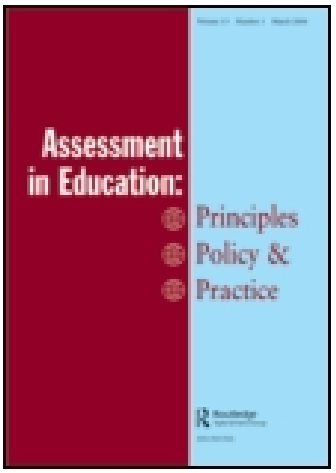


Assessment in Education: Principles, Policy & Practice
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/caie20

# Assessment and Classroom Learning: theory and practice

Carol Anne Dwyer [a]
[a] Educational Testing Service , Princeton, NJ 08541, USA
Published online: 28 Jul 2006.

To cite this article: Carol Anne Dwyer (1998) Assessment and Classroom Learning: theory and practice, Assessment in Education: Principles, Policy & Practice, 5:1, 131-137, DOI: 10.1080/0969595980050109

To link to this article: http://dx.doi.org/10.1080/0969595980050109

# Assessment and Classroom Learning: theory and practice

CAROL ANNE DWYER
*Educational Testing Service, Princeton, NJ 08541, USA*



Educational assessment has undergone a series of transformations in the 1980s and 1990s. A number of authors have ably characterized these transformations, including Tittle (1989), Mislevy (1993), Gipps (1994) and Black & Wiliam (1998). Some of these authors have characterized educational assessment as having undergone a paradigm shift. Although there remains considerable opinion to the contrary (e.g. Terwilliger, 1997), and considerable debate as to the very nature of paradigm shifts (e.g. Kuhn, 1970; Heshusius, 1989), it is clear that educational assessment has changed in fundamental ways. These changes remain, however, more in the realm of theory than of practice. One area of practice which many hope to see affected for the better by new kinds of assessment is that of assessment for classroom learning. Although Black & Wiliam (1998) have cited a number of studies linking assessment with improved learning, many barriers to realizing the potential of assessment for classroom learning remain. The principal barriers to implementation in the classroom are primarily psychological and social, however, not technical or psychometric.

## What Has Really Changed in Assessment?

It is clearly the case that assessment is currently undergoing a conceptual shift which carries with it considerable practical consequences. This shift is a complex social and technical phenomenon, however, and the movement towards new forms of assessment is not completely linear, nor is it even consistent. For example, in the United States today expanded use of performance assessments in state and local educational assessments is taking place concurrently with an expanded use of standardized, external, multiple choice testing. Major commercial test publishers report that 1997 was one of the strongest years in terms of volume of sales of traditional test batteries. At the same time that a new set of voluntary national school achievement tests are being proposed, critics of testing and many school reform advocates are arguing for an end to standardized testing. Many of those interested in school reform support testing conducted in classrooms, or they may support performance assessment whether it is conducted by classroom teachers or by external agencies.

Past barriers to improvements in assessments have been both conceptual and technical. These include the enormous complexity of the task of designing and implementing assessments; the discrepancies between psychometricians' and educator's views of knowledge and learning; and the widespread hegemony of control

0969-594X/98/010131-07 

objectives in a variety of settings that include the business world and the sciences, as well as education.

That changes in the prevailing form of testing are taking place now should not be surprising. A number of forces combine to support this move, and significant opportunities have been afforded by certain advances in the understanding of measurement. Messick's (1989, 1994) explications of assessment validity in terms of constructs, interrelated sources of evidence, and consequences have attained widespread acceptance. This view of validity greatly expands the notion of assessment quality to include study of school contexts and student characteristics. New models of learning processes are evolving from advances in cognitive psychology and related fields, and are beginning to be applied to assessment (e g. Frederiksen *et al.*, 1990; Mislevy, 1993; Snow & Lohman, 1993). Such scientific and technical developments have been very rapid and significant, and have buttressed a view of assessment that focuses on cognitive processes as well as subject-matter content. This view is compatible with changing views of assessment validity.

Rapidly expanding and affordable computing power has also opened a number of possibilities for using technology to overcome certain barriers to creating more sophisticated forms of assessment. These possibilities include computational and methodological advances such as the use of item response theory in test construction and scoring and the ability to deliver assessment tasks and accept student responses in a vastly wider variety of formats than paper-and-pencil assessment allows.

Much of the work in new models of assessment is involved with a search for a model of the entire assessment system, not just tests or the actual act of assessment. In traditional measurement terms, this is therefore not strictly a psychometric view (in the sense of a set of statistical principles), but a view that also encompasses psychological and educational considerations. That is, models are sought not just of the patterns of responses to questions, but also of the student's learning process and the social context in which both learning and assessment take place. Much of this very interesting work remains at the theoretical level, however. There is still considerable belief among test makers and educators that matters such as the nature of the learning process are beyond rational analysis, either as a matter of principle, or as a matter of practicality.

The notion of assessment design has become increasingly prominent (Embretson, 1985; Mislevy, 1993, 1994). Such an orientation typically involves attention to multiple facets of the assessment process, and to assessment as a search for evidence about specific phenomena such as student learning in a particular area. It also strongly implies attention to cognitive processes and models of student learning, rather than being limited to models of a discipline or curriculum area.

## Social and Psychological Forces in Assessment and Instruction

Despite the widespread attention it attracts, assessment itself is seldom the real issue in educational debates. Instead, the underlying questions are often found to be primarily of social psychological values.

Neither is the more circumscribed question of the relative value of formative and

summative assessment really the central issue, despite the prominence given to this topic in Black & Wiliam's paper. Black & Wiliam have reviewed a considerable body of literature that shows gains in learning associated with formative assessment, but have not made an equally strong case that summative assessment would not result in the same increases. Another problem is that the very nature of formative and summative assessment is difficult to specify precisely (although the terms connote a distinction based on purpose or intent of assessment, they also involve distinctions based on the size of the learning 'unit'). Sustained and systematic attention to the goals of learning, coupled with assessment that is sufficiently aligned with instruction, should be expected to produce increased learning whether the assessment is formative or summative in nature, given comparable social and psychological contexts in which the assessment takes place.

Important social and psychological variables include the perceived and actual users of the assessment information, the emotional valence of the assessment's purposes (e.g. fair/unfair, helpful/harmful to students), investment in control of the classroom and students, and views of the value of the teacher's autonomy, power, and right to exercise professional judgment. Tittle (1989) has provided an analysis of the complexities of the interactions among students, teachers, and assessment. In Tittle's view, which she links directly to modern conceptions of validity and to testing standards, the factors which impact on the integration of assessment into teaching and learning are not restricted to those typically identified by measurement specialists. Rather, they include such factors as students' perceptions of how testing impacts on them, students' and teachers' confidence in the veracity of test results, and students' and teachers' differing perceptions of the goals of the testing. Tittle stresses the importance of understanding how both students and teachers use assessment information, and the differences between frameworks typically invoked by scientist/test developers and by teachers.

House (1996) offers an alternative means of considering this problem of multiple viewpoints, in a highly-structured economic framework for appraising educational change that puts the context in which change occurs in a central role. House's view of educational change encompasses a variety of 'transactions' between those who would implement change and those who are stakeholders in the current process. The use of assessment to improve classroom learning is clearly an instance of such a transaction. Using House's framework, one would expect that an attempt to improve student learning through classroom assessment would invoke a series of transactions characterized by a combination of rationality, opportunism, and protection of perceived or real, tangible and intangible, assets such as independence, efficiency, and existing materials and procedures. Failure to understand such transactions can be expected to result in failure of the proposed reforms to take hold.

## Teachers and Assessment

The resistance to external assessment of classroom teachers in many countries and at all levels from primary through post-secondary, is well documented in the research literature and in everyday observation. Some of this resistance is based on

solid information about the actual or potential negative impact of certain assessments and assessment practices on teachers and students. Other resistance seems to be less data-based and related to teachers' own level of professional development and personal experience with assessment Whatever the source of the resistance, one outcome is a widespread pattern of practices that do not foster student learning, as Black & Wiliam have pointed out.

Black & Wiliam also allude to research which indicates that teachers' criticisms of prevailing standardized tests and testing practices do not necessarily result from a clear preference for other forms of systematizing the gathering of evidence about student learning. Nor does their dislike of standardized testing seem to result in their practising different and better forms of assessment in their own classrooms. In part, this certainly reflects the difficulty of developing high-quality assessments and integrating them effectively into classroom practice. Most teachers (and their supervisors or administrators) see this simply as being beyond the reasonable scope of their responsibilities and available time.

This certainly implies, of course, that many teachers do not view teaching as intrinsically involving an analytical process of collecting evidence about student learning. Assessment is most often viewed by teachers as being imposed by external forces (and as such, the responsibility of others). Moreover, they are disinclined to label their own collecting of evidence about students' learning as 'assessment', despite their strong endorsement of the importance of such activities as monitoring students' understanding in order to adjust instruction (Tracy & Smeaton, 1993; Dwyer, 1994).

In research on the knowledge and skill demands of beginning teachers, the alignment of teaching goals, classroom activities, and assessment was one of the most difficult for beginning teachers to master (Dwyer, 1994). Many beginning teachers do not clearly conceive of an integrated cycle of establishing learning goals and verifying student progress toward these goals. Novice teachers may simply not be distinguishing among important concepts related to evaluating student learning. In contrast to more experienced teachers, for example, novice teachers frequently fail to distinguish classroom activities from learning goals. In addition, they frequently select assessments that do not inform their judgments about progress toward their own learning goals. That is, the assessments that they give to the students may be unrelated to their instructional goals and the students' learning activities to the extent that the assessment cannot provide any meaningful formative or summative evidence about what learning has taken place.

Another issue that pertains to teachers' use of assessment is the way in which teachers are encouraged to conceptualize curricula. Much of teachers' education and many schools in which they teach emphasize the desirability of 'covering' the curriculum—ensuring that students are exposed to a body of knowledge or the contents of particular texts. Although these are clearly important goals, too often successful teaching is conceived as being limited to these goals, rather than including more attention to the structure of the material to be learned. Understanding the interrelationships of key concepts and other aspects of curriculum structure is a necessary prerequisite to the effective use of assessment evidence to understand

learning progress for individuals and groups of students, but it is often not reflected in classroom assessment.

It does not seem from such evidence that the paradigm of formative (or summative) assessment in aid of learning goals is a significant part of the focus of teacher education programmes. This accords with reports that formal coursework concerned with the evaluation of student learning is rapidly disappearing from teacher education curricula. This topic is, however, targeted in many professional development programmes for in-service teachers (e.g. Danielson, 1996), and is well understood by many successful experienced teachers.

## Grades and Classroom Assessment

From the teacher's point of view, another negative aspect of formal assessment, whether formative or summative, is that it increases the difficulty of their exercising certain kinds of professional judgment about student performance. Although preferences and practices vary with the age and educational level of the students, many teachers have a strong commitment to using grades to reflect motivation, improvement, effort, and other personal factors. Many teachers value the flexibility that they have to use grades as a reward, classroom control mechanism, and motivator of their students, and therefore prefer not to have tests determine too precisely their student evaluations.

Explicit reliance on students' performance on assessments usually involves a commitment to basing judgments on student performance on what is assessed, whether the assessment is under the control of the teacher or is imposed externally. Moreover, many elements of learning that are of interest to teachers are simply not addressed in the assessments that they use.

Another complicating factor in accepting assessment as playing a prominent role in classroom life is teachers' desire to make final evaluations of student learning in the light of very local 'norms' of attainment. Such comparisons may include the performance of a group of their students, or an entire class, or a set of previous classes in their experience. Most teachers do not have practical means at their disposal to make such explicit normative judgments, and thus have another reason to prefer to avoid formal assessment.

## Prospects for the Future

Black & Wiliam raise important issues about the role of assessment in learning. New developments in measurement theory have created opportunities for better assessment practice to occur, but these opportunities have not yet been realized. Social and psychological barriers to implementing instructional reforms are more likely to remain salient than are technical barriers. These aspects of teaching practice present barriers to many aspects of teaching, including but not limited to assessment.

Certain teaching reforms are gathering momentum, however, and may help us understand how assessment in the service of classroom learning may eventually cross the great divide from theory to practice. In the United States the work of groups

such as the National Council for Teachers of Mathematics, for example, has made considerable gains in the past decade in changing teachers' understanding of the importance of students' mathematical reasoning processes. Across many subject areas, teachers are increasingly incorporating project-based work and small-group learning activities in their classrooms.

A key issue in fostering such innovations is teachers' commitment to understanding students and their cognitive processes as well as the curricula that they teach. For a variety of reasons, allegiances of secondary school teachers have historically been stronger to curricula than to the psychology of learning, so the transitions that Black & Wiliam discuss may be more difficult at the secondary than at the primary school level. Improving student learning with the help of classroom assessment is a formidable educational challenge, which will involve considerable professional development efforts on the part of teachers. This challenge is likely to be met only through the concerted efforts of teachers, their teacher education programmes, teachers' professional organizations, and assessment specialists, and only if each of these key stakeholders can maintain a comprehensive vision of the learning process.



## References

BLACK, P. & WILLIAM, D. (1998) Assessment and classroom learning, *Assessment in Education*, 5, pp. 7–74, this issue.

DANIELSON, C.F. (1996) *Enhancing Professional Practice: a framework for teaching* (Alexandria, VA, Association for Supervision and Curriculum Development).

DWYER, C.A. (1994) *Development of the Knowledge Base for the Praxis III: Classroom Performance Assessments assessment criteria* (Princeton, NJ, Educational Testing Service).

EMBRETSON, S.E. (Ed.) (1985) *Test Design: developments in psychology and psychometrics* (New York, Academic Press).

FREDERIKSEN, N., GLASER, R, LESGOLD, A. & SHAFTO, M. (1990) *Diagnostic Monitoring of Skill and Knowledge Acquisition* (Hillsdale, NJ, Lawrence Erlbaum).

GIPPS, C. (1994) *Beyond Testing: towards a theory of educational assessment* (London, Falmer).

HESHUSIUS, L. (1989) The Newtonian mechanistic paradigm, special education, and contours of alternatives: an overview, *Journal of Learning Disabilities*, 22(7), pp. 403–413.

HOUSE, E.R. (1996) A framework for appraising educational reform, *Educational Researcher*, 25(7), pp. 6–14.

KUHN, T. (1970) *The Structure of Scientific Revolutions*, 2nd edn (Chicago, IL, University of Chicago Press).

MESSICK, S. (1989) Validity, in: R.L. LINN (Ed.) *Educational Measurement*, 3rd edn, pp. 13–103 (New York, Macmillan).

MESSICK, S. (1994) The interplay of evidence and consequences in the validation of performance assessment, *Educational Researcher*, 23(2), pp. 13–23.

MISLEVY, R. (1993) Foundations of a new test theory, in: N. FREDERIKSEN, R. MISLEVY & I. BEJAR (Eds) *Test Theory for a New Generation of Tests*, pp. 19–39 (Hillsdale, NJ, Lawrence Erlbaum).

MISLEVY, R. (1994) Evidence and inference in educational assessment, *Psychometrika*, 59, pp. 439–483.

SNOW, R.E. & LOHMAN, D.F. (1993) Cognitive psychology, new test design, and new test theory: an introduction, in: N. FREDERIKSEN, R. MISLEVY & I. BEJAR (Eds) *Test Theory for a New Generation of Tests*, pp. 1–17 (Hillsdale, NJ, Lawrence Erlbaum).

TERWILLIGER, J. (1997, November) Semantics, psychometrics, and assessment reform: a close look at 'authentic' assessments, *Educational Researcher*, 26(8), pp. 24–27.

TITTLE, C.K. (1989) Toward an educational psychology of assessment for teaching and learning: theories, contexts, and validation arguments, *Educational Researcher*, 29(2), pp. 149–162.

TRACY, S.J. & SMEATON, P. (1993) State-mandated assisting and assessing teachers: levels of state control, *Journal of Personnel Evaluation in Education*, 6, pp. 219–234.