

[Willis, D. (1993). Assessment Trends in 1992: Future Directions Handicapped by a Legacy from the Past? *New Zealand Annual Review of Education*, 2, 245-262]

Assessment Trends in 1992: Future Directions Handicapped by a Legacy from the Past?

DEBORAH WILLIS

There has been a noticeable increase in the pace of assessment and curriculum reform as a result of the National Government's Achievement Initiative. However, these developments should not be seen in isolation and are more accurately a continuation of earlier changes. One of the most significant developments was the abolition of the University Entrance Examination and its replacement with the internally assessed Sixth Form Certificate, the assessment of which has moved progressively towards achievement-based criteria. Other examples include the development of student records of achievement (Nightingale, 1990) and the increased use of diagnostic assessment in the junior school in areas such as reading and mathematics.

The reform process has not always been a smooth one. For example, intense debate took place in the years preceding the Labour Government's removal of the University Entrance examination. Indeed the merits of external examinations are still argued by individuals and groups from many sectors of society. The intensity of feeling is not surprising because although assessment is only one form of discourse about schooling practice it is "uniquely privileged" (Berlak, 1992, p. 191). This is because of the perceived importance of assessment for providing information about student, school and system performance. It represents a powerful mechanism of control. Furthermore, Berlak suggests that:

Particular forms of tests and assessments represent particular forms of discourse, that is they produce particular ways of talking and communicating with others about the schooling and educational process (p. 186)

246 Deborah Willis

Traditionally, assessment discourse has been informed by the terminology and assumptions of psychometrics – the science of measuring mental capacities and processes (Berlak, 1992). These technicist notions view assessment as an objective measure of predetermined and context-free constructs such as academic achievement or reading ability. Assessment (both norm and criterion-referenced) has been developed using a mathematical model designed to discriminate between individuals whether this be done normatively through the calculation of a standard score, percentile or (in the case of criterion-referenced measurement) by the establishment of standards of mastery. Such context-free forms of assessment have been seriously questioned as being inadequate indicators of a range of learning outcomes and processes (Willis, in press) as well as discriminating against particular groups of students (e.g., Wolf et al., 1991). Because of these and other weaknesses they do not "warrant making qualitative statements about a person's (or group's) performance" (Berlak, 1992, p. 11). Therefore, writers such as Blackmore (1988) and Berlak have been particularly critical of standardised testing and other forms of technicist assessment as vehicles for monitoring the achievement standards of students, schools and the system as a whole.

In an attempt to stimulate a more equitable assessment culture, Blackmore (1988) has argued for assessment to become an integral part of the educative process whereby it provides feedback to student learning "and as a dialogue between those actively participating in the learning environment" (p. 57). Bird and Willis (1992) have suggested that this educative assessment be characterised by the following themes.

1. Attempts are made to assess a range of learning (processes and outcomes) incorporating cognitive, affective and motivational aspects.
2. The assessment used relates directly to course-work carried out thus forming a link between curriculum and assessment.
3. Assessment takes account of the prior experiences both in terms of learning and wider cultural perspectives of all students.
4. A key purpose of assessment is the diagnosis of learning needs.
5. Students are involved in self-evaluations of their learning.

Mode of assessment alone will not of course ensure "a democratic or rewarding learning experience or fair outcomes" (Blackmore, 1988, p. 51). More far reaching change is required:

Fundamental changes in the way we think about education and the process of schooling must accompany the effort to rethink assessment theory and practices (Berlak, 1992, pp. 8-9).

This is particularly true in view of the inconsistencies which can be seen as the legacy of the technicist perspective and which may impede attempts to develop educative forms of assessment. This debate has not been lost on New Zealand educational policy makers (Irving, 1992) as the policy directions discussed here will show. The purpose of this paper is to identify some significant trends in New Zealand assessment policy and to locate them within a wider discourse.

The trends are interconnected and should not be seen in isolation. Firstly the trend towards separation of assessment of individual achievement and national monitoring. Secondly, the trend towards assessment within the classroom as a means of assisting in the diagnosis of student learning needs. Thirdly, the trend towards the recognition that assessment must be fair to all students. Finally the trend towards the professional development of teachers as an integral part of assessment reform. The discussion which follows does not extend to the assessment reform taking place in the senior secondary and tertiary sectors under the auspices of the New Zealand Qualifications Authority's Framework. However, some of the issues raised do have relevance for this paper.

Assessment Trends in 1992

Trend 1: Separating the assessment of individual achievement from national monitoring

The National Government's Achievement Initiative provides the policy umbrella under which the majority of government sponsored assessment developments are taking place. The Achievement Initiative seeks to meet the learning needs of individual students while at the same time addressing "the overall national monitoring of the education system" (Irving, 1992b, p. 1)

The conflict inherent in assessment systems designed to enhance the learning of individuals *and* monitor national standards has been much discussed. For example, I have highlighted elsewhere (Willis, 1992) the competing interests of professional, market and management accountability models each of which reflects particular ideological and epistemological assumptions. The professional model reflects a complex and context-bound view of the learning process that recognises the importance of process as well as non-cognitive aspects of learning. In contrast, both the management and market models are committed to an empiricist approach to education whereby "...knowledge can be broken down into a series of separate pieces through the operation of scientific methods commonly described as positivistic A failure to recognise the complexity of the learning and teaching process has contributed to assessment that is simplistic and concerned only with observable outcomes. Examinations and standardised tests play a key role in providing scores that serve as useful market indicators" (Willis, 1992, p. 208)

In a situation where national and individual assessments are incorporated into the same testing process the needs of the learner may be secondary to the requirements of national accountability (Murphy, 1990). The implications for learners are particularly serious when the testing is of the high stakes variety as it is in the United Kingdom.¹ The pressure resulting from high stakes testing is counterproductive to the creation of a positive assessment climate in which the learning needs of students can be met.

One of the most notable features of recent New Zealand assessment reform has been a recognition of this conflict and an articulation of the separate requirements of formative assessment and national monitoring:

monitoring individual student progress and overall national monitoring of educational standards are two forms of assessment carried out for different purposes and therefore requiring different assessment instruments and procedures (Irving 1992b, p. 2).

A. NATIONAL MONITORING

In New Zealand the intention of national monitoring is to:

build up a picture of student achievement nationally over time and thus provide clear information on how well national standards are

being maintained and where improvements might be needed (Irving, 1992b, p. 2).

For reasons of economy and practicality it is proposed that at four yearly intervals a sample (approximately 5%) of 8 and 12 year old children will be asked to take part in a range of assessment tasks. The assessment is to be standardised and will take place at a "set time during the year" (Irving, 1992a, p. 7). What is not yet clear is what areas of the curriculum will be covered by such tests although the terms of reference require that the assessment tasks span "the full range of outcomes likely to be of interest to educators, policy makers and the community". In an attempt to examine these outcomes a wide range of assessment tasks is to be included (Irving, 1992a).

It is to be hoped that the New Zealand model of national monitoring will avoid the pitfalls of those used elsewhere. These concerns reach to the heart of the culture of testing upon which such systems have been based. Traditionally assessment for national monitoring has made use of standardised procedures (Berlak, 1992). Recently the psychometric model on which the majority of norm and criterion-referenced testing is based has been questioned (e.g., Wolf et al., 1991; Berlak, 1992). Berlak argues that the move to school-based decision making is undermined by a continuing emphasis on standardised testing which is not only based on flawed assumptions about the nature of learning and assessment but perpetuates a centralising tendency. Berlak argues "that the psychometric paradigm embodies a particular regime of scientific truth" (p. 194). A number of assumptions lie at the basis of this paradigm. These serve to shape assessment discourse and practice in particular ways. The assumptions are as follows:

1. There are or can be universally accepted meanings of educational constructs, or where conflicts and contradictions exist, these may be transformed into technical problems which may be settled by experts;
2. Tests constructed according to established technical requirements are morally neutral scientific instruments which stand outside of history and culture;
3. Human cognition and affect may be separated at least for the purposes of measurement;

4. Schools and educational systems can and should be managed and controlled from the center (ibid, p. 195).

Given such assumptions one must examine the extent to which psychometrically based tests can tell us anything useful about the quality of educational provision or achieve one of the Minister of Education's, that

Form Two is the end of primary school for most students and so would enable us to gauge the overall health of the primary sector (Smith, 1992, p. 15).

Berlak explores the possibility of an alternative framework for national monitoring based on what he terms a contextual paradigm. It is beyond the scope of this paper to examine his proposals in detail. However, it is important to acknowledge the existence of alternative discourses during the process of policy development and to articulate their usefulness in terms of stated aims.

B. ASSESSMENT OF STUDENT ACHIEVEMENT

One of the most significant aspects of current proposals is the pre-eminent focus given to the needs of individual learners, the intention being to "help to improve (student) understanding" (Smith, 1992, p. 6). The impact of assessment on students has been documented by Crooks (1988) and acknowledged by the Ministry of Education. Assessment "should focus on the individual student rather than on comparisons between students" (Irving, 1992b, p. 1). There is no doubt that their needs are best served by assessment that is classroom-based, seeking to be both formative and diagnostic (Irving, 1992b). This trend towards classroom-based assessment will be discussed in the following section.

Trend 2. Classroom based assessment to identify learning needs

The second trend in assessment in 1992, and one that is clearly related to the first, is the recognition that effective formative assessment is located in the classroom. Not only is assessment being seen as belonging to the "teacher as part of the normal teaching and learning process" (Irving, 1992b, p. 1) but:

it should be part of the normal classroom situation and context as an integral, ongoing part of the curriculum to enhance each student's learning, and teacher understanding of that learning (ibid, p. 2).

This commitment to assessment as a means of enhancing student learning characterises a professional model of accountability (Willis, 1992a) whereby learning is understood as a complex process that takes place partly, although not exclusively, as a function of the relationships between students, teacher and the wider educational context. That is not to say, of course, that learning is the exclusive right of the school. Moreover, what is taught is not necessarily what is learnt.

By basing assessment in the classroom the opportunity is created for students to be participants in an educative assessment process (Blackmore, 1988). If children are to be assisted to reach their potential in all areas of the curriculum they need to be able to "demonstrate what they know, understand and can do" (Murphy, 1990, p. 48). Murphy goes on to suggest that:

the involvement of pupils and teachers in the assessment process is a positive and motivating experience which provides helpful feedback about progress at regular intervals, so that information can be useful in planning future teaching and learning strategies (p. 48).

Currently in New Zealand assessment exemplars are being developed which will assist teachers in diagnosing strengths and weaknesses in student learning. The Minister has made it clear that:

assessment especially at the primary school level, is not so much about establishing what a student knows or does not know but why a student understands or doesn't understand and how we can improve their knowledge (Smith, 1992, p. 6).

However, Simpson (1990) is critical of the ability of many criterion-referenced tests to diagnose. She argues that both norm and criterion-referenced assessment are based on the assumption "that differences in pupil attainment are the natural expression of differences in pupil characteristics" (p. 172). In addition, according to Simpson, both forms of assessment assume that students only learn what they have been taught. By testing only those clearly specified course objectives, criterion-referenced assessment fails to acknowledge the range of other

learning, understandings and misunderstandings that take place as the result of any learning experience (see for example Alton Lee and Nuttall, 1992; White, 1992).

While it is important for teachers to be clear about what it is they are trying to achieve, constructivist research has also highlighted problems associated with the use of pre-specified objectives and achievement structures that are external to the learner (Prawat, 1992). These seem likely to make meaningful self-evaluation in which students are directly involved in the development and assessment of what they are learning difficult to achieve (Codd et al., 1992).

It is also intended that individual student progress in the core areas of the National Curriculum be assessed at key stage transition points. This is to assist in the targeting of resources to particular groups of students and to allow schools to "determine the needs of groups of students as they enter new schools" and to "assess the relative performance of their students against national standards in the areas of learning covered by the items" (Irving, 1992b, p. 3).

Irving makes it quite clear that these tests will be used in combination with other forms of assessment to allow schools to identify the needs of entering students. Further clarification is clearly needed as to the nature of these other forms of assessment. In view of the importance given to in-class assessment, methods such as student profiling will be of use here. There is also a hint that such testing will be a further way of moderating national standards through the comparative use of the results. There is certainly no suggestion that this will be of the "high stakes" variety yet the use of standardised tests to measure quality of achievement and to assess need is questionable for the reasons outlined above.

Because of the instability of learning in the first couple of years it is not proposed to carry out formal testing with children entering school. Existing methods of assessing such things as language and social development will be used in a more systematic way. However, for those moving from Standard 4 to Form 1 and Form 2 to Form 3 it is proposed to administer more formalised tests. At present it is intended that these will take the form of "item banks of nationally standardised items" (Irving, 1992b, p. 3) although the final form of the items has yet to be decided. Initially the items will be in Mathematics and Science and later English, Te Reo Maori and Technology.

I have already highlighted some of the difficulties associated with the use of standardised testing including the difficulty in assessing complex forms of learning among which are non-cognitive aspects of affect and motivation. I will therefore move on to the structural base of the key stage tests. This rests on eight levels of learning described in the National Curriculum (1991). The draft National Curriculum makes it clear that:

the achievement levels are not to be confused with years of learning. In any class there will be some students who are achieving at higher levels than those of most of their peers...students will also be operating at different levels in different subjects, and even in the achievement aims of the same subject in any one year...teachers will be expected to know what levels of achievement learners are likely to reach at particular stages to assist them in their planning of appropriate activities (p. 22).

It has been assumed that there is a hierarchical relationship between levels of learning and that the learning set down in an earlier level establishes the base for later mastery. This "vertical" view of knowledge transfer has been questioned by recent constructivist research (Prawat, 1992):

[the] model of powerful ideas applied to specific aspects of the environment, is much more likely to result in the acquisition of well-organised and highly transferable networks of knowledge in youngsters compared with the traditional model that is based on a clear delineation between different types of knowledge (i.e., lower order/higher order) and between learning and application contexts (pp. 381-382).

With reference to New Zealand, Woods (1992) has suggested that the eight levels of learning are somewhat arbitrary. In addition Elley (1992) questions the assumption that:

children progress up a ladder of standards in the same sequence. In Mathematics you must learn number combinations before fractions. But there are few clear-cut sequences in Social Studies, or English, or Science. Even if there were, the actual setting of the standards can

only be arbitrary in most cases, and the gaps between them are bound to vary. There is very little research to support such a system (p. 2).

Although more specifically age related, the ten levels used in the United Kingdom bear a resemblance to the eight that underpin the New Zealand draft National Curriculum. These have been seen as a simplistic and inaccurate attempt to fit learning (which is inherently complex) into a convenient reporting mechanism (Murphy, 1990). In the first place, learning is rarely hierarchical nor is its progress smooth. Furthermore, in a paper that examines technicalities of the assessment system in the UK, Cresswell and Houston (1991) highlight the complex relationship between context and attainment:

Pupils of different ages make qualitatively different responses to the same task. Maturation thus involves more than simply growth along a single dimension of competence (p. 67).

In addition to identifying need, the key stage assessment tasks will be designed to assist in the "better targeting of resources" (Irving, 1992b, p. 3). The question needs to be asked who will be doing the targeting. Traditionally standardised tests have been associated with control by the centre. A centralising tendency may be moderated by close involvement of educators in the development of item banks. Indeed Irving (1992) makes the point that "the successful development of item banks will require the close involvement of schools and the participation of practising teachers" (p. 3). Furthermore, the chance of such monitoring becoming "high stakes" will be reduced if the decision to use them is one that rests with the school. Of course if test results are associated with the allocation of scarce resources few schools will wish to opt out of being involved in the tests. The nature of these tests will also be affected if the results are seen as providing essential input into the Education Review Office reviews.

Trend 3: Recognising the need for assessment to be fair to all students

A feature of the Achievement Initiative has been the acknowledgment that assessment both within and external to the classroom has a major impact on student motivation and self esteem as well as on the quality

of student learning. Not only is recognition given to the need to use a variety of assessment forms but differences in the cultural interpretation and valuing of forms of knowledge are acknowledged.

Assessment must recognise and value the differences in race, culture, language, gender, background and experience that students bring to their learning, so that it does not exclude some groups of students. This means that in introducing any new assessment procedures care must be taken to use methods that do not disadvantage particular students or groups. Educators must also ensure that the appropriate emphasis is given to the knowledge, skills and attitudes of Maori and Pacific Island students, and of both girls and boys (Irving, 1992b, p. 1).

Several meetings have been held in the past two years which have specifically addressed the assessment needs of girls and Maori students. In March 1991 a hui for Maori educators was held. Twelve months later a seminar on issues of assessment for girls was organised jointly by the Ministries of Education and Women's Affairs.

The scope of the discussion at the hui was wide ranging. The written record indicates that participants felt some disquiet at the extent to which assessment procedures have contributed to negative outcomes for Maori students. In addition to calls for increased participation and control over future assessment developments it is clear that assessment reform needs to recognise the existence of multiple sets of priorities arising from different educational contexts and aims (Prawat, 1992). Hui participants identified a number of successful methods of assessment for Maori students in mainstream education. These included the use of individual profiles and peer tutoring. Calls were made for the development of assessment and evaluation in Te Reo Maori in *all* areas of the curriculum. In practical terms as a result of the hui a survey has been undertaken to "investigate assessment procedures and practices currently used in Maori immersion classes and Kura Kaupapa Maori" (Irving, 1992b, p. 7).

A number of papers were presented at the seminar on assessment for girls. This meeting provided researchers, policy makers and teachers with an opportunity to examine the ways in which assessment has disadvantaged girls and to explore alternative approaches.

An obvious vehicle for the wider circulation of issues relating to fairness is the handbook on assessment for the Achievement Initiative being prepared by the Ministry of Education at the time of writing. It is intended that the handbook will be provided to all teachers to give guidance on a range of assessment matters. This document is to include sections on:

issues related to fairness such as the impact of assessment on boys and girls, Maori students, Pacific Island students and students with special needs such as ESL (Irving, 1992).

In addition, the terms of reference for the development of assessment for the key transition points and national monitoring make it clear that the test developer will consider "assessment issues relating to gender, cultural difference and the differently abled".

A commitment to fairness also extends into specific curriculum developments. For example, the draft Science Curriculum (1992) seeks to "make science more accessible to girls and to Maori students" (p. 11). It states that girls benefit more from science education when "the language strengths and preferred co-operative learning styles of girls are recognised in the range of teaching and assessment strategies used" (p. 10). Similarly Maori students are likely to benefit when:

the preferred learning and communication styles of Maori students are recognised in the teaching and assessment methods used, for example, co-operative learning, holistic approaches, recognition of oral contributions for assessment purposes (p. 11).

In the body of the curriculum document recognition is given to the subjectivity of knowledge and attempts are made to focus on students' diverse experiences, the learning context, and a range of assessment methods all of which are consistent with an educative approach to learning.

However, inconsistencies are apparent in this trend towards fairness. Firstly, the Achievement Initiative gives predominance to the core subjects of Mathematics, English, Science and Technology. Even more specifically, curriculum and assessment development priorities have focused on Mathematics and Science. In doing so a hierarchy is created within the National Curriculum. This devalues "the rest" to both

students, parents and the wider community. The students who have not been ranked highly in these subjects (including Maori students and girls) are likely to be disadvantaged unless significant changes can be made to the curriculum and assessment in these subjects. Crooks (1992) warns "if assessment changes de-emphasise these subjects relative to Maths, Science and Technology, the net effect might be to disadvantage girls in measured achievement and in school motivation" (p. 2).

Secondly, there is the disquieting increase in the use of multiple choice questions in the 1992 School Certificate English exam. Not only is it difficult for such questions to assess non-cognitive learning but they are less likely to require a holistic form of answer in favour of specific factual responses preferred by boys (Crooks, 1992). It is to be hoped that we are not witnessing a trend here.

Trend 4: Integration of assessment reform with teacher development

I have argued elsewhere (Willis, 1992b) that the implementation of educative assessment requires that teachers fully understand all aspects of the assessment system. This includes "an appreciation of the strengths and weaknesses of different forms of assessment and familiarity with the ideological and theoretical assumptions underpinning assessment models" (p. 8). Effective educative assessment also relies on a range of technical skills including an ability to use a range of assessment procedures, marking and moderation skills. In an assessment system that is to be largely based in the classroom, teacher development is critical to its success not only in the impact of assessment on individual students but also in ensuring that assessment operates evenly within and between schools.

A feature of recent reform has been the integration of teacher, assessment and curriculum development by policy makers. In May a National Workshop was held on Professional Development in Educational Assessment. Included in its 16 recommendations were the following:

that the Ministry of Education accord a high priority to teacher development in the use of educational assessment to support and optimise student learning;

that teacher development should promote informed reflective practice and a problem solving approach;

that a focus of professional development should be to increase awareness of ways that assessment can impede learning and to provide ways to encourage action to avoid these harmful effects;

that professional development on assessment take account of the needs of Maori, Pacific Island and other ethnic groups, girls and women and students from low socio-economic backgrounds (pp. 24-25).

Development of the procedures and instruments of assessment is well under way (albeit within the severe constraints of the PPTA moratorium). For example, assessment exemplars are being developed in the core areas of Mathematics and Science. Additional resources will include the assessment handbook referred to above. The list of contents suggests that the handbook contains plenty of information on the practicalities of testing:

- making judgements about student performance against the learning objectives specified in the new curriculum statements;
- the use and development of assessment examples/activities across the curriculum;
- issues related to fairness such as the impact of assessment on boys and girls, Maori students, Pacific Island students and students with special needs such as ESL;
- ways of recording and reporting student achievement in relation to the levels specified in the new curriculum statements (e.g., the development of individual and school profiles of student achievement);
- key assessment terms
- the features and development of a sound schoolwide assessment policy (Irving, 1992b, p. 5).

In conjunction with information designed to increase teachers' skills in the practicalities of assessment there needs to be a conscious effort made to include an exploration of the assumptions underlying various models of assessment. The recommendations of the National Workshop on Professional Development in Educational Assessment do suggest a view of the teacher as a participant in the learning process who should have

an understanding of factors that both encourage and impede the learning of particular groups of students. It is only then that educative methods of assessment can be associated with pedagogy that is:

concerned with the development of a sense of autonomy, self-confidence and respect, with gaining a greater understanding of the world with the nurturing of the ability to participate in the democratic management of society and with undertaking constructive and reflective thinking (Blackmore, 1988, p. 51)

It will be interesting to follow the implementation of teacher development programmes in 1993 to observe whether these recommendations are implemented in a way that is consistent with educative principles.

Involvement of teachers in some assessment activities, such as moderation, may assist this to occur. Ingvarson (1991) cites moderation of student learning as a means of increasing the professionalism of teachers. It is important to bear in mind that the moderation referred to by Ingvarson is not statistical moderation (currently carried out in the internally assessed School Certificate Modular Science). Alternative methods such as the use of visiting moderators and consensus moderation offer teachers the opportunity to develop forms of educative assessment in ways that do not compromise comparability. Thus:

A flexible and responsive approach to moderation could maximize the opportunity for teachers to experiment with different sorts of course-work assignments and gain experience of reflecting on the overall quality and effectiveness of their syllabus and teaching methods (Murphy and Torrance, 1988, p. 108)

Examples of this kind of moderation can be found within the existing Sixth Form Certificate system (see for example Codd et al., 1991). It is to be hoped that this can be adopted more widely.

Conclusion

During 1992 the direction of assessment reform has become clearer. I have suggested that in general terms these directions are likely to support student learning. However, some inconsistencies are still present such as a lack of theoretical justification for the levels of learning

used by the National Curriculum, and suggestions that standardised test questions might satisfactorily measure complex forms of learning and the quality of educational processes and outcomes. These inconsistencies can be seen as the legacy of technicist views of assessment. In New Zealand there are signs that the technicist approach is being challenged by the emphasis increasingly placed on the needs of individual students, recognition of the importance of learning contexts and calls for all teachers to understand the assumptions and processes underlying assessment. Hopefully by the end of 1993 it will be possible to find specific examples of assessment that use the educative characteristics proposed earlier in this paper. Indeed it is important that further moves be made to replace restrictive technicist notions with educative models of learning and teaching.

Note

1. A regime of high stakes testing occurs in situations where assessment is used to make comparisons between individuals and schools. Significant "rewards" in the form of resources are made available by Government agencies on the basis of test scores. In addition the publication of scores allows participants in the educational marketplace (i.e., parents) to make choices. See Gipps (1990) and Broadfoot (1989) for an analysis of such forms of testing in the United Kingdom.

References

- Alton-Lee, A. and Nuttall, G. "Students learning in classrooms: curricular, instructional and sociocultural processes influencing student interaction with curriculum content". Paper presented at the annual meeting of the AERA, San Francisco, 1992.
- Berlak H, "The need for a new science of assessment" in Berlak, H., Newmann, F., Adams, E., Archbald, D., Burgess, T., Raven, J. and Romberg, T. (eds) *Toward a New Science of Educational Testing and Assessment*. Albany: State University of New York, 1992.
- Bird, L. and Willis, D. "Assessment at upper secondary school: silenced voices of students in an era of reform". Paper presented at AARE/NZARE Conference, Geelong, 1992.
- Blackmore, J. *Assessment and Accountability*. Geelong. Vict.: Deakin University Press, 1988.
- Broadfoot, P. "Changing priorities in assessment: an international reform agenda." Paper presented at Conference on Qualifications for the 21st Century, Wellington, 1992.

- Broadfoot, P. "The Significance of Contemporary Contradiction in Educational Assessment Policies in England and Wales", Paper, University of Bristol, 1989.
- Codd, J. McAlpine D., and Hansen, S. "Achievement and consensus: an evaluation of the Sixth Form Certificate practical art moderation trial", Palmerston North: Massey University, 1990.
- Codd, J. McAlpine, D. and Poskitt, J. "The achievement initiative: myths and mysteries of assessment." Paper presented at the NZARE Conference, Dunedin, 1991.
- Codd, J., McAlpine, D. and Poskitt, J. "The reform of student assessment in New Zealand: whose agenda? what effects?" Paper presented at AARE/NZARE Conference, Geelong, 1992.
- Cresswell, M. J. and Houston, J. G. "Assessment of the National Curriculum – Some fundamental considerations" in *Educational Review*, 43, 1, 1991, pp. 63-78.
- Crooks, T. "Some factors affecting assessment which may act against the interests of girls" Paper presented at Seminar on Issues of Assessment for Girls, Wellington, March, 1992.
- Crooks, T. "The impact of classroom evaluation practices on students" in *Review of Educational Research*, 58, 4, 1988, pp. 438-481.
- Elley, W. "Is standards-based assessment feasible? Some blunt words about our blunt instruments", Paper presented at the National Workshop on Professional Development in Educational Assessment, May 1992.
- Gipps, C. "The social implications of national assessment" in *The Urban Review*, 22, 2, 1990, pp. 145-159.
- Hui for Maori Educators, Wellington, Ministry of Education, March 1991.
- Ingvarson, L. "Enhancing professional skill and accountability in the assessment of student learning". In Stake, R. (ed) *Advances in Program Evaluation*. Vol. 1 Part B Effects of Mandated Assessment on Teaching, Greenwich, Conn: JAI Press, 1991.
- Irving, J. "Educational assessment and monitoring in New Zealand" Ministry of Education, Wellington, August, 1992b.
- Irving, J. "Monitoring the education system" Paper presented at the Auckland Primary Principals' Association Conference, Auckland, April, 1992.
- Ministry of Education *Science in the National Curriculum* (draft). Wellington: Learning Media, 1992.
- Ministry of Education *The National Curriculum: A Discussion Document*. Wellington: Learning Media, 1991.

- Murphy, R. "National assessment proposals: analysing the debate", in Flude, M. and Hammer, M. (ed) *The Education Reform Act*, 1988. London: Falmer Press, 1990.
- Murphy, R. and Torrance, H. *The Changing Face of Educational Assessment*. Milton Keynes: Open University Press, 1988.
- Nightingale, D. "School Leaver documentation developmental project: evaluation report on the trial in nine New Zealand secondary schools." 1989-1990 Wellington: Ministry of Education, 1990.
- Prawat, R. S. "Teachers' beliefs about teaching and learning: a constructivist perspective" in *American Journal of Education*, 100, 3, 1992, pp. 354-395.
- Rowan, B. "Commitment and control: alternative strategies for the organizational design of schools" in *Review of Research in Education*, 16, 1990, pp. 353-389.
- Simpson, M. "Why criterion-referenced assessment is unlikely to improve learning" in *The Curriculum Journal*, 1(2), 1990, pp. 171-183.
- Smith L. "Assessment and evaluation in the primary school" Speech presented at the University of Canterbury, May 1992.
- White, R. T. "Implications of recent research on learning for curriculum and assessment" in *Journal of Curriculum Studies*, 24 (2), 1992, pp. 153-163.
- Willis, D. "Learning and assessment: exposing the inconsistencies of theory and practice." *Oxford Review of Education* (in press).
- Willis, D. "School-based assessment: underlying ideologies and their implications for teachers and learners." Education Department, Victoria University of Wellington, 1992b.
- Willis, D. "Educational assessment and accountability: A New Zealand case study" in *Journal of Educational Policy*, 7 (2), 1992a, pp. 205-221.
- Wolf, D., Bixby, J., Glenn, J. III & Gardner, H. "To use their minds well: investigating new forms of student assessment" in *Review of Research in Education*, 17, 1991, pp. 31-74.
- Woods, M. "New assessment skills for the National Curriculum and the Qualifications Framework" Paper presented at the National Workshop on Professional Development in Educational Assessment, Wellington: May 1992.

The author

Deborah Willis is a Senior Lecturer in the Education Department at Victoria University of Wellington. Her research interests lie in the area of assessment and accountability. She is currently engaged in a research project on Learning and Assessment in fifth form modular science.