

Primary school assessment in New Zealand: Some issues and concerns

Gregory Lee, Howard Lee and Debbie Hill

“During the last fifteen years, educators and the public alike have witnessed the introduction of a whole host of educational ‘reforms’ covering, besides other things, school administration, curricula, and assessment.”

During the last fifteen years, educators and the public alike have witnessed the introduction of a whole host of educational ‘reforms’ covering, besides other things, school administration, curricula, and assessment. Each was introduced ostensibly to enhance educational ‘efficiency’ and achievement ‘standards’, to promote greater ‘transparency’ in educational and administrative activities, and to secure more ‘accountability’ on the part of the teaching profession (Broadfoot, 1996; Butterworth & Butterworth, 1998; McKenzie, 1997 and 1999).

Historical background

While much of the thinking underpinning these reforms has been examined by academics and other commentators,¹ (see Lee & Lee, 1999a, 1999b; Lee & Lee, 2000), the fact remains that few issues have attracted more scrutiny and controversy than that of assessment in both primary and secondary schools. There is little doubt, in the public mind, that ‘assessment’ usually represents one of the most important activities in which teachers and pupils will be engaged. Research into New Zealand’s history of education reveals, however, that there is nothing new about the public obsession with assessment within the nation’s schoolrooms (Lee & Lee, 1998; McKenzie, Lee & Lee, 1996; Openshaw, Lee & Lee, 1993). There is plenty of evidence to conclude that from the introduction of the primary school standards’ examinations in 1878, parents, employers, and other interest groups have taken a special interest in assessment (Lee & Lee, 1998). And there is every reason to predict that this interest is not likely to diminish in the foreseeable future.

How might we explain the high

level of public interest in school-based assessment? The answer appears to rest, in large part, with the long-held, intimate association between school curricula, assessment, and pupil and teacher ‘accountability’, independent of whether or not a curriculum is to be examined externally (as has traditionally been the case in New Zealand secondary schools) or internally. In other words, the practice of determining or ‘measuring’ pupils’ progress in the various subjects prescribed in the national curriculum – (in existence since 1877 in New Zealand primary schools) – seen as a means of ascertaining pupils’ ‘real understanding’ of officially prescribed ‘knowledge’ (information) – has steadily captured the public imagination (Lee, 1991). Having gained this status, subsequent efforts to persuade the public about other uses to which assessment can and ought to be put were likely to meet with resistance (McKenzie, Lee & Lee, 1996; Openshaw, Lee & Lee, 1993).

At a time (in the late nineteenth century) when primary teachers were expected to prepare youths as candidates for standards’ examinations, wherein successful performance at the more advanced levels enabled young girls and boys to gain qualifications for employment and, later, for further schooling purposes, there was little questioning about whether or not pupils were *actually educated* as opposed to *formally schooled and certificated* (Cooney, 1999; McKenzie, 1999). In recent years, however, educators have gained a greater appreciation of some of the distinctions that can and should be made between education and schooling (see Hill, 1997). These distinctions are, of course,



Gregory Lee
School of Education
University of Waikato



Howard Lee
School of Education
University of Otago



Debbie Hill
School of Education
University of Waikato

sharpened once educators become more aware of the very real tension existing between conflicting notions of assessment. More importantly, perhaps, the conflict between summative and formative assessment philosophies², as Nisbet (1993) laments, tends to produce conservative outcomes:

In the attempt to reconcile... two very different requirements, the demand for accountability, for certification and selection, for hard evidence, will take precedence over the equally legitimate requirement that assessment should promote learning...the two cannot be readily reconciled.

(cited in Hill, 1999, p.177).

Assessment: policies and practices

Mary Hill's (1999) research, in a New Zealand primary schooling context, provides additional insight into the competing assessment discourses. It also examines the strategies that New Zealand primary teachers have employed to cope with "assessment expectations [that] were driven from above" (p.180). Like Nisbet, Hill notes the definite tension that exists between "the summative and formative functions of assessment" (p.177), and concludes that the former has gained ground at the expense of the latter. She explains that summative assessments are often undertaken in response to the externally-imposed pressures and constraints that frequently accompany national schooling systems, and that notions of 'performance', 'effectiveness', and 'accountability' invariably become part of schools' practice whenever summative considerations dominate (as they tend to do).

Hill's work also attests to a decrease in emphasis on formative assessment within Ministry of Education publications between 1990 and 1994, and a move toward more summative ones, despite much official rhetoric to the contrary. The emphasis on individual learners and how assessment and evaluation can enhance pupils' learning and

development that was evident throughout the 1970s and 1980s (for example, in Department of Education documents) gave way to a more prescriptive and interventionist approach, one that foreshadowed the introduction of specified achievement objectives and the national monitoring of standards, and placed more weight on conducting external reviews of school 'efficiency' (via the Education Review Office (ERO)), under the guise of promoting greater institutional and teacher accountability (Ministry of Education, 1991; Ministry of Education, 1993; Thrupp & Smith, 1999). As Hill shrewdly observes, these policies constitute "forms of surveillance over both students' achievement and teachers' practice" (1999, p.179). Furthermore, they have led to "changes in teachers' classroom assessment practices" (p.179). Some of these practices, she argues correctly, have been antithetical to "the primary purpose of assessment", which is "to improve learning and learning programmes" (p.179).

Primary teachers' assessment strategies

How have primary teachers coped in the assessment environment of the 1990s? Hill's research is again invaluable in this connection. She reveals that teachers in her study were aware of the 'top down' nature of policy delivery and of the official expectation that they keep regular and comprehensive records of their pupils' achievements and progress therein. This detailed record-keeping was deemed necessary because of the achievement objectives that were prescribed in the New Zealand Curriculum Framework (Ministry of Education, 1993), and the knowledge that ERO staff expected to see written 'evidence' of teachers' assessment activities and performance (Hill, 1999; McKenzie, 1999; Thrupp & Smith, 1999). Nevertheless, Hill explains that in her study, there were distinct differences between the

teachers' approaches to assessment (see pp.181-183). Some chose to "keep most of their ongoing assessment information in their heads" (p.183); others "focused almost entirely on checking progress against the achievement objectives" (p.183), while the remainder tried to achieve some sort of balance between formative and summative assessment. This balance was achieved not only by "[systematically] planning for formative assessment during teaching and learning activities" (p.182) but also by being alert and responsive to unplanned situations in their classrooms in which assessment information could be gathered spontaneously. Hill observed, however, that the achievement objectives specified in the various curriculum documents were *not ignored* by teachers. In fact, these achievement objectives underpinned their teaching objectives.

Given the Ministry of Education's (1991) assertion that "assessment is an essential part of schooling" and that "[it] is central to the effective implementation of the National Curriculum" (p.24), it is entirely understandable that teachers would wish to be more informed about the range of assessment options available to them, and their respective strengths and weaknesses. We suggest, in this context, that it is vital for teachers to be able to distinguish between assessment that is conducted for educational reasons (formative assessment) and that which is undertaken merely to satisfy technical and accountability requirements in an uncritical manner (Hill, 1999; McKenzie, 1999). Knowledge of this distinction (and its importance) will assist teachers, in Hill's words, "to clearly articulate [their] beliefs about learning and how [their] assessment fit[s] with these" (1999, p.184). Teachers could then work collectively to ensure that school policies "build in assessment strategies known to provide the feedback crucial to student learning" (Hill, 1999, p.184). For this to happen, academic commentators have

suggested that ERO would be required to modify its existing approach toward reviewing and reporting on school assessment activities (Hill, 1999; McKenzie, 1999; Thrupp & Smith, 1999). It is scarcely surprising, however, that ERO officials view the matter differently (Aitken, 1996; Aitken, 2000).

Brazier (1999), for example, reports on “the drive for more tests and school league tables [that] is sweeping the world” (p.21), one that he alleges is already seriously undermining the ability of school teachers “to look after the wider needs of the whole child” (p.22; see McTamney, 2000). The “omnipotent obsession” with testing

industry. He argues that, despite claims to the contrary, this industry is mostly dedicated to “rooting out creative, child-centred teaching and learning, in order to resuscitate many of the practices of The Good Old Days” (p.24). But this “Golden Age”, Brazier reminds us (p.24), was all too often Dickensian in nature; that is, it was based on a



Assessment and education

The research findings outlined above clearly reveal that school-based assessment has been, and remains, a highly contested domain. They also point to the reality that “when the [assessment] stakes are raised for purposes of comparison” (Hill, 1999, p.184) – as is predicted to happen if ever externally referenced, standardised national testing of pupils is introduced into New Zealand primary schools (Lee & Lee, 1998) – parental and societal expectations are such that the great majority of teachers will be practically unable, as Hill expressed it, “[to] resist the performance imperative” (1999, p.184).

Such a response is educationally counterproductive, some commentators suggest. Christopher

pupils, Brazier observes (p.22), has meant that primary teachers are less likely to instinctively place a high value on fostering children’s artistic and musical appreciation, and to assist them to develop “a moral and critical understanding of the world they live in” (p.22). In short, teachers ‘know’ that careful test preparation of their pupils and good test results (‘outputs’) must be accorded priority.

One of the consequences of introducing a concerted testing campaign (associated with the expressed desire to raise ‘achievement standards’ in schools) that has attracted a strong politically conservative following throughout much of the Western world, Brazier suggests, is that education has been turned simply into another type of

mediocre, Gradgrindian pedagogy (see McKenzie, 1999). We might ask, why, then, has the notion of a Golden Age (re)gained popular appeal in several quarters? Brazier offers two possible explanations: the “diminished status and demoralised state of the teaching profession worldwide” (p.24), and the inability of policymakers and educationists to “focus on the vast difference between what is taught in schools and what is learned” (p.33). He concludes:

Teachers, like children, have become subject to the productivist view of education which runs the risk of seeing teachers as just another ‘input’, not quite as costly as a classroom but much more costly than a textbook. (p.24)

The purposes of assessment

On the basis of research conducted into testing in New Zealand and overseas, what should New Zealand teachers think about specifically in their assessment work? We suggest, first of all, that they ought to be fully aware of (and develop an informed stance on) competing viewpoints about the purposes and misuses of assessment. As Ross St. George (1981) said, teachers must be aware of “differences over the social functions served by educational tests and over the aims of education” (p.3). Although many commentators accept that one of the main purposes is to ascertain whether or not certain kinds of learning have actually taken place, there is likely to be disagreement over any claim that *every activity* can, and should, be assessed either by teachers or by a system(s) of public testing (see Renwick & Gray, 1995; White, 1999). Andrew Davis (1995), for example, maintains that “rich knowledge” (that which is acquired by means of coming to understand the concepts associated with true beliefs and their justification) can be validly assessed only by teachers and not by nationwide testing.

Other commentators, however, have argued that teachers’ assessment of pupils is, by its (limited) nature, an inadequate accountability instrument (Winch & Gingell, 1996). What are essential, they claim, are public tests which are both reliable (Turnbull, 1981) and valid (Education Forum, 1998), if individual schools and teachers are ever to be held accountable (as they should be) for the public money expended on them (see, for example, St. George, 1981). But, as John White notes (correctly), such a suggestion does nothing to “bring the work of the teacher closer to that of the parent” (1999, p.210). All that public testing can do is to assess “only the thinnest skills and recall of information” in a valid way (White, 1999, p.210). The “spiritual, moral and cultural development of pupils” is, therefore, better assessed by parents

and teachers than by national assessment, White concludes (p.210).

Control over assessment

Teachers ought also to be asking questions about the extent to which, as White puts it, “the school curriculum should be under professional or political control” (p.211; see Jesson, 1999). He maintains that monitoring and recording pupils’ accomplishments, along with using diagnostic testing and other assessment instruments, “[properly] belong to the professional sphere” (p.211). White contends that “teachers have a legitimate expertise...[in deciding on] the aims and broad content of school education” (p.211), and that they should be free from “political control of [the] more specific procedures” in their professional work (p.211).

Kelvin Smythe, in writing about New Zealand primary schooling, has expressed a similar view. He claims that “[assessment] in schools should largely be the domain of consenting teachers and parents” (1999, p.86), simply because school communities in the post-Picot era have been officially given responsibility for making major schooling decisions. These communities, Smythe suggests, ought to “decide on the nature and extent of the kind of [assessment and] evaluation that suits them” (p.86), rather than the Education Review Office. These sentiments were echoed by Emily Nelson (1998), who, in reflecting on the significant impact of New Right thinking on New Zealand schooling, declares that too many external controls have been applied to schools over the last ten years. She laments the restrictions placed on “teachers’ professional autonomy” (p.6), and “the marginalisation of teachers in the process of education policy development at a national level” (p.5), both of which she attributes to the mistaken belief by government officials that teachers are nothing other than “a self-interested group whose ideas cannot be trusted” (p.5). Nelson concludes that a pressing need exists for

teachers to “reclaim the[ir] profession by remaining informed, and by speaking out whenever and wherever [the] occasion permits” (p.6).

Authentic assessment

Although the debate over who will ultimately control school-based assessment looks set to continue, many educationists are insistent that, in any event, assessment must be ‘authentic’. Elizabeth McCabe (1998) explains that authentic assessment “acknowledge[es] the personal, social, cultural and political contexts of assessment, and relat[es] them to what happens to people in assessment situations” (p.41). But she is quick to note that authentic assessment ideals “must be tempered with a touch of reality” (p.41), because of the current preoccupation with measurement and outputs. Authentic assessment can occur, McCabe observes, when teachers have “rich and sometimes complex kinds of assessment data” available (p.42), which will reveal (besides other things) students’ understandings of a task, recognise differences in the cultural values and backgrounds of students and their assessors, and acknowledge the importance of assessment validity over reliability. Nevertheless, the stranglehold still exercised by conservative assessment practices cannot be ignored easily, as McCabe remarks:

Some parents still [seek] the comparability and comfort of formal, but often by nature less in-depth, scores from procedures like international, national, or school-based tests. (p.48)

Evidently, these parents had not been persuaded as to the wisdom of the Stewart Committee’s (1989) recommendation that in any and all assessment designed for better learning, “emphasis should be given to identifying and reporting educational progress and growth, rather than to comparisons of individuals or schools” (The Stewart Report, Department of Education, 1989, p.26). According to William Turnbull (1981)³, some people all

too readily subscribe to “the micrometer fallacy” (p.1), whereby parents and others ascribe an undeserved precision and an infallibility to test results, as well as adhere to “the whole person fallacy” (p.1). The latter refers to “the tendency to read into achievement test scores much more than they really tell” (p.1). Turnbull concludes by observing that “those who would ascribe to tests given at school age some magic that enables them to divine genetic intelligence or ability to learn should forget it” (p.3).

National testing in primary schools

Surprisingly perhaps, the above cautions (and others) did not stop the Education Forum (1998) from declaring its support for (re)introducing national testing into New Zealand primary schools. Ignoring the available historical and other literature on New Zealand primary school assessment (see Lee, 1991), the Forum declared that comparisons between schools will serve as “a motivating factor” for schools and teachers to “adjust their behaviour towards enabling the children to do well in the tests” (p.40). The explanation given in support of this behavioural psychology is that “[all] parents wish to see their children at a good school” (p.40). Accordingly, it was suggested that once teachers overcome their “initial resistan[ce] to tests” (p.36), they will come to appreciate that “children benefit

from the tests” (p.36) and that the test results will not be seen as “a measurement of the teachers themselves” (p.37).

Such sentiments are not shared by Lester Flockton, the co-director of the National Education Monitoring Project, however. After emphasising that standardised tests ought not to be seen as comprehensive, effective diagnostic instruments, Flockton urged the Ministry of Education to devise nationally referenced procedures for school-based assessment (“Teachers Need,” 2000; see Flockton, 1999). Nonetheless, this recommendation did not stop him from warning teachers that national primary school testing, as favoured by the Shipley government, could resurface at some point. If this was to happen, we suggest that primary teachers may have to remind those politicians and others supportive of national primary testing about Nick Smith’s comments, as a former Minister of Education, on such testing early last year to a Rotorua Secondary Principals’ Association meeting:

Schools are not widget factories where you can easily measure the product. We need to be cautious of creating a bureaucratic monster which has teachers spending more time assessing than they do teaching.

(Laxon & Young, 1999, p.A3)

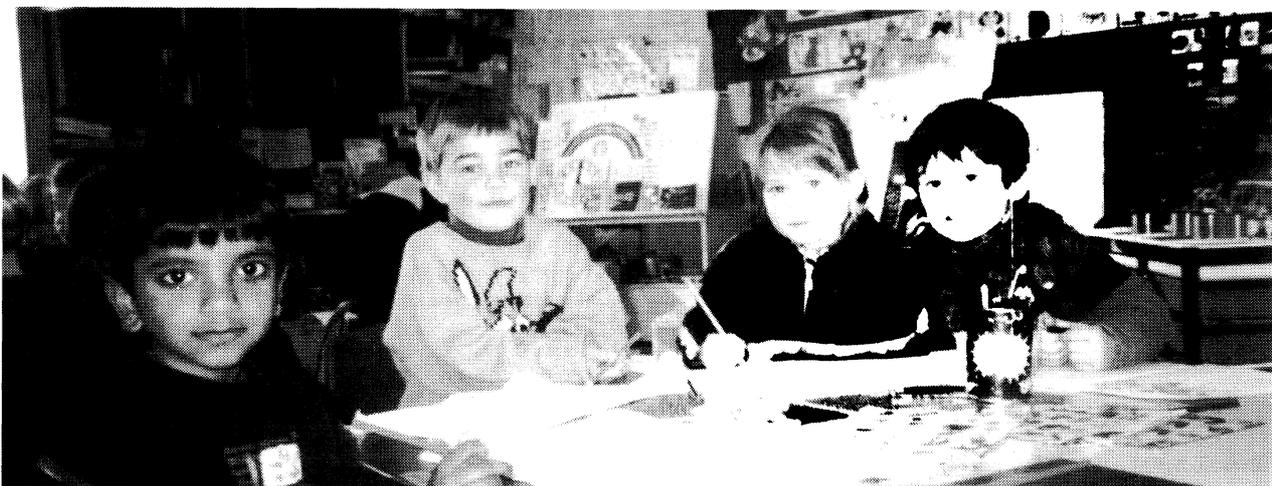
Readers may remember that the New Zealand Educational Institute (NZEI) fully endorsed the then

Minister’s view. But, erring on the side of caution, it decided to embark on a campaign to critique the Shipley government’s primary school assessment policies and proposals. Part of this campaign involved informing parents about the wide range of instruments that already resided in a primary teacher’s “assessment tool box” (NZEI, 1998). We suggest that teachers refamiliarise themselves with the NZEI’s assessment literature in case politicians change their stance on assessment. Doing this will also enable teachers to challenge ministerial assertions (and those of neo-liberal policymakers and commentators) that education policy documents are not “political” (see Smith, 1999, p.1), when their content is “written by the Ministry of Education” (Smith, 1999, p.1).

A lesson from the past

It is timely to recall John Ewing’s (1972) pertinent observation, recorded nearly three decades ago, that when the Commission on Education (The Currie Report, Department of Education, 1962) expressed support for ‘checkpoint’ achievement tests to be administered to primary pupils at the end of Standard 1, Standard 4, and Form 2 (formerly Standard 6), the Commissioners, too, believed that this testing would successfully identify pupils’ strengths and weaknesses. Ewing summed up the consequences of the Commissioners’ recommendation as follows:

Although the Commission was



at pains to explain 'that it had not in mind the mechanical application of examination barriers of the old standard type', many teachers felt that the checkpoint tests could regress into regular external examinations and, if administered nationally, could have a narrowing effect on the curriculum. (p.27)

Primary teachers, parents and politicians would do well to heed Ewing's warning.



Notes

1. NZJES, 34(1), 1999 is a special issue, edited by Martin Thrupp, titled "A Decade of Reform in New Zealand Education: Where to Now?" This issue contains several articles whose authors critique the education reforms implemented in the decade 1989-1999.
2. The Ministry of Education (1994) defined formative assessment as "a range of formal and informal assessment procedures (for example, the monitoring of children's writing development, anecdotal records, and observations) undertaken by teachers in the classroom as an integral part of the normal teaching and learning process in order to modify and enhance learning and understanding" (p.48). By comparison, summative assessment related to "a final assessment decision, for example: the assessment made prior to assigning a mark, grade, or descriptive comment to a student's performance for the mid-year or end-of-year report; the assessment made at the end of a course or section of a course; [and] the grades assigned to students' performance in external examinations" (p.50).
3. At the time of writing William W. Turnbull was President of the Educational Testing Service (ETS) based in Princeton, New Jersey, USA. The EFTS (founded in 1947) had become the world's largest private, non-profit organisation devoted specifically to educational measurement and testing. See Turnbull (1981), p.3.



References

Aitken, J. (1996, March 8). Achievement tracking: The role of the Education Review Office. Paper presented at the Principals' Centre Seminar, Auckland.

Aitken, J. (2000, April 12). Probability or proof-inference or information. Paper presented at the Teachers' Refresher Course Committee Seminar, Dunedin.

Brazier, C. (1999, August). Testing, testing, 123. *The New Internationalist*, 315, 21-25.

Broadfoot, P.M. (1996). *Education,*

assessment and society. Buckingham: Open University Press.

Butterworth, G., & Butterworth, S. (1998). *Reforming education: The New Zealand experience, 1984-1996.* Palmerston North: Dunmore Press.

Cooney, M. (1999, October). Against primary examinations. *Principals Today*, p.2.

Davis, A. (1995). Criterion-referenced assessment and the development of knowledge and understanding. *Journal of Philosophy of Education*, 29(1), 3-21.

Department of Education (1962). *Report of the Commission on Education in New Zealand.* Wellington: Author.

Department of Education (1989). *Assessment for better learning: A public discussion document.* Wellington: Author.

Education Forum (1998). *Policy directions for assessment at the primary school level: A submission on the government Green Paper, Assessment for success in primary schools.* Auckland: Author.

Ewing, J. (1972). The Currie Report: Ten years later. *Education*, 21(6), pp.1-50.

Flockton, L. (1999). Connecting theory and practice in national monitoring assessment. *Teachers and Curriculum*, 3, 23-30.

Hill, D. (1997). What do we teach? Teaching philosophy of education. *Waikato Journal of Education*, 3, 5-19.

Hill, M. (1999). Assessment in self-managing schools: Primary teachers balancing learning and accountability demands in the 1990s. *New Zealand Journal of Educational Studies*, 34(1), 176-185.

Jesson, J. (1999). Battling against the odds: The changing nature of the teacher unions. *New Zealand Journal of Educational Studies*, 34(1), 133-143.

Laxon, A., & Young, A. (1999, March 25). Government scraps school test plan – and Labour has ERO in its sights. *The New Zealand Herald*, p.A3.

Lee, H. (1991). The credentialled society: A history of New Zealand public school examinations, 1871-1990. Unpublished doctoral thesis, University of Otago, Dunedin.

Lee, H., & Lee, G. (1998, December). Raising 'national standards': Some inconsistencies in the Green Paper on Assessment for Success in Primary Schools. Paper presented at the annual conference of the New Zealand Association for Research in Education, Dunedin.

Lee, G., & Lee, H. (1999a). Book review: Butterworth, G., & Butterworth, S. (1998). *Reforming education: The New Zealand experience, 1984-1996.* Palmerston North: Dunmore Press. *Teachers and Curriculum*, 3, 89-92.

Lee, G., & Lee, H. (1999b, August 27). Education policy dynamics ignored. *New Zealand Education Review*, p.8.

Lee, G., & Lee, H. (2000, February). A response to Pat Lynch. *Principals Today*, p.2.

McCabe, E. (1998). Authentic assessment: Its application to mathematics education. *Teachers and Curriculum*, 2, 41-49.

McKenzie, D., Lee, H., & Lee, G. (1996). *Scholars or dollars? Selected historical case studies of opportunity costs in New Zealand education.* Palmerston North: Dunmore Press.

McKenzie, D. (1997). The cult of efficiency and miseducation: Issues of assessment in New Zealand schools. In M. Olssen & K. Morris Matthews (Eds.), *Education policy in New Zealand: The 1980s and beyond* (pp.47-64). Palmerston North: Dunmore Press.

McKenzie, D. (1999). The clouded trail: Ten years of public education post-Picot. *New Zealand Journal of Educational Studies*, 34(1), 8-17.

McTamney, J. (2000). Have fellowship – will appraise performance. *Good Teacher*, 2, 13.

Ministry of Education. (1991). *The national curriculum of New Zealand: A discussion document.* Wellington: Author.

Ministry of Education. (1993). *The New Zealand curriculum framework.* Wellington: Author.

Ministry of Education. (1994). *Assessment: Policy to practice.* Wellington: Author.

Nelson, E. (1998). New right: Wrong for teachers. *Developmental Network Newsletter*, 3, 5-6.

New Zealand Educational Institute. (1998). *When you want answers about your child's achievement at primary school.* Wellington: Author.

Openshaw, R., Lee, G., & Lee, H. (1993). *Challenging the myths: Rethinking New Zealand's educational history.* Palmerston North: Dunmore Press.

Renwick, M., & Gray, A. (1995). *Implementing the New Zealand curriculum in primary schools.* Wellington: New Zealand Council for Educational Research.

Smith, N. (1999, October 20). *Media release: Parents have right to information says Smith.* Wellington: Author.

Smythe, K. (Ed.). (1999). *The classroom and beyond: A celebration of New Zealand primary teachers and teaching.* Hamilton: Developmental Publications.

St. George, R. (1981). Troubled times for educational testing. *Set: Research Information for Teachers*, No.2, Item 15.

Teachers need to remain vigilant: NEMP. (2000, June 12). *Eduvac: The Education Weekly*, p.1.

Thrupp, M., & Smith, R. (1999). A decade of ERO. *New Zealand Journal of Educational Studies*, 34(1), 186-198.

Turnbull, W. (1981). Assessment test scores in perspective. *Set: Research Information for Teachers*, No.1, Item 12.

White, J. (1999). Thinking about assessment. *Journal of Philosophy of Education*, 32(2), 201-211.

Winch, C., & Gingell, J. (1996). Educational assessment: reply to Andrew Davis. *Journal of Philosophy of Education*, 30(3), 378-388.

