how this tension may be eased, if not resolved. Some commentators (e.g. Harlen et al., 1992) argue that formative and summative procedures and data must be kept separate, as the use for summative purposes of information collected with the intention to support learning would severely impair its formative role. My understanding that it is the use to which assessments are put that determines whether the procedure is formative or summative (to which I referred above), however, makes me refute this position in favour of those (e.g. Wiliam & Black, 1996) who argue that to have two separate sets of procedures would produce an intolerable burden on teachers and that, in any case, in such circumstances the summative would always overshadow the formative because of accountability demands.

The awareness that ‘If formative assessment is to prosper, initiatives aimed at supporting a positive link between formative and summative work are sorely needed’ (Broadfoot & Black, 2004, p. 17) leads me to argue that teachers need to establish what Black (1998) calls a ‘workable relationship’ between their two distinct roles. This approach, which surely calls for goodwill, skill and discernment on their part, can help to avert the strain in the relationship between the teacher and students that, as Gipps (1994) warns, can result when teacher assessment is used for summative purposes, as the teacher may then be seen by students as judge rather than facilitator.

Classroom assessment in the alternative assessment paradigm

The current paradigmatic shift away from the traditional assessment paradigm that privileges assessment for non-professional (or non-learning) purposes inherently refutes the continuation of the system in which assessment is seen purely in terms of its product, its results, and the use of which results may be put in managing or even driving school systems (see Torrance, 1995). Assessment in the alternative paradigm is seen, in fact, as a process almost wholly integrated with teaching and learning (Torrance, 1995). This repositioning—which is in line with the deep transformation in our conceptions of learning, of assessment and of what counts as achievement—requires teachers and students alike to distance themselves from what Ellis (2001) calls an ‘objectified sense of assessment’, that is, the traditional view of assessments as events or objects that stand apart from teaching and learning.

I therefore find myself agreeing with Gipps (1994) when she argues that, according to our current understanding, the appropriate assessment model inside the classroom is one that is designed to support the teaching and learning of important skills and concepts at both basic and higher levels. This newly emerging culture—which calls for an alternative way of comprehending classroom assessment away from the traditional psychometric model—is now generally known, at least in the UK, as ‘assessment for learning’. Contrary to the traditional ‘assessment of learning’ for the
purposes of grading and reporting that has its own well-established procedures (Assessment Reform Group (ARG), 1999), ‘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’ (ARG, 2002).

The fact that the phrase ‘assessment for learning’ has come to refer to ‘any assessment for which the first priority is to serve the purpose of promoting students’ learning’ (Black et al., 2003, p. 2) may suggest, wrongly in my view, that ‘assessment for learning’ and formative assessment are necessarily one and the same thing. I agree with Black et al. (2003) that it is best to distinguish, even if subtly, between these two terms. While it is true that most definitions of formative assessment (e.g. Gipps & Murphy, 1994) essentially echo and respect the spirit of assessment for learning, some others do not. This is particularly so with definitions (e.g. Nitko, 2001) which move beyond the immediate classroom environment to include scenarios such as when the teacher revises lessons or learning materials by using information obtained from their previous use. The point is that if assessment information about learning is used, say, for recording purposes or for long-term curriculum improvement without helping the learning of the students currently involved, it might be formative for the teacher, but not for the students (Black et al., 2003). Assessment for learning, on the contrary, is primarily about addressing the needs of the present students.

The classroom assessment reform vision upheld by the current drive towards assessment for learning, which Gipps captured so well in her book Beyond testing in 1994, inevitably calls for changes in the traditional assessment roles of teachers and students. Just as the teacher is no longer considered to be a transmitter of knowledge but a facilitator of student learning and the student is no longer considered to be a receiver but a constructor of knowledge, so too are they no longer expected to behave respectively in classroom assessment solely as the ‘one who checks’ and the ‘one being checked’.

It is thus no longer acceptable to see classroom assessment, as has traditionally been the case, simply as the gathering of information about students by the teacher in order to aid his or her decision-making process. The new paradigm calls instead for classroom assessment to be seen as the gathering of information by both the teacher and students about their teaching–learning situation in order to help them in their decisions. This emerges clearly from Cross’s attempt to provide an indication of their new roles in the reconceptualized classroom assessment process:

```
Classroom assessment informs teachers how effectively they are teaching and students how effectively they are learning. Through classroom assessment, teachers get continual feedback on whether and how well students are learning what teachers hope they are teaching. And students are required, through a variety of classroom assessment exercises, to monitor their learning, to reflect on it, and to take corrective action while there is still time left in the semester. (Cross, 1998, p. 6)
```

Apart from the emphasis on the link between assessment and learning, the other key element of the new classroom assessment scenario highlighted by Cross (1998) is
the realization that students cannot rely exclusively on assessments made by teachers. Even though these assessments may provide them with good quality formative feedback, students still need to become self-monitoring learners if their learning is to improve (Sadler, 1989).

The active involvement of students in their assessment process is linked to the constructivist view that ‘it is essential to grasp the goals of one’s work and compare them with one’s present understanding if learning is to be meaningful and permanent’ (Black, 1999, p. 126). The understanding is that if students are to become effective learners, they need to progress in their knowledge of themselves as thinkers and learners, in their understanding of particular tasks and in their strategic knowledge of how to go about the improvement of their own learning (Alexander et al., 1991). The recognition that students’ progress depends on their coming to understand their strengths and weaknesses, and how they may deal with them (Harlen & James, 1997) leads me to argue that the capacity of students to judge their own work is more than a bonus in good formative assessment. It is instead a crucial component for developing the complex understandings through reflective habits of mind (Black, 1998) that they need to possess in order to share with the teacher as far as the responsibility for the quality of learning is concerned.

However, students cannot become the hoped-for independent and autonomous learners unless they first absorb standards of quality (Sadler, 1989) and develop the capacity for self-assessment (Black, 1998; Stefani, 1998). The onus that students are in a position to assume responsibility over their own learning falls, in my view, on the teacher as part of his or her mentoring role demanded by the constructivist theories of learning. Instead of remaining a provider of assessment information per se, the teacher should therefore also guide and counsel students about the quality standards expected from them and how to reach these standards. This would require the teacher ‘to share expertise, work in collaboration with students over setting assessment criteria, and provide exemplars and models so that students develop more of an understanding of the learning goals for different levels of attainment’ (Stefani, 1998, p. 348).

It is, however, important that the assessment criteria being used are either criterion- or self-referenced. This would allow each student to learn respectively if he or she can do a specific task or range of tasks (Gipps & Stobart, 1993) and about his or her progress over time, building in the process a personal profile (Black, 1998). By doing away, inside the classroom, with the traditional norm-referenced criteria that supposedly measure how much better or worse a student’s performance is in relation to that of other students, we would be contributing towards the creation of a classroom assessment environment that fosters self-esteem and motivation in students, as well as encourages cooperation among them—what I see as essential ‘for learning’ ingredients that have been much neglected under the traditional paradigm.

It needs to be said, however, that the new vision presented here for classroom assessment cannot happen unless the ‘old’ undemocratic relation inside the classroom between the powerful (i.e. the teacher) and the powerless (i.e. the students) is
replaced by dialogue in an environment in which the teacher and students alike feel comfortable to reveal their thinking (including the limitations of what they know). It follows that teachers have to let go their traditional inalienable right to rule their classroom as they move towards classroom dynamics in which teachers and students form a more collaborative community of learners (Peterson & Stack, 1998)—an important characteristic of the new paradigm that Cross (1998) seems to have missed (see above). Only then can students become insiders rather than consumers of classroom assessment (see Sadler, 1989).

But the ensuing assessment partnerships between teachers and students cannot be effective without also developing partnerships in the teaching and learning process (Stefani, 1998). This explains why I am particularly in favour of creating the so-called ‘communities of shared practice’ in which the students become ‘participants in a learning curriculum where understandings concerning what they are doing, or, in relation to assessment, what is being done to them are shared’ (Elwood & Klenowski, 2002, p. 246; see also Wiliam, 1998). In such classrooms, the students not only learn to reflect on their work and their learning, and make critical self-judgements, but they also critique the work of their peers and use productively the critique of others (NCTM, 1995). This helps them to become self-directed lifelong learners who can carry on when they are out of formal education and no longer have people to direct them where, when or what to learn (Cross, 1998).

**Assessment for learning inside the classroom: where do we stand?**

I argued in the previous section that students’ learning is best served when classroom assessment is guided by the principles of assessment for learning. This implies that the quality of classroom assessment, at least as understood within the new paradigm, depends on the extent to which there is a strong commitment towards the formative use of assessment practices in favour of the students currently in class.

A look at the literature reveals that, although assessment reform has now been a major educational issue for almost two decades in many countries, assessment practices inside these countries’ classrooms have not changed that much. The overall picture—which attests to how hard it is to introduce effective formative assessment into classroom practice (see Wiliam & Black, 1996)—depicts the formative assessment scenario in many countries as generally one of weak practice (Black, 1998). The main weaknesses are:

- Classroom evaluation practices generally encourage superficial and rote learning, concentrating on recall of isolated details, usually items of knowledge which pupils soon forget.
- Teachers do not generally review the assessment questions that they use and do not discuss them critically with peers, so there is little reflection on what is being assessed.
- The grading function is over-emphasized and the learning function under-emphasized.
- There is a tendency to use a normative rather than a criterion approach, which emphasizes competition between pupils rather than personal improvement of each.
The evidence is that with such practices the effect of feedback is to teach the weaker pupils that they lack ability, so that they are de-motivated and lose confidence in their own capacity to learn. (Black, 1998, p. 111)

Although such weaknesses sadly persist in, educationally speaking, satellite countries such as Malta (see Grima & Chetcuti, 2003; Buhagiar, 2005), I find it particularly striking that the assessment situation in countries such as the US and UK, which have been and still are at the forefront of the assessment reform efforts, is none the less gloomy. This is, for instance, how Delandshere describes the present US assessment reality:

> educational assessment practices still reflect for the most part the legacy of the past: the purposes are narrow and the methods used generate limited data. The assumptions and theories of learning are implicit; examinees submit to the process without active and equal participation (e.g. critique, reflection, self-reflection), and secrecy, reward, and punishment remain key concepts. (Delandshere, 2001, p. 130)

Delandshere reports, in fact, that ‘alternative’ assessment practices in the US are more the exception than the rule, and that they are not particularly encouraged or supported in most schools.

On similar lines, in the UK, the broader conception of educational assessment that was stimulated by the curriculum reform movement of the 1960s did not take hold and continues not to in view of the public’s need to get down to brass tacks, to go back to the basics, to measure, to monitor and to mandate (Eisner, 1993). It seems that these residual values—which clearly mirror what Broadfoot (1995) calls the ‘Zeitgeist of a previous age’, that is, the set of unquestioned assumptions concerning the merits of particular techniques in the traditional paradigm—continue to this day to act as a major stumbling block to assessment change.

The implication is that classroom assessment practices cannot change in line with the demands of the new assessment paradigm unless we first undergo a major change in the way we think about assessment in particular and education in general. This emerges clearly, in my view, from our manifest inability to move from the generally agreed-upon policies to related practices—a reality that seems to justify Eisner’s (1993) assertion that the problem is not one of correct policy formation but of practice.

**Looking forward with hindsight**

In the last two decades or so, our thinking about educational assessment in general and classroom assessment in particular has come a long way in response to our evolving understanding of what learning entails. We now speak, in fact, about the emergence of an alternative assessment paradigm that links assessment to helping students learn as opposed to the traditional emphasis on linking assessment to classifying and grading students.

From my point of view, though, the problem with these theoretical developments is twofold. First, although they enjoy the support of policy-making bodies, they have
failed so far actually to capture the imagination of teachers, students, school administrators, parents and other interested individuals. Second, although we continue to dish out matching, even if laborious, alternative assessment practices, such as records of achievement, profiling and now portfolios, no real effort has been made to create the right environment for their successful implementation. My fear is that to continue ignoring the practitioners and to keep adding to our tried and found lacking list of innovations would risk throwing overboard the whole assessment reform with tragic consequences for learning.

To dispel this possibility from becoming an uncomfortable reality, I propose that we embark on a well-targeted and marketed educational campaign aimed at promoting the centrality of learning in education and how curricular, pedagogical and assessment practices need to be reorganized to serve this end. In this process, academics evidently need to reach out beyond colleagues, their students and policymakers; their expertise needs to become visible, yet clearly open to discussion and modification, within the school grassroots, students’ families and the wider community. This participatory agenda for change would hopefully be the true beginning of a new educational era.

References


