A Conceptual Evaluation of Primary Assessment Policy and the Education Policy Process in the Republic of Ireland

Kathy Hall

A Conceptual Evaluation of Primary Assessment Policy and the Education Policy Process in the Republic of Ireland

KATHY HALL, *Leeds Metropolitan University*

ABSTRACT   *This paper offers a conceptual evaluation of primary assessment policy and the policy process in the Republic of Ireland. It considers, in particular, the National Council for Curriculum and Assessment's policy document,* A Programme for Reform: curriculum and assessment policy towards the new century *and within that, the policy statement,* The Curriculum at Primary Level: programme development and pupil assessment *and it also refers to the primary teachers' union stance on assessment. It critically evaluates the Irish policy position on assessment in the light of recent literature in this field. The main argument of the paper is that the proposed policy is preoccupied with measurement issues and it does not sufficiently recognise the complexity of implementing formative assessment. The paper also seeks to explain the nature of current policy in Ireland with reference to the decision-making process and comparisons with other systems outside Ireland, particularly England/Wales, are made, where appropriate.*

Introduction and Rationale

It has been argued that little attention has been paid to the shaping of educational debate within the Irish policy-making community and that the process of policy decision-making receives inadequate attention from scholars (O'Sullivan, 1992; Walshe, 1997). The intention of this paper is to evaluate the substantive area of primary assessment policy in Ireland but, in doing so, to have regard to the themes, interpretations and authorities which form the background against which the policy itself was created.

   The first part of the paper chronicles the currently proposed assessment policy and it teases out various pressures bearing on its shape. It fully describes the proposed assessment policy. In the second and major part of the paper, the strengths and weaknesses of the proposed policy are examined in the light of contemporary learning and assessment theory. This is followed by a critical focus on the decision-making process. This part highlights the fact that while some of the influences on the assessment reforms are shared by other systems, others are particular to the national context itself. The different political agendas obtaining in Ireland and England/Wales are noteworthy, especially in view of the traditional tendency of Irish curriculum and assessment policy to follow, eventually, trends in Britain. The conclusion includes general recommendations for the kind of assessment policy that is most likely to raise education standards and it attempts to identify features of the decision-making process that may be important in mounting policies that are informed by relevant research.

Chronology of Events and an Account of Proposed Assessment Policy

In 1990 the Review Body on the Primary Curriculum (Department of Education, 1990), which was set up by the then Fine Gael Minister of Education to review the primary curriculum, identified what it saw as three *kinds of assessment*: individual pupil assessment, teacher and school evaluation, and system evaluation. It clearly states that it is 'inappropriate to rely unduly on measures of pupil attainment to gauge the individual teacher's effectiveness' (p. 80). It goes on to say that an 'extensive use of attainment tests is inappropriate … and would be especially prejudicial to the needs of disadvantaged pupils' and that teacher accountability should be determined through an evaluation of teaching (pp. 80–81). It lists several principles of and recommendations for assessment that add up to a clear endorsement of curriculum-oriented and formative assessment. It advocates the use of a combination of 'informal teacher assessment and the judicious use of standardised tests (criterion-referenced and norm-referenced)' where both types are curriculum linked. It is explicit about the purpose of *informal* assessment which it equates with teacher assessment: 'the central purpose of teachers' assessment should be for motivational, diagnostic and guidance purposes' (p. 83). It is *not* explicit, however, about the purpose of *formal* assessment which it equates with standardised testing.

It devotes considerable space to recommending procedures for the recording and reporting of assessment information to parents. It advises that assessment information destined for the post-primary school should be *summative* by which is meant the provision of a summary mark/grade/comment that would indicate an overall level of performance in the subjects on the curriculum. It advises that these summary ratings should be standardised through such procedures as 'standardised test information, group moderation, verbal description of prototypes (allocation of pupils to one of three broad bands of achievement within each subject)'. The resource implications identified to support their recommendations include teacher training in the theory and practice of assessment and the need to devise a range of standardised tests and *item banks* in various subject areas.

The Green Paper of June 1992 (Department of Education, 1992) was produced by the Fianna Fail/Labour coalition government in which the education portfolio was controlled by the more conservative, Fianna Fail element. It did not adhere to the advice of the Review Body, adopting a much narrower conception of assessment which it largely equated with testing. This position should be viewed against what government officials saw as a 'lack of adequate quantitative data on educational attainment, particularly in the early years of schooling' by which it meant the primary years (p. 4). The latter, together with its perceived lack of an adequate system of quality assurance, were listed among shortcomings in the current education system in Ireland and so a key aim of the government was stated as follows: 'to ensure great openness and accountability throughout the system, and maximise parent involvement and choice' (p. 27). The purpose of each school publishing its school plan would be to assist parents in making 'a more informed choice in the selection of schools for their children' (p. 28). Each school, it was stated, should produce a written annual report which should include the outcomes of assessment—'[i]t should contain, as an appendix, a summary analysis of relevant statistical data' (p. 114). This was interpreted to mean that each school would be expected to make assessment results available to parents.

Unlike the *Report of the Review Body*, the discourse of which was pedagogical and curricular, the discourse throughout the Green Paper, like that of reforms in England and

Wales (see Broadfoot & Gipps, 1996), is market-driven. Competition, choice, marketisation, management techniques, auditing, contractual arrangements, performance indicators and accountability are either implicit or explicit features of this discourse. This official document recasts education in a market mould and a market in education could only operate if parents (not yet cast as *consumers*) had information with which to select and make judgements about schools. A key feature of the assessment reform endorsed in the Green Paper, therefore, was system accountability, so the provision of public information on school performance becomes vital. It acknowledges the value of *testing* (by which it means standardised tests) 'as a diagnostic aid' (p. 175) but sees tests for 7 and 11 year olds as 'most appropriate for this purpose'. Moreover, it states that the test results would be made available in aggregated form to the Department of Education in order to 'provide valuable public data on which to base remedial efforts and the targeting of such efforts' (p. 176).

The stance in the Green Paper towards assessment in particular (and towards curriculum and learning, more generally) closely parallels the Conservative-led position on reform in England and Wales and undoubtedly was influenced by events here. It is also in keeping with trends in other Western and developing countries (Little & Wolf, 1996). However, the Fianna Fail–Labour government collapsed at the end of 1992 and, without an election, a new centre–left administration was installed comprising Fine Gael, Democratic Left and the Labour Party. For the first time in the state's history, the Labour Party secured the education portfolio. In line with other countries, Green Papers in Ireland are intended as discussion documents while White Papers outline the government's decisions. The new Minister for Education, Niamh Breathnach, extended the deadline for the receipt of responses to the Green Paper to April 1993.

By that deadline, over 1000 written responses had been received and several public and private seminars and conferences had been convened (see, for example, Hogan, 1995a). The education correspondent for *The Irish Times*, John Walshe (1997), described the scale of responses and the diversity of organisations from which they came as *extraordinary* and he explains this reaction in terms of the growing public interest in education as a means for personal and societal advancement. Unlike the situation in England where it was mainly teachers who objected to the marketisation of education, there was overwhelming opposition from many sectors of Irish society to what was perceived as an instrumental and technical orientation to education. For example, the Catholic Primary School Managers' Association argued that the Paper was inspired by 'economic pragmatism, acquisitive individualism and functional efficiency'; that it lacked an 'explicit philosophy'; it also argued, as did the three main teacher unions, against national testing for 7 and 11 year olds. The National Parents' Council (NPC) supported the use of standardised tests provided they formed part of an in-school assessment package (see Walshe, 1997). The scale of the public response to the Green Paper prompted the Minister to prolong the consultation process through a national convention on education, the reports from which were extensively covered in the newspapers, at conferences including, for example, the British and European Educational Research Association Conference in 1995 (Hogan, 1995b) and in one substantial volume edited by the convention's Secretary General (Coolahan, 1994). The composition of the secretariat is noteworthy—seven were academics in senior positions in higher education or research institutions in Ireland (John Coolahan, Patrick Clancy, Sheelagh Drudy, Damian Hannan, Tom Kellaghan, Seamus McGuinness and Maire Ui Mhaicin) and there were three *international members*—Malcolm Skilbeck, Peter Mortimore and Galo Ramirez.

The National Education Convention (Coolahan, 1994, p. 1) was defined as 'an unprecedented, democratic event in the history of Irish education' and brought together representatives from 42 Irish organisations. It debated key issues of educational policy in Ireland in 1993, among them, assessment. It rejected the proposal of the Green Paper that all pupils should be tested through standardised tests at ages 7 and 11 and objected, particularly, to the proposal that test results for schools should be made available to the Department of Education. The grounds for this rejection stemmed from the convention's view that such an approach would lead to 'a high stakes accountability situation which would have a narrowing effect on the primary school curriculum' (Coolahan, 1994, p. 71), that the design of tests to fulfil both formative and accountability purposes would be too difficult, and that the difficulties of controlling the testing conditions would render it unfeasible. For the purpose of system monitoring the convention advocated the non-controversial procedure of testing a representative sample of pupils from time to time.

It did endorse the need for schools to have a clear assessment, recording and reporting system and that assessment results should be considered on a whole school basis. It went on to claim that '[t]here is an obvious place for standardised tests in such a system to provide normative and diagnostic information as a basis for devising teaching programmes' (p. 72). Because the assumption was that standardised testing has an *obvious* place, the convention did not feel the need to justify its place beyond the presumed need to compare children's performance with group norms (school, national, etc.) It is difficult to see how norm-referenced testing could provide meaningful diagnostic information that would inform teaching, but I return to this point below. However, the *Report of the National Education Convention* notes that it was not successful in identifying an assessment system that would be acceptable to all parties (Coolahan, 1994, p. 72).

The National Council for Curriculum and Assessment (NCCA) was established in 1987 to advise the government of the day on educational issues. This body was explicitly and directly representative of major educational interests. In 1993 the council published *A Programme for Reform* which incorporated policy statements and recommendations on curriculum and assessment. Because a significant section of this document, entitled *The Curriculum at Primary Level, Programme Development and Pupil Assessment: a policy statement*, is the most recent and detailed official statement on assessment in the Republic and informs the stance to assessment in the White Paper (Department of Education, 1995) and in the draft curriculum documents (1997), it merits in-depth analysis.

In line with the *Report of the Review Body* it aspires to a model of assessment that gives priority to formative purposes and to *the central role* of the teacher. It advocates the use of 'various instruments and methods of assessment' and cautions against the introduction of *reductionist* procedures that would have a *backwash* effect on the curriculum (para 5.1.2). It is clear and explicit about the use of assessment for accountability purposes: 'The results of [formative and diagnostic] assessments should not be used as a mechanism for teacher or school accountability' (para 5.1.4) and '[r]esults of standardised or other forms of assessments should not be published for school, class or individual' (5.5.1) and also, assessment records 'should not be used for accountability purposes' (para 5.6.1).The policy statement endorses the use of nationally norm-referenced and criterion-referenced assessments on the grounds that they 'fulfil different yet complementary functions' (para, 5.2.3). Among the functions that *systematic* assessment are expected to serve are the evaluation of learning outcomes at individual, group, class and school level; the monitoring of teaching and planning

(*programme effectiveness*). It is stated, too, that alongside *teacher assessments,* the use of 'standardised assessment tests' will help 'provide pupils and their parents with accurate and differentiated information' (para 5.3.7). System monitoring, it is suggested, can be done through national sampling so it is not envisaged that the results of tests/assessments would be passed on to the Department of Education for accountability purposes.

The time scale for the introduction of the new (curriculum) and assessment policy is not specified but the statements that the introduction will involve 'dissemination of information and distribution and discussion of documentation', that there should be a defined time gap between introduction and implementation (para 3.2) and that 'a process of evolutionary development' (p. 7) is seen as desirable suggest that change will not be thrust upon unprepared teachers or imposed in a diktat fashion.

The White Paper (Department of Education, 1995) was published while Niamh Breathnach was still Minister for Education and it incorporated quite a detailed philosophical framework. It is clearly influenced by the NCCA's assessment advice and by some, although not all, of the arguments on assessment put forward by the national convention. Government thinking on primary assessment as evidenced in this document is very much in line with the advice of the NCCA and the market discourse of this Paper, unlike its predecessor, the Green Paper, is considerably tempered, if not entirely gone. In particular, the earlier intention to give parents access to a school's aggregated results so they can make informed choices in selecting schools is now dropped.

However, there is one crucial continuity with the Green Paper and discontinuity with both the NCCA document and the view expressed by the national convention—the passing on of assessment information to the Department of Education. It states that:

> [a]ggregated assessment outcomes for each school, in accordance with nation-
> ally agreed guidelines, will be available on a confidential basis to boards of
> management, to education boards and to the Department of Education for the
> purpose of quality assurance, the identification of special learning needs and
> the targeting of resources. These data can be of assistance in enhancing the
> quality of education regionally and nationally. (p. 30)

It is arguable, on the one hand, that the promise of resources may invite misuse of the assessments in certain circumstances. On the other hand, it is arguable that such information, once made available, could quite easily find its way into the hands of journalists and newspapers and be used for other purposes, particularly to rank schools, albeit against the intention of the Department of Education. Moreover, once available to so many agencies and particularly to the Department of Education, there is no guarantee that the assessment results would not (in time) be used to make individual teachers and schools accountable. A new government could decide to publish the results in response to pressure from parents. The content and tone of the *Reservation* of the NPC which was included in the *Report of the Review Body* and the fact that the NPC also expressed reservations about the NCCA statement, which was not included in its report (NCCA, para 5) suggest that this is far from an unlikely scenario. However, it must be noted that the White Paper, in line with the advice from both the NCCA and the national convention, advocates a system of monitoring standards, based on 'the regular assess-ment of the performance of a representative sample of schools', and schools participating in these monitory exercises would not be identified publicly (p. 30).

The government decisions outlined in the White Paper were, by and large, the culmination of a genuine consultation process. Educationists undoubtedly managed to

exert considerable influence. The nature of the proposed assessment reform (together with its pacing and introduction into schools) are in marked contrast with the situation in England and Wales (see Black (1997) for an account) where the pace of reform was both radical and rapid and where some obviously interested parties were largely omitted from the decision-making process. As Ball (1994, p. 50) eloquently expresses it—the teacher in England and Wales is 'an absent presence' in education policy, 'an object rather than a subject of discourse'. In this regard also, it is noteworthy that, unlike the case of England and Wales, there is formal and frequent recognition, especially by ministers of education, of the contribution of teachers to education in Ireland. Several ministers, for example, have expressed support for a Teaching Council (Steering Committee on the Establishment of a Teaching Council, 1998) and very frequently one finds endorsement of the high quality of the education system and of teachers in statements by government agencies and multi-national corporations. Moreover, I should point out that the impetus for assessment and the wider educational reforms in England and Wales stemmed from different political agendas—the over-riding impetus for change stemmed from the perceived need among Conservative politicians for the marketisation of the education system and a desire on their part to make the system more accountable and to make the curriculum more concentrated on the *basics* of literacy and numeracy. The scale of political interference in education was unprecedented and the government frequently rejected many of the principles on which professionals had reached a consensus (e.g. Cox, 1995). One particular quotation serves to summarise for an international audience the marginalisation of educationists in the birth of the National Curriculum in England and Wales. At the time that the legislation for the introduction of the National Curriculum was going through Parliament, the then Secretary of State, Kenneth Baker, described the bill as follows: 'I would sum up the bill's 169 pages in three words, standards, freedom and choice … We must give consumers of education a central part in decision-making. That means freeing schools and colleges to deliver the standards that parents and employers want ...' (Hansard, 1987). As noted above, the push towards a market in education was attempted in Ireland in the Green Paper but was wholly rejected by the public.

The assessment elements of the proposed Revised Curriculum (draft form) and various subject areas in Ireland, designed by the NCCA (1997), conform to the principles and procedures detailed in their earlier policy statement, although, fortunately, the draft of the Revised Curriculum[1] is more explicit about some of the procedures for formative assessment. To exemplify, the draft English document recommends that a variety of assessment tools can be used including: teacher observation, teacher-designed tasks and tests, work samples, portfolios, projects, curriculum profiles, diagnostic tests and standardised tests. Each of these is described in the document although diagnostic *testing* gets rather more space than any of the other tools.

The recommended function of standardised testing is noteworthy. It states: standardised assessment (both norm- and criterion-referenced) 'can quantify the extent to which the child is performing in relation to particular language skills and confirm less precise judgements made using the more informal assessment tools'. It continues: 'Standardised tests contribute to the accuracy of the teacher's monitoring and help to identify the needs of individual children and the appropriate learning targets they require' (p. 51). Referring to 'a balanced approach to assessment' it states that the:

> … principal function of assessment is to provide the teacher with an accurate picture of the child's language development. This will enable him/her to create

the learning contexts and design the teaching strategies most appropriate to the needs of individual pupils. A great proportion of the teacher's assessment will involve the use of less structured methods and will be an integral part of the teaching process. Assessment techniques like observation, teacher-designed tasks and tests and the use of work samples/portfolios are, by their nature, subjective. It is important, therefore, that teachers moderate their standards and criteria against a wider base of teacher experience. Staff discussion and school-based in-career development can help to provide the teacher with a wider perspective and more objective standards of reference for those forms of assessment. (p. 51)

(I challenge the assumptions about relative accuracy and precision of standardised assessments and the subjectivity of *less structured* procedures later in the paper). There is plenty of guidance in the document on the reporting of assessment information about individual children to their parents and this is uncontentious, but there are no references to what, if any, information should be forwarded to the Department of Education. At the time of writing I am unaware of any decision having been made regarding the exact role of pupil assessments in system accountability.

Recognition of Formative Assessment: the major strength

The first point to make here is that assessment in primary schools in Ireland had not been fully explored and exploited for the benefit of pupils, their parents and, of course, their teachers so attempts to address this are recognised. More specifically, aspirations to accord primary status to formative assessment, to the teacher as the central player in the assessment enterprise, the emphasis on the need to use a variety of methods of assessment, and the aspiration to avoid reductionist methods that have a negative and narrowing effect on what is taught and how it is taught are all very much in line with contemporary research on assessment from the USA and Britain (Nuttall, 1987; Crooks, 1988; Stiggins, 1992; Harlen *et al.*, 1992; Gifford & O'Connor, 1992; Graue, 1993; Gipps, 1994; Harlen, 1994; Torrance, 1995; Murphy, 1996; Black & Wiliam, 1998). The recommended use of pupil profiles, although their form and function only became clear later, in an Irish National Teachers' Organisation (INTO)-commissioned chapter (Shiel & Murphy, 1997), is also a strength as this approach aims to set clear standards against which pupils' performance can be assessed; these standards can be shared with pupils, particularly as they progress through their school careers, and, with the mediation of the teacher, pupils themselves can be encouraged to monitor their own performance.

In addition, the decision to implement the assessment (and curriculum) changes on a very gradual basis fits well with the literature on the management of educational change (Fullan with Stiegelbauer, 1991). It is noteworthy in this regard, for example, that work on the development of curriculum profiles for English is ongoing since 1995, that teachers have been directly involved in the piloting of these procedures in their classrooms, and that their views about their usefulness have been sought (Shiel & Murphy, 1997), thus maximising the likely match between policy and practice in due course.

While it is important to acknowledge the key strengths of the proposed assessment system, it is equally important to recognise the weaknesses, particularly since those weaknesses are likely to detract from the potential impact of the system's strengths. It is to these critical issues I now turn.

Conceptual Clarity and Epistemological Weaknesses

The NCCA statement advises the use of different forms of assessment and asserts that there is a 'continuum of assessment, ranging from classroom observation … to standardised tests' (5.2.1). It claims that no single form of assessment is superior to another, that each form has advantages in certain situations. However, the nature of the continuum—formality, validity, reliability—is not explained although I believe 'soft' to 'hard' and 'inaccurate–accuracy' is implicit in view of the emphasis on and the nature of the statements made on standardised assessment in what immediately follows in the text. And it does not clarify the circumstances in which the various forms should be used.

Despite the assertion that all forms of assessment should have parity of esteem, the document devotes more space and comment to standardised tests and their functions than to teacher-designed assessments. It is explicit about the functions of standardised tests, it is not explicit about the role of teacher assessments. This means that inadequate and unbalanced advice is offered in relation to the forms that ought to be used in different situations. This is most unfortunate as it communicates certain messages about the status of the different forms of assessment and the faith one can have in their results, despite the rhetoric that they are equal in status. It should not be entirely surprising, therefore, to observe that the White Paper seeks aggregated test outcomes from schools on which to base decisions about the system and particularly about the allocation of resources.

The policy statement claims that nationally norm-referenced and criterion-referenced assessments 'should be used, as they fulfil different yet complementary functions'. What these different and complementary functions are is not sufficiently elaborated or explained. Lack of clarity is a feature of the policy statement as a whole. Despite the space given to standardised assessment, it is not at all clear what the status of this type of assessment should be (relative to other types), or what the nature of it should be (whether norm-referenced or criterion-referenced) and there are confusions and contradictory statements about how often pupils should experience it (5.7.4 and 5.7.5). The net effect of this confusion and lack of coherence in meaning together with the unbalanced attention to standardised testing legitimated the White Paper's [and indeed the INTO's (1997)] subsequent endorsement and unjustified faith in standardised tests. In my view the government did not get a sufficiently clear steer on the appropriate role of standardised assessment. However, I also acknowledge that policy documents are texts which are capable of being interpreted in different ways depending on the context in which they are read (Codd, 1988) and the new government, like its predecessor, saw this form of assessment as providing it with greater control over the education system.

Along with what I see as a lack of balance between standardised assessment and teacher assessment, the policy statement tends to emphasise recording and reporting (to parents and other agencies) at the expense of the act of assessing and giving feedback to the learner. While not wanting to take from the importance of those elements of the assessment process, these could be considered more procedural and routine than principles of how to assess well, how to interpret results, how to feedback to the learner, etc. It is important, for example, that recording and record keeping do not become ends in themselves.

The Forward to the NCCA statement claims that all pupils are 'entitled to a broad and balanced programme of education' and that '[i]t is a dangerous folly to set up one aspect of the curriculum against another' (p. 3). Despite this, the curriculum areas of English, Irish and mathematics are singled out for assessment through 'a range of standardised assessment tests'. The suggestion that 'the application of standardised tests to other areas

of the curriculum should be explored' (5.7.2) is sufficiently tentative to be conveniently ignored—a good thing in my view—but what I am highlighting here is the contradiction. On the assumption that what you test is what you get, this reinforces the opposite stance to the one endorsed in the Forward.

The proposed priority attached to standardised testing in Ireland as reflected in the policy documents and more recently in the publication on assessment of the primary teacher's union (INTO, 1997) reflects an outmoded model of assessment, learning and knowledge, and directs resources and energy away from more complex and demanding aspects of the assessment process, thus running the risk of undermining the potential that now exists to institute the kind of assessment system that would promote high-quality learning. It is not intended to rehearse here the extensive literature that demonstrates the inadequacies of testing (see Gipps (1994) for a review). In summary, this literature identifies *a paradigm shift* from a testing culture with its associated emphasis on psychometrics and measurement to an assessment culture with an emphasis on the assessment of learning and a realisation that assessment is an inexact science that depends crucially on judgement and interpretation.

A critical review of no less than 578 studies demonstrates conclusively that the kind of assessment that promotes learning is formative assessment—the kind of assessment that is integrated into teaching and learning and is characterised by attending to the learner's understandings, learning strategies and dispositions to learn, by attending to the learners' responses in relation to her/his expectations and assumptions about the classroom process and learning context, by attending to how the learner interprets the assessment tasks set and the criteria for success (Black & Wiliam, 1998). Their conclusion is unequivocal: '[t]he research reported here shows conclusively that formative assessment does improve learning. The gains in achievement appear to be quite considerable, and …amongst the largest ever reported for educational interventions' (p. 61). However, among the key findings are that this kind of assessment is not well understood by teachers and is weak in practice and that its implementation calls for 'deep changes both in teachers' perceptions of their own role in relation to their pupils and their classroom practice' (p. 20). This and many other recent studies on the practice of assessment clearly demonstrate that there can be no effective change at the level of the classroom without schools and teachers being provided with the necessary training and resources (See *Assessment in Education*, 4(3), 1997).

It follows, therefore, that a priority for the Irish policy-makers must be the development of teachers' professional judgement and interpretive skills in their pedagogical and assessment interactions with their pupils. The NCCA policy statement is weak on this kind of teacher development, adopting a more instrumentalist perspective in emphasising the training of teachers in test procedures (para 5.8) while INTO (1997) seems to subscribe to a psychometric model of assessment and is equally technicist in its model of teacher development (see Chapters 1 and 2 and pp. 74–75).

Of particular concern in the case of the INTO perspective is the assumption that teachers should 'buy into' and, apparently, defer to the discourse of other professionals who speak the language of measurement:

> Increasingly, teachers must liaise with a range of other professionals such as Department inspectors, psychologists, speech and language therapists, doctors, public health nurses, social workers and solicitors. This frequently necessitates the exchange of information and professional judgements on the intellectual, academic and social development of pupils. These judgements must be capable

of being substantiated with evidence, evidence which is often unlocked by an assessment instrument. It is increasingly necessary for teachers to be conversant and comfortable with the assessment vocabulary and techniques of these related professionals. While this dialogue assumes test literacy as a minimum, the professionalism of the teacher is enhanced where s/he can draw upon the valuable information which can be gleaned from appropriate assessment instruments. (pp. 12–13)

This stance clearly does not empower teachers or see them as equals, much less, leaders in the enterprise of assessing learning. It demonstrates unjustifiable confidence in the potential of testing instruments to inform learning and ignores the potential evidence inherent in informal, classroom assessment tasks, observation, profiles, portfolios, and the potential of moderated teacher assessment in yielding evidence about pupils' achievements. These procedures have greater potential to motivate pupils because the tasks involved are more meaningful; they also have greater capacity to target higher-order thinking and problem-solving skills (e.g. Herman *et al.*, 1997).

What the *paradigm shift* in assessment implies is that policy and practice need to move away from a notion of testing as a thermometer to a notion of assessment as a feedback mechanism for learning. All the Irish assessment statements place great emphasis on the *identification* of learning difficulties (e.g. NCCA, para 5.1.3) and assume that the best way of doing this is through testing. The *identification* of those pupils with learning difficulties or the identification of what these children do not know or cannot do, do not, arguably, pose such a problem for teachers. They rarely need a test to tell them which children in the class are not making adequate progress. It is what to do to alleviate those difficulties, how to intervene effectively—these are the areas that they need help with, and here, tests offer little, if any, guidance.

INTO makes the point that the 'undoubted rise' in the use of standardised tests in Irish schools is the result of 'voluntary decisions by teachers' (p. 4). While this is the case, the reason for the increased incidence is because teachers were not encouraged to consider alternatives, specifically, tools that have the capacity to yield more authentic assessments. Another reason is that funding tends to be directed towards the production of norm-referenced and criterion-referenced *tests* rather than towards the promotion and moderation of teacher assessments and the production of standard assessment *tasks*. The main point here is that teachers need more than assessment instruments, they need support in interpreting and responding to their assessment results in a formative way (Black & Wiliam, 1998). This is not to deny that they also need to be supported in acquiring a critical awareness of psychometric models of assessment.

I have already noted that one of the strengths of the NCCA statement is the emphasis on using a range of assessment methods. However, on the basis of research conducted at the Assessment of Performance Unit (APU) in England, the policy should be reviewed to take account of the need to adopt a range of assessment modes—presentation, operation, and response (see DES, 1988; and see Gipps & Murphy, 1994, Chapter 5). It needs to be much more explicit, too, about how to assess the range of achievements that it considers important.

The assumption made in the NCCA policy statement and repeated in the INTO publication that standardised tests yield *hard* and *accurate* data while teacher assessments offer *soft* and less *accurate* (if more differentiated) information (NCCA, para 5.3.7 and 5.2.2; INTO, p. 11; and Draft English, p. 51, quoted above) is unhelpful for it ignores the most important consideration of all assessment procedures—validity—and

more specifically, the rich and more unified view of this concept that has been made available over the past decade (Messick, 1989; Wiliam, 1996; Crooks *et al.*, 1996). Messick's view of validity greatly expands the notion of assessment quality to include a study of school contexts and pupil characteristics.

The Irish model assumes that the results of standardised tests are not subjective. A study conducted over 20 years ago is instructive in this regard. Leiter's study (detailed in Stierer (1990)) notes that one of the reasons advanced for the use of standardised tests is the elimination of teachers' subjective knowledge when assessing their pupils. However, he found that, in order to make the test scores meaningful, teachers located them in just the kind of background knowledge the scores were meant to replace. This knowledge pertains to home backgrounds, pupils' classroom behaviour, views about curriculum and learning, and teachers' experience and understanding of the test itself. In addition, different kinds of knowledge were used in the case of different pupils and the same kinds of knowledge were sometimes used to create different meanings from different pupils' scores. Leiter, however, does not criticise teachers for this or urge them to be more objective, rather he suggests that the use of subjective knowledge is the only way to make sense of *objective* test scores. As Stierer rightly observes, the promotion of numbers/test scores (standardised or otherwise) to indicate competence 'invites and authorises subjectivity in a way which less quantifiable assessment measures may go some towards preventing' (1990, pp. 154–155). Teachers *give* meaning to their pupils' scores just as they give meaning to their actions and responses.

As is clear from the research literature referred to above, if learning is to be improved, assessment has to be integrated into teaching/learning. This presupposes that how children acquire knowledge and how that knowledge can be accessed should be taken into account (Murphy & Moon, 1994). Contemporary learning and epistemological theory, and specifically, constructivism, hold that knowledge is not passively transmitted to learners but that it is the result of an active construction by them in relation to their prior knowledge and experiences and in relation to the context of their learning. If we accept that this is the case, then there can be no independent, objective reality, since no two individuals can have exactly the same set of prior experiences (von Glasersfeld, 1994). It follows from this that knowledge is not separate from the knower or from the context of its occurrence. The implications of this for assessment are that attention has to be paid to the way learners interpret their assessment context. Murphy puts this well:

> …if assessment practice is to be fair to pupils attention needs to be paid to the context and activities used to develop the achievements being assessed as these will define in part the achievements of the pupils. Such a view of knowledge as situated also raises a question mark about the validity of generalizing about achievement, or indeed lack of achievement, from a small number of assessment instances. Good assessment practice has to recognise the tentative nature of judgements made about pupils' achievements. (1996, p. 179)

The problem with the NCCA's emphasis on testing, especially norm-referenced testing, is that, in representing the most rudimentary versions of behaviourism, it is a poor match for constructivist models of learning and teaching.

Explaining the Shape of Assessment Policy and the Real Decision-makers

The shape of Irish assessment policy is explained with reference to a number of factors—the composition of the decision-making committees, the influence which some

of the interested groups appeared to be able to exert on those committees, and the inadequate level of expertise of those who made the decisions and who advised the government.

Reference has already been made to the democratic nature of the NCCA committees. These were constituted of representatives from the teacher unions, the Department of Education, management groups, industry, and the NPC. In addition, the assessment committee consulted national agencies and other interest groups about the issue of assessment. It is noteworthy that a survey of NCCA participants in general showed that respondents overwhelmingly identified the greatest strength of the NCCA committees as being their representative nature (Granville, 1994). This helps explain the consistency across the advice offered to the government and the latter's adoption (if not total) of that advice in the White Paper (Department of Education, 1995).

Policy research shows that *social corporatism*, whereby councils composed of professionals and interest groups determine the direction of educational reform, enables the country to act in a consensual manner (Kvavik, 1974; Rust, 1990) and this, in turn, is likely to support the continuity between policy and practice. Current (and past) educational policy-making in Norway, for example, represents a good example of this approach (Hall *et al*., 1998). The committee charged with advising on assessment decision-making in England/Wales, in contrast, was composed mainly of people with strong curriculum/assessment expertise, people who had a very strong academic profile, and, incidentally, that committee had no teacher union representative (Black, 1997). (In any event its advice was not taken by government.) Recent research in England demonstrates the folly of introducing a system of assessment without the support of the teaching profession (Brown *et al*., 1997). The case of England and Wales is also a classic example of how initiatives which are not in harmony with the aims and understanding of the government of the time do not survive (Broadfoot & Gipps, 1996).

It is worth speculating on why, bearing in mind that the overall purpose of assessment is clearly stated as improving and informing learning, there is a relatively greater emphasis on standardised assessment over other types of assessment in the NCCA document. The effect of this is that resources, instruments and techniques are (unwittingly) prioritised over the mediating role of the teacher. I speculate that there are at least three inter-related factors relevant here and discuss each one briefly.

On the basis of this textual analysis, the NCCA (and previously, the Review Body and the national convention) assumed that since teachers are already engaged in teacher assessment for formative purposes through such informal means as observations, teacher-set tasks and tests, and classroom interaction, they are also doing it well and, therefore, further guidance is unnecessary. [This stance is not unique—there is no doubt that official reports in Britain (see Hall, 1997) assume that there is an automatic and simple link between diagnostic assessment and the capacity to promote learning.] Standardised testing, on the other hand, is perceived as being unfamiliar to teachers or, at least, as a more recent aspect of their practice. Studies of practice in England challenge the assumption that informal assessment by teachers for formative purposes is an area of strength (Black, 1993, 1998; McCallum *et al*., 1993; Desforges *et al*., 1994; James, 1994; Radnor, 1994; Macrae *et al*., 1994; Torrance, 1995). And a most comprehensive, recent review of the international research on the subject concludes that the state of formative assessment is far from healthy (Black & Wiliam, 1998).

On the question of teachers' familiarity with standardised assessments (letting aside the merits of this form of assessment) there is some evidence that teachers are already well practised in the application of standardised tests. A study in 1994 showed that 90%

of fifth and sixth class teachers in 54 schools in Cork city and county had a policy of regular use of standardised testing in English and mathematics (Curran, 1994). Furthermore, it is likely that any training (as opposed to critical understanding) that might be needed in relation to the use of this form of assessment is more technical than conceptual (again letting aside its merits).

I speculate that a second explanation for the relatively high status on testing stems from previous reports, particularly the *Report of the Review Body* (Department of Education, 1990) which had the strong backing of both the Department of Education and the primary teachers' union. In my view the NCCA statement is unduly influenced by the 1990 report. But this begs the question of why standardised testing enjoys such status in those earlier reports. To this, one can point to the stranglehold of tradition and to what Patricia Broadfoot calls 'the myth of measurement' (1995)—*myth* because it has the task of depoliticising language; myth makes things innocent and gives a *natural* justification. (The word *obviously* in relation to standardised tests in the *Report of the National Convention* is telling, see above.) Broadfoot exposes the myth in arguing that educational assessment can never be scientific but that the widespread acceptability of educational measurement as manifested in so-called psychometric tests lies not in its scientific worth, nor in the capacity of techniques to measure accurately what they claim to measure, but simply because we believe in them.

The third explanation that I advance is a sensitive one but has to be recognised—it relates to the level of curriculum and assessment expertise amongst the decision-makers. Reference has already been made to the representative composition of the committees in the NCCA and how this acts to enable a consensus to occur, thereby securing a closer fit between advice to the Department of Education and adopted policy. However, it is arguable that this is at once a strength and a weakness, the weakness being that the inevitable lack of expertise across such a committee means that outmoded and traditional perspectives win the day. Granville's (1994) survey of NCCA committee members lends support to this view. He notes that 'the most revealing finding' of his study was the 'low rating' respondents gave themselves in the area of curriculum (of which assessment is a key component). He goes on to say (presumably from the perspective of an education officer) how difficult it is to stay committed to a project 'among personnel who do not consider themselves qualified in the essential skills and processes required' (p. 84).

This is not the only voice expressing reservation about the effect of representation: in 1992, Burke claimed that 'teacher unions and other special interests (e.g. management bodies) enjoy a virtual veto on the formulation of national educational policy'. He wrote: '[t]he NCCA … lacks a sufficient quota of independent voices with known expertise in the fields under investigation' (p. 201, cited in Granville, 1994). In view of the status of standardised testing in the NCCA advice to government *and* in the primary teachers' union stance on testing, Burke's argument about the influence of the unions fits well. This view is supported when one takes into account that the union representation doubled in the new council membership that was set up in 1991 while the representation of other groups remained at the 1987 level. The union membership occupied 29% of the total representation from 1991. Apart from the democratic concern of this balance on the council, there remains the major concern of the likely lack of expertise on the assessment committee, in the light of the overemphasis on standardised testing. I consider that there are two key lessons to be learned from this, first that it behoves those individuals who are charged with the decision-making to be fully aware of the available research in the relevant field, second, that a proper balance of informed views be facilitated by the committees and this means that a reasonable proportion of the committees should be

composed of *independent* educationists, and that the teaching profession should be represented by more than merely their union.

Recommendations and Conclusions

This final section offers several general recommendations for the design of assessment policy. On the basis of the theoretical, empirical and policy literature, these would appear to have application across countries—these are not just applicable to the Irish case, although they are directly in keeping with the NCCA's aspirations of designing assessments that enhance good-quality learning. The key difference is that, unlike the current proposals, the recommendations are grounded in constructivist notions of learning and knowledge. This section also reflects on the interpretation of *policy formation* and identifies further avenues for policy research in relation to assessment in Ireland.

On the basis of the research on assessment, the development of teacher skills in formative assessment, with reference to conceptions of learning and what constitutes valuable knowledge in the various curriculum areas, needs to be prioritised. Pre-service and in-service provision could attend to both cognitive processes/learning strategies and subject matter content. Whatever funding may be available for the development, distribution, administration and collation of standardised test instruments/results outside of school would best be redeployed, in my view, for this purpose. Since there is no point in standardising formative assessment—the type of assessment to be prioritised—there is little need for standardised testing. Attempts to standardise formative assessment are unnecessary and would only confuse teachers about its purpose (see Harlen and James (1997) for a full discussion). The policy in relation to summative assessment whereby achievements over a period of time (say a year) need to be summarised for purposes of reporting to parents could be developed. In particular, teachers could be provided with opportunities to share and arrive at consistent interpretations of pupils' achievements, collected in say, portfolio format.

A great deal needs to be done in helping the key partners in the education enterprise loosen their grip on what they undoubtedly perceive to be the *tried and tested* assessment methods. Education professionals who have expertise and understanding of the relevant fields of research will have the major role here but they ought to be supported by the media. Implicit in the above is that policy-makers need to recognise how learning occurs and be clear about what knowledge they deem to be important, since the assessment system must reflect both these dimensions.

The Irish assessment policy process is one that is ostensibly democratic in that the main interested parties are well consulted; yet, as this analysis argues, the teacher unions seemed to exert inordinate influence on the process and the outcome, and the current proposals are inadequately informed by contemporary research. Ultimately, however, I, like Broadfoot (above), believe that the widespread acceptability of educational measurement in the form of standardised tests lies at the root of the Irish assessment policy and is the root of the problem. In order to counter this, it is imperative that those who are allowed to lead and most influence the decision-making are not only properly informed by the available assessment research literature but are also sufficiently socially and intellectually skilled to persuade their committee colleagues and their political leaders towards more enlightened perspectives.

In this paper I have attempted to interpret and evaluate the intentions in various policy texts as well as understand some of the influences on the production of those texts. I have

highlighted the tensions around different kinds of assessment and noted, in particular, the problems of emphasising standardised tests at the expense of formative assessment methods. What I have not done is to examine the differing effects of assessment policy in practice. If we take a view of policy as a process rather than an instruction for unequivocal implementation and accept that policy writers cannot control the meaning of their texts but that this has to be subjected to interpretation and to recreation in the process, according to the values and experiences of those who do the implementing (Bowe *et al.*, 1992; Ball, 1994), then it follows that how different interests get prioritised or marginalised in practice merits close investigation. In the case of assessment in Ireland, the task for further research is to document the way teachers, both as a group at school level and as individuals at classroom level, implement and recreate the official policy, and, most importantly, to document the impact of assessment, whatever its format, on learners.

Acknowledgement

Correspondence: Kathy Hall, Faculty of Cultural and Education Studies, Leeds Metropolitan University, Beckett Park, Leeds LS6 3QS, UK.

NOTE

[1]   At the time of preparing this paper (September 1999) the Revised Curriculum has not been disseminated to teachers and schools.

REFERENCES

BALL, S. (1994) *Education Reform: a critical and post-structural approach* (Buckingham, Open University Press).

BLACK, P. (1993) Formative and summative assessment by teachers, *Studies in Science Education*, 21, pp. 49–97.

BLACK, P. (1997) Whatever happened to TGAT?, in: C. CULLINGFORD (Ed.) *Assessment versus Evaluation,* pp. 24–50 (London, Cassell).

BLACK, P. (1998) Learning, league tables and national assessment: opportunity lost or hope deferred?, *Oxford Review of Education,* 24(1), pp. 57–68.

BLACK, P. & WILIAM, D. (1998) Assessment and classroom learning, *Assessment in Education,* 5(1), pp. 7–84.

BOWE, R. & BALL, S. WITH GOLD, A. (1992) *Reforming Education and Changing Schools* (London, Routledge).

BROADFOOT, P. (1995) Educational assessment: the myth of measurement, in: P. WOODS (Ed.) *Contemporary Issues in Teaching and Learning,* pp. 203–230 (London, Routledge).

BROADFOOT, P. & GIPPS, C. (1996) Assessment developments in England and Wales: the triumph of tradition, in: A. LITTLE & A. WOLF (Eds) *Assessment in Transition,* pp. 134–153 (Oxford, Pergamon).

BROWN, M., MCCALLUM, B., TAGGART, B. & GIPPS, C. (1997) The validity of national testing at age 11: the teacher's view, *Assessment in Education*, 4(2), pp. 271–293.

BURKE, A. (1992) Teaching: retrospect and prospect, *Oideas* 39 (Journal of the Department of Education, Ireland).

CODD, J. (1988) The construction and deconstruction of educational policy documents, *Journal of Education Policy*, 3(3), pp. 235–247.

COOLAHAN, J. (Ed.) (1994) *Report of the National Education Convention* (Dublin, Government Publications).

COX, B. (1995) National Curriculum English, in: B. RABAN-BISBY WITH G. BROOKS & S. WOLFENDALE (Eds) *Developing Language and Literacy* (Stoke-on-Trent, Trentham Books/UKRA).

CROOKS, T. (1988) The impact of classroom evaluation practices on students, *Review of Educational Research,* 58(4), pp. 438–481.

CROOKS, T., KANE, M. & COHEN, A. (1996) Threats to the valid use of assessments, *Assessment in Education*, 3(3), pp. 265–285.

CURRAN, M. (1994) A study of the assessment practices of a selected sample of senior cycle primary school teachers, with particular reference to standardised testing. Unpublished MEd thesis (UCC).

DEPARTMENT FOR EDUCATION AND SCIENCE (1988) *Task Group on Assessment and Testing (TGAT). A report* (London, DES).

DEPARTMENT OF EDUCATION (1990) *Report of the Review Body on the Primary Curriculum* (Dublin, Department of Education).

DEPARTMENT OF EDUCATION (1992) *Education for a Changing World: Green Paper on education* (Dublin, Stationery Office).

DEPARTMENT OF EDUCATION (1995) *Charting our Educational Future: White Paper on education* (Dublin, Stationery Office).

DESFORGES, C., HOLDEN, C. & HUGHES, M. (1994) Assessment at key stage one: its effects on parents, teachers and classroom practice, *Research Papers in Education,* 9(2), pp. 133–158.

FULLAN, M. WITH STIEGELBAUER, S. (1991) *The New Meaning of Educational Change* (London, Cassell).

GIFFORD, B. & O'CONNOR, M. (Eds) (1992) *Changing Assessments: alternative views of aptitude, achievement and instruction* (Boston, Kluwer).

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

GIPPS, C. & MURPHY, P. (1994) *A Fair Test: assessment, achievement and equity* (Buckingham, Open University Press).

GRANVILLE, G. (1994) Professionalism and partnership: the NCCA committee system as a mechanism for curriculum policy formation. Unpublished MSc dissertation (Trinity College Dublin).

GRAUE, M. (1993) Integrating theory and practice through instructional assessment, *Educational Assessment*, 1(4), pp. 283–309.

HALL, K. (1997) Towards assessment for raising standards in the twenty first century: lessons from England and Wales, *Oideas: Journal of the Department of Education, Ireland,* 45, pp. 70–83.

HALL, K., OZERK, K. & VALLI, Y. (1998) Curriculum reform in contemporary English and Norwegian official documents, with particular reference to primary literacy, *European Journal of Intercultural Studies*, 10(1), pp. 85–104.

HANSARD (1.12.1987) House of Commons Reports, columns 780, 771–2.

HARLEN, W. (Ed.) (1994) *Enhancing Quality in Assessment* (London, Paul Chapman Press).

HARLEN, W. & JAMES, M. (1997) Assessment and learning: differences and relationships between formative and summative assessment, *Assessment in Education*, 4(3), pp. 365–379.

HARLEN, W., GIPPS, C., BROADFOOT, P. & NUTTALL, D. (1992) Assessment and the improvement of education, *The Curriculum Journal,* 3(3), pp. 215–230.

HERMAN, J., KLEIN, D. & WAKAI, S. (1997) American students' perspectives on alternative assessment: do they know it's different? *Assessment in Education*, 4(3), pp. 339–352.

HOGAN, P. (Ed.) (1995a) *Partnership and the Benefits of Learning: a symposium on philosophical issues in educational policy* (Maynooth, Educational Studies Association of Ireland).

HOGAN, P. (Convenor) (1995b) Educational research and the conduct of change: the Republic of Ireland in a European context, Symposium at the British Educational Research Association/European Conference on Educational Research Conference, Bath.

IRISH NATIONAL TEACHERS' ORGANISATION (INTO) (1997) *Teaching and Assessment: issues in assessment* (Dublin, INTO).

JAMES, M. (1994) Can we save formative assessment? Paper presented at the British Educational Research Association regional conference, University of Cambridge Institute of Education.

KVAVIK, R. D. (1974) Interest groups in a 'cooptive' political system: the case of Norway, in: M. O. HEISLER (Ed.) *Politics in Europe* (New York, David McKay).

LITTLE, A. & WOLF, A. (Eds) (1996) *Assessment in Transition: learning, monitoring and selection in international perspective* (Oxford, Pergamon).

MACRAE, S., BLISS, J. & ASKEW, M. (1994) Deconstructing scaffolding. Paper presented at the British Educational Research Association Conference, Oxford.

MCCALLUM, B., MCALISTER, S., GIPPS, C. & BROWN, M. (1993) Teacher assessment at key stage 1, *Research Papers in Education*, 8, pp. 305–327.

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

MURPHY, P. (1996) Integrating learning and assessment—the role of learning theories?, in: P. WOODS (Ed.) *Contemporary Issues in Teaching and Learning*, pp. 173–193 (London, Routledge).

MURPHY, P. & MOON, B. (Eds) (1994) *Developments in Learning and Assessment* (London, Hodder and Stoughton).

NATIONAL COUNCIL FOR CURRICULUM AND ASSESSMENT (NCCA) (1993) *A Programme for Reform: curriculum and assessment policy towards the new century* (Dublin, NCCA).

NATIONAL COUNCIL FOR CURRICULUM AND ASSESSMENT (NCCA) (1997) *Draft Curriculum: English* (Dublin, NCCA).

NUTTALL, D. (1987) The validity of assessments, *European Journal of the Psychology of Education,* 11(2), pp. 109–118.

O'SULLIVAN, D. (1992) Shaping education debate: a case study and an interpretation, *The Economic and Social Review*, 23(4), pp. 423–438.

RADNOR, H. (1994) The problems of facilitating qualitative formative assessment in pupils, *British Journal of Educational Psychology,* 64, pp. 145–160.

RUST, V. D. (1990) The policy formation process and educational reform in Norway, *Comparative Education,* 26(1), pp. 13–25.

SHIEL, G. & MURPHY, R. (1997) Curriculum profiles, in: *Teaching and Assessment: issues in assessment,* pp. 34–57 (Dublin, INTO).

STEERING COMMITTEE ON THE ESTABLISHMENT OF A TEACHING COUNCIL (1998) *Report of the Steering Committee* (Dublin, Stationery Office).

STIERER, B. (1990) Reading tests, in: T. HORTON (Ed.) *Assessment Debates* (London, Hodder and Stoughton).

STIGGINS, R. (1992) High quality classroom assessment: what does it really mean?, *Educational Measurement: Issues and Practice,* 4(4), pp. 263–273.

TORRANCE, H. (Ed.) (1995) *Evaluating Authentic Assessment* (Buckingham, Open University Press).

VON GLASERSFELD, E. (1994) Learning as a constructive activity, in: P. MURPHY & B. MOON (Eds) *Developments in Learning and Assessment* (London, Hodder and Stoughton).

WALSHE, J. F. (1997) From consultation to legislation: a review of recent developments in Irish education. Unpublished MEd thesis (University College Cork).

WILIAM, D. (1996) National Curriculum assessments and programmes of study: validity and impact, *British Educational Research Journal*, 22(1), pp. 129–141.