J. CURRICULUM STUDIES, 2000, VOL. 32, NO. 5, 623–650

Assessment: the engine of systemic curricular reform?

MARY BARNES, DAVID CLARKE and
MAX STEPHENS

We sought to examine empirically the prevailing assumption that changing assessment can leverage curricular reform. This assumption has been significantly confirmed by our research for the case of mandated high-stakes assessment. Two studies were conducted in the two most populous Australian states, New South Wales and Victoria. In the final two years of secondary school in both states, courses of study and assessment arrangements are mandated for all schools, including the private sector, by the state's Board of Studies. Congruence between mandated assessment and school-wide instructional practice was found in two states whose high-stakes assessment embodied quite contrasting values.

The assessment agenda

Whatever metaphor is employed to characterize the function of assessment within or upon the curriculum, the centrality of assessment is universally acknowledged. This paper makes use of the findings from two related studies to establish and elaborate two fundamental observations regarding the function of assessment:

- attempts at curriculum reform are likely to be futile unless accompanied by matching assessment reform; and
- assessment can be the engine of curriculum reform, or the principal impediment to its implementation.

Mary Barnes, formerly Director of the Mathematics Learning Centre at the University of Sydney, Australia, is now a freelance mathematics education consultant and a doctoral student at the University of Melbourne. Her interests in mathematics education include gender equity, teaching of calculus, assessment, and applications of technology. Her current research is on students' experiences of collaborative learning.

David Clarke is an Associate Professor and Associate Dean of Education (Research and Graduate Study) at the Faculty of Education, University of Melbourne, Grattan St., Parkville, Victoria 3010, Australia (e-mail: d.clarke@edfac.unimelb.edu.au). His research interests have centred on assessment, learning in classrooms, and teacher professional development. His publications include three books on assessment and a chapter on assessment in the *International Handbook of Mathematics Education* (1996). He is currently directing an international study looking at learner practices in 8th-grade mathematics classes.

Max Stephens was responsible for curriculum and assessment in mathematics at the Victorian Board of Studies from 1991–1998. His career has included teaching in schools and universities, working at the Ministry of Education, with a special focus on mathematics education and assessment. He is currently a senior associate in the Department of Science and Mathematics Education at the University of Melbourne.

Journal of Curriculum Studies ISSN 0022–0272 print/ISSN 1366–5839 online
http://www.tandf.co.uk/journals

In the studies reported here, these points are established with reference to the mathematics curriculum and the associated assessment. It is our contention, however, that the messages arising from these studies apply to most curriculum domains and speak generally to any school system seeking to implement a reform agenda.

Reardon *et al.* (1994: 1) identified two assumptions which have gone largely unchallenged in the public discourse but are central to the contemporary debate on assessment: first, 'that changes in assessment policies can be used as a powerful lever for reforming schools'; and, secondly, 'that new "authentic" forms of assessment ... are inherently superior to traditional standardized, multiple-choice tests'. The second of these assumptions was addressed explicitly by Supovitz and Brennan (1997), in their study of the equitability of portfolio assessment relative to standardized tests. Of at least comparable importance is the first assumption: that changing assessment policy can leverage school reform. This assumption provides the focus for this paper.

It is widely acknowledged that plans for curriculum change often fail to do anything more than generate new policy documents or exchange one form of professional rhetoric for another without any substantive change in classroom practice. It remains a challenge to identify substantive curriculum change in the form of changed instructional practice. Another widely accepted item of folklore is that mandated assessment drives instruction — that is, that the goals of instruction, and even the tasks employed, are derived from the anticipated assessment. The question as to whether instructional practice mimics assessment warrants immediate attention. In particular, does the form taken by 'high-stakes' assessment mandated by an authority external to the school significantly affect the patterns of classroom instruction? This issue is of contemporary importance, because assessment is being employed in many countries in the belief that it will be a catalyst for curriculum change.

The association between assessment and instruction has been assumed in much of the rhetoric of the education community (and the mathematics education community in particular). Where the form and content of assessment have been seen to exert a strong influence over what is taught, this is usually seen as undesirable and gives rise to prejudicial metaphors such as 'the assessment tail wagging the curriculum dog'. We argue that the performances privileged by assessment should be precisely those performances that constitute the goal of the curriculum. Where this is the case, it is entirely appropriate that there be a close link between assessment and instruction.

Irrespective of whether the link between assessment and instruction is seen to be desirable or not, the assumption of its existence has provided the motivation for the substantial deployment of personnel and resources in many countries in attempts to redefine curricula through changes to high-stakes assessment (Ministry of Education 1986, Burkhardt and Resnick 1991, Organization for Economic Cooperation and Development [OECD: PISA] 1998). It is, therefore, essential that the education community describe and analyse the link between assessment and instruction. We cannot simply assume such a causal connection. We must identify the

mechanisms by which assessment is translated into classroom practice. We need to ask: Is assessment an engine of curricular change or simply one manifestation of a larger change in the curriculum, the educational climate, or the wider culture?

Haertel (1999: 666) has asserted that 'The power of high-stakes testing to influence classroom instruction has been oversold'. It is worth noting, in passing, that the important debate on the role of assessment as either catalyst or engine for reform is not assisted by the exclusive identification of assessment with testing. In the discussion that follows, we present evidence of the pervasive influence of externally-mandated high-stakes assessment on both classroom assessment and instructional practice throughout the secondary school systems of the two most populous Australian states.

It is our contention that assessment should be recognized, not as a neutral element in the curriculum, but as a powerful mechanism for the social construction of competence. The imperative is to realize and exploit the significant role that assessment plays in this process. Investment in quality assessment offers governments and school authorities a powerful, cost-efficient means to model exemplary practice, while meeting the evaluative obligations of public accountability.

Assessment and alignment

One way of bringing these ideas together is to explore the concept of alignment as the desired convergence between a system's expectations, as expressed in its course documents and recommendations for teaching and learning, and what is actually mandated for assessment. Our basic premise is that all elements of a school system should work together to give consistent messages to teachers, parents, students, and the wider community about what is being valued. If a system's expectations are not well aligned with the means by which its outcomes are assessed, confusing messages are sent to these various parties about what is truly valued and what is expected of them.

The implementation of any reform agenda in education will be severely hampered by any lack of credibility in either the content of the proposed curriculum, or the means by which its outcomes will be assessed. The California Assessment Programme (California State Department of Education 1989) was an explicit commitment to aligning assessment practices with the concepts presented in the state's curriculum documents (California State Department of Education 1985, 1987). Public credibility of a system's curriculum and assessment arrangements is likely to be enhanced when these two elements are aligned.

However, mere alignment between expectations for curriculum and teaching and what is assessed will not necessarily reduce controversy and challenge. This was clearly demonstrated in the public reaction to both the California Mathematics Framework and the associated California Assessment Programme. More recently, public reaction to the new California Mathematics Standards (see, for example, Wu 1998) further indicates that controversy may indeed follow when people disagree with a system's

expectations or with the particular assessments by which the realization of those expectations is documented. Similarly, despite the great care taken to align curriculum and assessment objectives, the early years of the Victoria (Australia) Certificate of Education also attracted public controversy (Brown and Ball 1993). Alignment is, thus, no guarantee that what is valued by the school system will be universally respected by the general public, but it does point to the obligation on systems to avoid creating dissonance and confusion by having curriculum expectations and assessment out of alignment.

Webb (1997: 2) stressed the importance of a careful alignment between expectations and assessments as a means of avoiding 'overassessing a few outcomes at the expense of ignoring others'. While we would agree that benefits accrue from the alignment of expectations and assessment, the curriculum resides in the realization of these expectations in practice. This leads us to the key point that *mandated assessment mediates between the expectations of the system and their embodiment in classroom practices.* This accords an agency to assessment as a powerful lever for shaping school practice. Our research has shown that, where expectations and assessments are well aligned, there is a strong causative influence on classroom practices, and where this alignment is lacking or weak, assessment practices (particularly mandated assessment) appear to exert a far more coercive influence on classroom practice than the curriculum documents that articulate the system's expectations.

Whether or not assessments capture fully a system's expectations, the practices of teachers and students derive from an anticipation of what will be assessed and the form the assessments will take. In practice, this can lead to teaching based on a restricted set of goals, which misrepresent the richer expectations framed in curriculum documents. This is particularly true where the assessment is externally mandated in some way. Such assessments exert an influence on a school's curriculum and teaching, by the de-emphasizing of elements that are not endorsed by the assessment. Equally, in cases where the forms of mandated assessment do match more closely the expectations of a system's curriculum documentation, this conjunction acts as a powerful force for system-wide coherence between expectations, classroom practice and assessment.

Ideally, assessment tasks should reflect the range of performance specified in curriculum documents, syllabuses, or courses of study. Similarly, the rigour of the assessment tasks should match the expectations of the system. If, for example, the course documents expect students of statistics to work with and to interpret real data, but in assessments students are given only decontextualized data in order to meet the constraints of a particular timed test, then there is a degree of misalignment in the system. Webb (1997) suggests that, even if formal or large-scale assessments are unable for practical or technical reasons to cover the full range or rigour of expected performance, alignment can still be achieved if teachers incorporate into their classrooms alternative assessment practices not included in those formal or large-scale assessments. The two studies reported in this paper cause us to ask why teachers would do other than reproduce externally-sanctioned forms of assessment. The findings of our

research show that teachers either refrain entirely from undertaking types of assessments which are not part of the established regime or do so only at the fringes, if time and effort allow. Further, the inclusion of particular performance types in external mandated assessment supports an emergent consensus as to the nature of these valued performances. Where performance types lack this form of institutionalized endorsement, their enactment is consequently varied, not just in occurrence but in character.

In summary, alignment refers to how well all the policy elements of a system work together to guide both instruction and what is assessed. The determination of the degree of alignment can be carried out in a narrow sense through an analysis of the system's existing documentation of curriculum expectations and assessment arrangements. Both studies discussed in this paper have used this technique, although they moved beyond document analysis in two ways: First, to examine the messages that a school system sends to teachers through the means by which success is measured in public terms for teachers and students, and by which teachers are held accountable. Such mandated assessment arrangements shape teachers' beliefs about what is important, and bear more directly on the implemented curriculum than documentation alone. Secondly, and more importantly, both studies employed a combination of actual classroom documents, teacher questionnaires, and interviews to examine what teachers actually do in the classroom, that is, their instructional and assessment practices as distinct from their espoused beliefs.

Contexts of assessment and curriculum reform in mathematics

In recent years, the goals of mathematics education in most countries have been broadened to include such activities as problem-solving, investigation, mathematical modelling, and the communication of mathematical ideas (National Council of Teachers of Mathematics 1989, Australian Education Council 1991, Schoenfeld 1994, Burton 1996). Contemporary mathematics curricula represent efforts to translate this expanded conception of mathematical activity into practice. Mathematics educators in many countries are currently working to develop assessment schemes capable of modelling, monitoring, and informing these new curricula (Clarke 1996).

A consistency can be seen in the trends in mathematics assessment in communities as geographically dispersed as Australia (Victorian Board of Studies 1995), the Netherlands (Van den Heuvel-Panhuizen 1996), the Pacific region (Pacific Resources for Education and Learning 1997), Portugal (Leal and Abrantes 1993), Sweden (National Agency for Education 1995), the UK (Close *et al.* 1992), and the USA (National Council of Teachers of Mathematics 1995). The common elements of these assessment initiatives include the use of open-ended tasks, the use of contextualized settings for all tasks, the use of technology in instruction and its presence in assessment, and the expansion of the means of assessment beyond time-restricted examinations. This consistency derives from a new conception of the mathematics curriculum and the consequential demands for assessment

which is sensitive to the new standards in mathematics. These various national trends have been drawn together in significant international documents (e.g. OECD: PISA 1998) which recommend a broader framework for assessing mathematics than that found in traditional tests.[1]

In general, there is an international commonality of purpose in mathematics curriculum and assessment reform. As a consequence, each community can look to developments elsewhere with the confident anticipation that the new techniques can be adapted successfully to local conditions and priorities. In our opinion, this commonality of purpose extends across the school curriculum and, as a consequence, the implications of research into the instructional consequences and correlates of assessment initiatives in mathematics warrant the attention of the entire educational community. Strong suggestions of such causal associations between assessment and instruction have recently emerged in the Australian context.

Curriculum and assessment in Australia

In Australia, major responsibility for the provision of education, including advice on curriculum and assessment, resides with the states. Unlike the USA, each state's Department of Education is the direct employer of teachers in government schools and is responsible for funding their schools' operations. There is also a well-established independent sector that includes many schools affiliated with religious denominations.

Australians are apt to focus on the differences in educational provisions between the states, but to any outsider the similarities are probably greater than the differences. While the details of practices vary from state to state across Australia, all states have in place high-stakes externally mandated assessments taken by all students in the final year of secondary schooling. These assessments are used both to signify successful completion of secondary (high) school and also as the principal measure for deciding entry to university or college. This study focuses on the two largest states, New South Wales (NSW) and Victoria, which have a combined population of over 11 million out of the total Australian population of 18.6 million. (In each of these states about 66% of students attend government schools.) In both states, although the administration of government-administered schools is the responsibility of a state Department of Education, curriculum and assessment procedures are prescribed by a separate government-appointed Board of Studies. These Boards prescribe courses of study and assessment arrangements for all schools, including the private sector, in the final 2 years of secondary school. Entry to university or college is based almost entirely on the results of these assessments. Final grades are determined from a combination of school-based assessments and end-of-year examinations, with the latter playing a major role in determining the overall result. The end-of-year examinations in mathematics typically consist of extended-response questions based on a statewide syllabus. Students are expected to show substantial working in answering these questions. Both states, however, also make some use of multiple-choice items.

In Australia, mathematics is studied as an integrated course and is not divided into separate courses such as algebra and geometry as in the USA. Commercially-developed textbooks are written to meet the curriculum requirements of each state. However, any comparison between high school mathematics textbooks used in the two states would reveal substantial similarity of content to any outside observer.

Different courses are provided for students according to their mathematical background and intended courses of study at university. There is a high degree of similarity in the content of courses offered in the two states for students who intend to enrol in a university course for which mathematics is a required component. These courses all include substantial components of calculus, algebra, coordinate geometry and trigonometry. As in Years 7–10, these components continue to be integrated within a broad course description rather than taken as separate subjects.

In both NSW and Victoria, almost all students in the 11th grade take one or more mathematics subjects. However, in the 12th grade, about 70% of students in Victoria take some course in mathematics compared with over 90% in NSW where there is a broader range of mathematical offerings, some of which are not generally recognized for university entrance.

The international reform agenda in mathematics, referred to above, is most evident in the structure of the Victorian senior secondary assessment. Thus, a major difference between NSW and Victoria is the requirement in Victoria that school-based assessments in mathematics in the 11th and 12th grades include significant components of extended problem-solving and investigative work. These requirements apply to all mathematics courses, including the most challenging. By contrast, although schools in NSW are allowed to choose the form of school-based assessments, there is no requirement (and little encouragement) to use forms of assessment other than pen-and-paper tests.

Some specific features of mathematics courses in Victoria

For the purposes of this paper, the most important features in Victoria are the mathematics curriculum and associated assessment procedures used in the final 2 years of secondary school for the Victorian Certificate of Education (VCE).[2] The VCE acknowledges the very different types of performance from which 'mathematical competence' is constituted and employs a multi-component assessment instrument which is intended to capture the major features of that competence through the use of very different instrument types. An underlying principle of VCE mathematics is that all students will engage in the following mathematical activities:

- *Problem-solving and modelling*: the creative application of mathematical knowledge and skills to solve problems in unfamiliar situations, including real-life situations;
- *Skills practice and standard applications*: the study of aspects of the existing body of mathematical knowledge through learning and

practising mathematical algorithms, routines and techniques, and using them to find solutions to standard problems; and

- *Projects*: extended, independent investigations involving the use of mathematics.

These three learning activities are incorporated into all courses for VCE mathematics in Years 11 and 12 as formal *work requirements*. These work requirements are intended to promote key aspects of mathematical behaviour and to guide the work of teachers and students. The three work requirements are directly linked to the ways in which mathematical performance is assessed. They are intended to be used in an integrated way to develop understanding of concepts, communication skills, and a capacity to justify mathematical claims. Work requirements, therefore, provide a framework for planning, teaching and assessing in each course. At least 20% of class-time must be devoted to each work requirement and satisfactory completion of all work requirements is necessary for students to receive credit for courses taken in any semester.

In Year 11, schools are responsible for carrying out assessment tasks based on these work requirements. In Year 12, all mathematics courses have three Common Assessment Tasks. One is a centrally-set school-assessed task undertaken over a period of at least 2 weeks in the middle of the school year using a mixture of in-class and out-of-class time. The other two tasks are end-of-year examinations conducted under formal test conditions. All three Common Assessment Tasks contribute equally to students' overall scores for the course. School-assessed tasks take the form of an investigative project or a problem-solving task for which students have to prepare written reports according to guidelines. Reports are then graded by teachers following prescribed criteria. Sampling procedures are used to ensure quality control of grading and comparability of assessments across schools.

The study of VCE mathematics, thus, embodies a three-way link between the objectives of the course, work requirements and the range of performance assessed. Work requirements ensure that course objectives are translated into time spent on teaching and learning. Assessment tasks are in turn closely linked to work requirements.

In policy documents provided for Years 7–12 in Victoria, there has been a clear focus on having students become familiar with a range of problem-solving techniques, justifying their solutions, and writing short reports of mathematical investigations. These emphases are not simply the result of changes at the top. Greater emphasis on problem-solving in Years 7–10, for example, has been consistently advocated by all states and territories in their advice to schools, and was strongly endorsed in *A National Statement on Mathematics for Australian Schools* (AEC 1991). In Victoria, however, the VCE seems to have endorsed, or some would say required, a change of practice that these other policy documents merely encouraged.[3] It was the assumption of the Victorian study that we will be reporting here that this expanded range of assessment tasks in Years 11 and 12 had a direct and profound impact on the teaching of mathematics in the preceding years of secondary school.

Some specific features of mathematics courses in NSW

The most significant differences between NSW and Victoria lie in the curriculum and assessment arrangements for the final 2 years of secondary school. The comparable qualification to the VCE in NSW is the Higher School Certificate (HSC), which is intended by authorities in NSW to be seen as 'a credential with significant standing nationally and internationally' (Stanley 1998). A high importance is placed on maintaining this standing.

Syllabus documents for the HSC focus exclusively on content rather than process and make no suggestions about teaching or student work requirements. Students' grades are determined from a combination of school assessments and externally-set end-of-year examinations, with equal weight given to each. The guidelines for school assessments make it clear that for the 'high-status' courses (i.e. those recognized for university entrance), schools *may* if they wish use forms of assessment other than timed written tests, but there is no requirement or recommendation to do so. However, to ensure that comparisons of assessments by different schools are fair, performance on the examination is used as a statistical moderator to scale school assessments. This use of the examination for moderation purposes has the effect of placing additional emphasis on its importance.

Formal arrangements for the teaching of mathematics in the junior secondary years (Years 7–10) in NSW are very similar to those in Victoria. The main difference is that there are three distinct courses for Years 9 and 10, designed for students of different levels of achievement. These are distinguished by depth of treatment of largely common topics. As a result, from Year 9 onwards (earlier in some schools) all NSW schools practice some form of 'setting' in mathematics, placing students in different classes according to their perceived ability.

In addition to the 'high-stakes' external assessment at the end of Year 12, NSW has had for many years an additional statewide examination in mathematics, English and science in Year 10. At the time of the study, this examination was known as the School Certificate Reference Test and used only for the purpose of statistical moderation of school-awarded grades. In mathematics, this examination consists mainly of multiple-choice questions, although a smaller number of free-response questions appear towards the end of the paper, but most of these require the answer only and no working need be shown. Thus, the emphasis is on getting 'right answers', rather than communicating mathematical thinking — product rather than process.

Results from the School Certificate Reference Test have little practical importance for most students, although in the past many students left school at the end of Year 10 and this examination provided their main school-leaving credential. Possibly, as a result of this historical significance, the School Certificate Reference Test is still regarded as important by many schools, and much effort is expended on preparing students for it. No comparable assessment occurs in Victoria.

It appears that, in contrast to the situation in Victoria, teachers in NSW receive contradictory messages about what the system expects of them in

mathematics. On the one hand, a 'Statement of Principles' has been incorporated into all recent curriculum documents which discusses, among other issues, the nature of mathematics learning, emphasizing that students learn mathematics best through interaction with other people, through investigation, and through the use of language to express mathematical ideas. The syllabuses for Years 7 and 8 are well-aligned with this statement, emphasizing problem-solving, investigative approaches, and communication.[4] On the other hand, there is no *requirement* at any level to incorporate specific investigative, problem-solving, modelling, or communication tasks into school assessments. There is a clear implication in the various curriculum documents that assessment solely by means of examination is perfectly acceptable. This is seen as appropriate because it is widely believed that the only way to maintain standards is to ensure that the work assessed is the result of the student's unaided effort and that this can only be achieved under examination conditions.

Documenting the influence of mandated 'high-stakes' assessment

Two related studies were carried out to document the influence of the mandated 'high-stakes' assessment associated with the HSC in NSW and the VCE in Victoria. The Victorian study (Stephens *et al.* 1994, Clarke and Stephens 1996) hypothesized a relationship between mandated assessment and instruction, described as 'The ripple effect'. The intended image is one of a pebble or stone dropped into a pond at one point, and the resulting ripples spreading across the entire surface of the water. Here, the stone represents the introduction of the new VCE assessment procedures, and the ripples suggest the effect on teaching throughout the secondary school. A replication of the Victorian study was carried out in NSW (Barnes *et al.* 1995, 1996), driven by the belief that the relationship between mandated assessment and instruction could be better understood if informed by comparative data obtained in a state with a very different mix of curriculum advice and mandated assessments. In combination, the findings of these two studies confirm the significant role played by mandated assessment in determining the shape of instructional practice and assessment in the years of school preceding that assessment.

Sample selection

Schools in both states were selected by a process of theoretical sampling to include rural and metropolitan schools, government and non-government schools, and a variety of social demographic characteristics (including ethnicity and language). The variation in schools with respect to these characteristics was encompassed effectively in a sample of 11 schools in Victoria and 12 in NSW. Only schools providing instruction for Years 7–12 were included in the sample.

The three-stage design

Each study was carried out in three stages, based on three different sources of data: documents, questionnaires, and interviews. The results from each stage were used to inform the design of the instruments to be used in the next stage.

Stage one
The first stage of each study involved the analysis of documents related to the mathematics curriculum, and in particular to the teaching, assessment, and reporting practices in each school. The senior teacher of mathematics in each school, or his or her delegate, was asked to select documents which would help to establish a profile of present instructional, assessment, and reporting practices at the Years 7–10 level in the school, and which would indicate the extent to which these practices may have been influenced by mandated assessment at senior secondary level. Informants were asked to use their judgement to choose documents that would be most likely to provide relevant information. The intention was to identify the occurrence (if any) in such documents of terms, tasks, and structures derived from mandated high-stakes assessment. The inclusion of teachers' worksheets, teacher-designed assessment tasks, and other classroom material provided important corroborative detail supplementing the self-report data obtained by questionnaire and interview.

Stage two
In the second stage, all mathematics teachers in each of the selected schools were invited to complete questionnaires. The resulting sample included teachers with a broad variety of experience, including some who had been teaching for only a few years and others with many years of experience. The sample included some teachers who were not at the time teaching Year 12 mathematics, as well as many who were. The questionnaires sought to investigate current teaching and assessment practices at the junior secondary level (Years 7–10). In particular, the questionnaires invited teachers to comment on the importance of problem-solving and investigations in their teaching; the range of different forms of assessment they employed; and the values they attached to specified features of teaching and learning associated with problem-solving and investigative work in mathematics.

Stage three
The final stage took the form of interviews with some of the teachers who had completed questionnaires, undertaken with the aim of exploring in more detail their perceptions of the influence of external assessment on their teaching, assessment, and reporting practices at junior secondary level, as well as a number of additional questions arising from the responses to the questionnaires. The interviewer sought to ascertain the meanings and values which teachers attached to various teaching and assessment practices, and the extent to which they thought these practices might be helpful or otherwise to their students.

The 'ripple effect': assessment as change agent in Victoria

The Victorian study focused on:

- *Impact on school policy*. To what extent does school policy for the teaching and assessment of mathematics in Years 7–10 replicate the structural elements of the new VCE?
- *Adoption of VCE-derived practices*. To what extent are teaching and assessment practices inherent in VCE mathematics evident in the teaching of mathematics in Years 7–10? To what extent do teachers attribute classroom practice in Years 7–10 to those practices inherent in the VCE? Is there widespread use of a new vocabulary, demonstrably derived from the VCE by which mathematical activity at Years 7–10 is described?
- *Shared beliefs and common values underpinning changed practice*. Is there evidence of congruency of meaning and purpose in the way assessment and instruction is conceived in the VCE and in Years 7–10?

The following were the criteria by which a specific practice was taken to constitute evidence of the hypothesized 'ripple effect':

- A specific term introduced through VCE documentation was replicated in school documents, questionnaire responses or interview statements in relation to instruction or assessment at Years 7–10;
- The form in which a particular instructional or assessment practice was embodied in Years 7–10 closely resembled the form in which that practice had been introduced through the VCE;
- Statements were made by teachers on questionnaires or in interviews which explicitly linked these practices to those advocated within the VCE; in particular, statements indicating that particular assessment and instructional practices in Years 7–10 were derived or adapted from corresponding practices in Years 11 and 12.

The document analysis

The results of the document analysis can be reported as indicative of change in four key areas: policy; curriculum and teaching; assessment; and reporting. In their policy documents for Years 7–10, all 11 participating schools had adopted the VCE practice of using the three work requirement categories of extended mathematical projects, problem-solving and skills-practice to describe their mathematics courses. Within the course documents produced by the mathematics departments in all participating schools, problem-solving tasks and investigative projects were elaborated for teachers and students in very specific terms, strongly resembling those used in VCE curriculum documents.

VCE-related assessment practices were evident in:

- the tasks which comprised a school assessment package;

- what students were expected to do in those tasks; and
- how the tasks themselves were assessed.

In all 11 schools, course documents outlined the clear expectation that the assessment of problem-solving and investigative projects, together with the assessment of standard mathematical skills and applications, should contribute to students' overall assessment in mathematics.

There was clear evidence in the Victorian sample that the assessment of problem-solving tasks and investigations had been incorporated into the formal structures of assessment and reporting. As important, for this research, was evidence from an analysis of teacher-generated worksheets and assessment material that teachers of Year 7 and 8 mathematics (as well as Years 9 and 10) included problem-solving tasks and investigations in their repertoire of instructional activities.

The questionnaire

The total Victorian sample comprized 50 teachers. The number of teachers per site ranged from two to six, with a median of five.

Problem solving activities and investigations

The proportion of class-time which teachers reported giving to small-scale investigations and to problem-solving activities at Years 7–10 is recorded in table 1. For the majority of teachers, the reported level of use was in close correspondence to the proportion required in Years 11 and 12, although no such prescription applied in Years 7–10.

Respondents were asked to indicate the degree of importance they attached to various classroom activities related to problem-solving and investigative work in Years 7–10. A four-point scale was used (*highly important*, *of some importance, beneficial but not essential*, *of little importance*) to allow teachers to indicate their purposes in using problem-solving activities and investigations. Consistently high levels of approval (with 50% of respondents indicating a 'highly important' rating) were given by all respondents to those aspects which were strongly endorsed by VCE curriculum advice and assessment practice. Two items, 'Students posing their own problems' and 'The regular completion of student mathematical journals', which were not derived from VCE practice, received the smallest

Table 1. Proportion of class time reported by teachers as being given to small-scale investigations and problem-solving activities in Years 7–10.

Percentage of class-time reported as given to the activity type	Small-scale investigations *n* of respondents (% of sample)	Problem-solving activities *n* of respondents (% of sample)
0–< 10	7 (16)	6 (14)
10–< 20	21 (49)	20 (45)
20–< 30	12 (28)	15 (34)
30–< 40	3 (7)	3 (7)

Table 2. Percentages of respondents rating each item as 'Highly important' in each state.

	Item	NSW	Victoria
	The application of mathematics to real-world contexts	67	66
	The use of different mathematical skills in combination	64	72
	Presenting problems spanning a range of content areas in mathematics	56	45
*	Teaching problem-solving skills	51	69
*	Presenting problems which require a range of problem-solving techniques	49	67
	Using problems specific to the topic being taught	49	45
*	Students developing investigative skills	40	74
*	Using problem-solving to develop mathematical skills	28	54
*	Giving students substantial written comment on their problem-solving attempts	21	36
	Students posing their own problems	19	22
*	Students developing report writing skills	10	55
*	Students undertaking an extended mathematical activity (over 1 week)	5	32
*	Students undertaking open-ended mathematical activities	5	30
*	The regular completion of student mathematical journals	3	18

Table 3. Types of formal assessment strategies employed by current VCE teachers and others (percentages of column totals).

Formal strategies	VCE teachers	Non-VCE teachers
Tests or exams		
• In use	100	100
• Attributed to VCE	34.5	41.7
Projects reports		
• In use	93.8	92.9
• Attributed to VCE	92.6	81.8
Students problem-solving reports		
• In use	93.5	90.9
• Attributed to VCE	85.7	75.0
Student self-assessment		
• In use	48.4	16.7
• Attributed to VCE	13.6	20.0

proportion of 'highly important' responses, 22% and 18%, respectively. These results are presented in table 2, where they can be compared with responses for the NSW sample. Items for which there is a discrepancy of 15% or more between the NSW and the Victorian responses are highlighted with an asterisk.

Assessment and reporting practices

Respondents were asked to indicate which formal assessment strategies they employed and whether the use of these strategies by themselves or others was a consequence of VCE influence (see table 3). The formal strategies presented were: *Tests or exams*; *Project reports*; *Student problem-solving reports; Student self-assessment.* Consistently high levels of use and attribution to the VCE were reported for *Project reports* and *Student problem-solving*, regardless of the teachers' VCE experience.

Taken together with the document analysis, the findings from the second phase of this study substantiated the hypothesized 'ripple effect' in Years 7–10, revealing general and consistent instructional practices in significant correspondence with (and attributed to) changed assessment practices in Years 11 and 12. The uneven occurrence (and widely different interpretations and implementations) of non-VCE practices such as student self-assessment can be interpreted as significant confirmation of the research hypothesis.

The interviews

Shared meanings

Teacher interviews revealed a substantial commonality of meaning for those terms and practices that were integral to VCE mathematics. Problem-solving, for instance, was consistently associated with open-ended mathematical activity rather than merely 'textbook problems'. None of the teachers saw the completion of routine problems from a textbook as constituting engagement in problem-solving activity.

It might be argued that the inclusion of problem-solving in the curriculum of Years 7–10 can hardly be attributed to the influence of the VCE, when it has received such wide endorsement from the wider educational community. Clearly, the inclusion of problem-solving, in itself, is not evidence of the impact of the VCE. Our argument is based rather on the particular emphasis given to problem-solving in the participating schools, and the consistently shared beliefs among teachers about how problem-solving activities are to be implemented. It is because these particular beliefs and practices are so consistent with and so closely reflect the practices and approaches to problem-solving endorsed by VCE mathematics that the 'ripple effect' argument is strongly supported by the data. In the interview data, teachers consistently employed similar language to describe classroom problem-solving activity, its structure and the means by which it was assessed. In particular, the terminology employed was predominantly the terminology of the VCE problem-solving assessment guidelines. In the following interview excerpts, some of these key terms have been italicized:

> [In commenting on problem-solving reports] [I] expect to see the *aim* and *information* given or any *assumptions* . . . and then *solution* and *justification* . . . why the *solution* should be accepted . . . how they've reached the answers.
>
> Then we would look for the application of an appropriate scheme, . . . and then some discussion of the *reasonableness of answers*. . . . And along with that, providing students with the *criteria for assessment* before the task.

Summary of the Victorian study

Taken together, the three phases of this study substantiate the hypothesized ripple effect in Years 7–10, due to changed assessment practices in

Years 11 and 12. In addition to their confirmatory value, the second and third phases of the study have provided a more detailed characterization of this ripple effect and of the mechanisms by which it occurs. In particular, specific terminology and practices associated with the VCE could be found consistently in mathematics instruction at every level of the secondary school. Other practices, such as student self-assessment and the use of student journals, while endorsed by the informed mathematics education community, did not have the explicit sanction of inclusion in high-stakes assessment. The level of use of such practices was uneven (and generally low) across the sample and inconsistent in form of implementation.[5] It is a specific finding of this study that teachers are reluctant to embrace new assessment and instructional practices unless these have the endorsement of inclusion in high-stakes assessment.

'Maintaining standards': constraints on change in NSW

The purpose of the NSW study was to investigate the influence of mandatory high-stakes assessment on teaching practice in circumstances where the conjunction of assessment in Year 12 (the HSC) and curriculum advice in Years 7–10 was rather different from that in Victoria. The research questions for this study can be summarized as follows:

- *Impact on school policy*: Could any specific terms or practices be identified in documentation for the HSC which were replicated in school instruction and assessment policy in the junior years?
- *Influence of mandated assessment on school practice*: To what extent is instructional practice in the junior years influenced by the need to prepare students for the HSC? To what extent does the School Certificate Reference Test (Year 10) influence teaching and assessment in the junior secondary years, and how can this influence be explained?
- *Teacher beliefs and values*: To what extent do teachers value practices such as extended problem-solving, investigations and the communication of mathematical ideas, which are endorsed by the mathematics education community generally and in some NSW curriculum documents, but which are not explicitly incorporated in the HSC assessment?

Since the main goal of the NSW study was to generate findings that could usefully be compared with those from the earlier Victorian study, the research instruments were adapted from those used in the Victorian study. The wording of questions was altered where it was felt necessary to ensure that the instruments were appropriate for the NSW context, and some additional items were added which addressed issues specific to that state, notably the influence of the Reference Test in Year 10.

The document analysis

The documents supplied by schools showed that they all relied heavily on timed written tests and examinations for assessment at all levels. Six schools sent copies of policy statements on assessment, indicating that, at junior secondary level, assessment for reporting purposes was based from 80–100% on tests and examinations (the percentage varied from school to school and from one grade level to another). One school's policy statement made the reasons for this explicit:

> Pen and paper tests remain the most common formal method of assessment at [this school], and this simply reflects tradition and the dictates of the external assessment in Years 10 and 12. [Later in the same document] Examinations help prepare Years 9–12 for the end-of-course external exams which they must face.

With few exceptions, written examinations in Years 9 and 10 (and in some cases Years 7 and 8) closely followed the format of the School Certificate Reference Test. It was clear that most schools accorded high priority to familiarizing students with the form of the Reference Test and the style of questions likely to be asked in it. In their reports to parents, none of the schools mentioned any alternative assessment procedures such as reports of investigations, student journals, or student self-assessment. In only one case was explicit reference made to problem-solving. It cannot, of course, be inferred from this that such assessment procedures were not used in these schools, only that it was not thought important to include information about them in reports to parents.

In summary, the emphasis on tests and examinations as the major form of assessment, and the formats used for school reporting, indicated the influence of external assessment. (This influence was made explicit in one school's policy statement.) In particular, the modelling of examinations on the School Certificate Reference Test demonstrated that this test exerted a very strong influence on school assessment in Years 8–10.

The questionnaire

Responses were received from 59 teachers in 12 schools — 32 males and 27 females. Apart from a small country school from which only one response was received, the number of respondents per site ranged from three to 10, with a median of five.

Problem solving and investigations

Teachers were asked what they perceived as the benefits and costs of using problem-solving and investigations with students in Years 7–10. The responses placed more emphasis on the difficulties and costs involved in implementing these activities than on the benefits. The cost most frequently mentioned (by 76% of respondents) was the time required for these activities. Two typical responses were:

> Whilst we have overloaded syllabuses which require the acquisition of numeric and algebraic skills, there is little spare time for some students to experience much problem-solving.
>
> The necessity to complete all work before exams . . . means that little time can be spent on open-ended investigations.

Assessment and reporting

Respondents indicated that the formal assessment strategies employed at the Years 7–10 level strongly emphasized tests and examinations. All respondents reported using tests and examinations with frequency varying from every 2 weeks to every 6 weeks or so. For 50% of respondents, this was the only formal strategy used.

The teachers were also asked about their schools' practices in determining the school assessment mark or grade for the HSC. The replies indicated extremely heavy reliance on tests and examinations; eight of the 12 schools reported that tests and examinations were the *only* form of assessment used for the higher-level (calculus-based) courses taken in Years 11 and 12. None indicated the use of investigative project-work or problem-solving carried out over a reasonable period of time with opportunities for revision. One stated explicitly that they did not use any form of take-home assessment 'as these cannot be guaranteed to be the student's own work'.

Factors informing teaching practice

The questionnaire sought to explore the opportunities respondents might have had to become aware of new ideas and trends in mathematics education. The responses indicated that by such means as attending courses, reading journals, and talking with other teachers, most respondents had had many such opportunities.

A closer study of the data, however, suggests that this conclusion needs to be qualified slightly. Over 40% of those who had taken part in professional development activities specifically reported attending courses, conferences or meetings dealing with teaching or assessing the HSC. The most frequently mentioned was an annual one-day 'Meet the examiners' conference held early in each academic year to discuss the previous year's HSC papers. One teacher reported that the most significant influence on her teaching had been her experience as an HSC marker. Thus, many of the significant influences cited were focused on external assessment.

In summary, the questionnaire results provided evidence of teachers' constant awareness of external assessment, and in particular of the need to complete the syllabus in good time and give students plenty of practice in the type of questions likely to be asked in external examinations, both the School Certificate Reference Test and the HSC. These results provided further confirmation that external assessment does indeed have a powerful influence on teaching, assessment and reporting at the junior secondary level.

The interviews

As in the Victorian study, the purpose of the interviews was to explore explanations for some of the questionnaire responses in more detail, and in particular to probe the values and beliefs underlying the teaching and assessment practices which the teachers referred to in the questionnaires.

Perceived influence of the HSC

In general, teachers felt that the HSC tended to make them more content-oriented in their teaching in the junior years. They saw it as important to build a sound foundation for work in Years 11 and 12 by emphasizing the basics and perhaps by adopting a more theoretical approach than they might otherwise have chosen.

Apart from content issues, teachers generally felt that the need to prepare for the HSC made them more examination-oriented. Some indicated an awareness of the importance of the HSC right from the beginning of Year 7. Especially in some non-government schools, the culture of the school placed a strong emphasis on training students to perform well in written examinations. As one teacher expressed it:

> The school community, the parents and the students, are so focused on the Higher School Certificate that everything is focused on it, it's a cultural thing ... right from Year 7. ... Our approach seems to be that this is very real and earnest, that it's important, and the parents are very concerned that the students are moving forward in what they see as traditional maths, so we tend to stay with traditional approaches.

It should not be thought that a traditional approach implies rote learning. Many of these schools emphasized the development of a deep understanding of mathematics. As a teacher from one of these schools said: 'We don't want them to do things by rote, we want them to do things by understanding, and we want them to understand at the highest level that they can'.

Perceived influence of the School Certificate Reference Test

Four teachers claimed that the Reference Test had hardly any influence at all. These teachers came from two non-government schools and all had already indicated that they saw the HSC as the major goal for their students. For others, the Reference Test exerted a strong influence on both teaching and assessment in Years 9 and 10. The main effect on teaching was the pressure to complete the syllabus in time for the test, and to prepare students for the format of the test. Typical responses included the following:

> I suppose to a certain extent that you tend to 'coach' .... I think, to be fair to the students, you have to do that, because I'm quite sure that ... the majority of schools, will be doing the same thing. So we don't want to put them at a disadvantage.
>
> We make sure that our students always have, in exams, multiple-choice questions to practice, to work on, so that they are accustomed to them and they're not thrown by them and don't treat them as multiple guess!

What if there were no external assessment?

Interviewees were asked: 'Thinking more about process and teaching style than about content, would you teach differently in Years 7–10 if you did not have to bear in mind the requirements of external examinations?' Approximately a quarter of the teachers said that it would make little difference to what they did. The rest expressed, in a variety of ways, the idea that removing the constraints of external examinations would free them to be more experimental. It would give them the chance to try more exciting things, to undertake more extension or enrichment work, more 'hands-on' practical activities, more investigative activities or group work, or to relate mathematics more to the students' interests. For example:

> You would probably tend to do more interest-type activities, more outdoor-type activities, that you wouldn't worry about in an exam, and the children might enjoy it more, find it more interesting.

Problem-solving and investigative activities

Among NSW teachers, unlike their Victorian counterparts, there were diverse interpretations of the nature of problem-solving and investigations. There was little agreement about what these terms mean, how important the activities are, how to incorporate them into teaching programmes, and how to assess them. In many schools, problem-solving was assessed either informally or not at all, thus sending a message to students that it was not as important as those activities that were assessed by means of conventional tests and examinations.

There was also little consensus among the NSW teachers about the purpose of written reports in mathematics, and great variability in their use. In many schools, grading and reporting procedures appeared to attach considerably less value to this type of activity than to more conventional forms of assessment. These findings can be linked with the lack of models of problem-solving, investigations or any activities requiring the writing of reports in mandated external assessment procedures, and with the lack of any guidance in an official syllabus or in other documents on how to implement and assess these activities.

Alternative forms of assessment

Respondents were asked about other alternative assessment techniques — specifically self-assessment by students and the use of student journals or learning logs. Only one teacher reported using either of these regularly, and many were doubtful about their potential benefits.

The responses to these questions revealed the high value placed on what is most likely to help students to prepare for the type of questions likely to be encountered in external assessment, especially the HSC. One reason that alternative assessment procedures are not widely used is that they are not perceived by most teachers as directly relevant to that aim. As one teacher said, 'I don't see a student journal as being very relevant to the style of HSC exams'.

Summary of the NSW study

The HSC clearly emerged as the goal that teachers and students have in their sights from quite early in secondary school. Its major influence appears to be the high value which it causes schools to place on tests and examinations as a form of preparation for the high-stakes HSC examination. Few mathematics teachers are prepared to experiment with alternative approaches to teaching or alternative forms of assessment — mainly because of pressure of time, but also because they do not see a clear pay-off for their students in terms of improved grades in the HSC. In this way, the present system in NSW tends to inhibit educational innovation.

This study also suggests that such assessment as the School Certificate Reference Test may have an unintended impact on the mathematics curriculum, because of its mandated, external character. Most teachers acknowledged its lack of significance, while simultaneously reporting that they stressed its importance to their students. The importance of the Reference Test was variously associated with its implications for the reputation of the school and with the maintenance of 'standards'.

Values accorded to different mathematical activities: a comparison between NSW and Victoria

Questionnaire respondents in both states were asked to indicate the degree of importance they attached to a number of different aspects of teaching and learning associated with problem-solving and investigative work in mathematics. A four-point scale (*highly important*; *of some importance*; *beneficial but not essential*; *of little importance*) was used. Table 2 lists the items, showing the percentage in each state who scored each item as 'highly important'.

The greatest difference between the responses of the NSW and Victorian teachers to these statements is seen in the importance they give to students developing report-writing skills. Fifty-five per cent of Victorian teachers regarded it as highly important as compared with only 10% of NSW teachers. NSW teachers also gave very much less support than Victorian teachers to students developing investigative skills, the item supported most strongly by Victorian teachers. These two statements reflect aspects of doing mathematics which are emphasized in VCE assessment procedures, but which are of little importance in preparing students for the HSC examinations or the School Certificate Reference Test. The same applies to students undertaking extended and open-ended mathematical activities. Such activities are endorsed by the way in which problem-solving and investigations are built into and assessed by the VCE. Teachers in NSW do not attach comparable importance to these activities, possibly because they cannot be tested by means of traditional examinations.

The use of student mathematical journals, extended and open-ended mathematical activities, and the development of report-writing skills received very limited endorsement by teachers in NSW. There was also a

low level of endorsement for students posing their own problems and teachers providing substantial written comment on problem-solving attempts. All of these represent aspects of reflective mathematical activity. A possible interpretation of these results is that extended, reflective and investigative activities are not valued by NSW teachers because they cannot be tested effectively by means of standard tests and examinations.

Mathematics teachers in NSW appear to hold many shared values in relation to their teaching which differ greatly from those of their colleagues in Victoria. We emphasize that the difference between the two states is *not* the presence or absence of standard tests and examinations across the two states; as has been noted, traditional examinations constitute a major component of assessment in VCE mathematics. Likewise, the differences between the two states cannot be attributed to differences of culture or community, since they are similar in their level of affluence and cultural diversity and have comparable demographic profiles. The principal difference between Victoria and NSW is that the VCE mandates problem-solving and investigative tasks as significant components within its suite of high-stakes assessments and the HSC in NSW does not.

Assessment and systemic curricular change

We sought to investigate whether the implementation of the curriculum was facilitated, or hindered by, or independent of, mandated high-stakes assessment, i.e. to examine empirically the prevailing assumption that changing assessment can leverage curricular reform. This assumption has been supported by our studies for the case of mandated high-stakes assessment. In a school sample that included a wide range of demographic contrasts, private and public schools, rural and metropolitan schools, those with significant non-English speaking backgrounds and those without, the degree and nature of the impact on instruction of mandated high-stakes assessment was effectively uniform across all school types. That congruence between mandated assessment and school-wide instructional practice should be found independently in two Australian states whose high-stakes assessment embodied quite contrasting values constitutes the most compelling outcome of these studies.

The effects of mandated assessment at the end of secondary school are quite different in the two states. In both, assessment of students at the end of secondary school consists of a mix of school-based assessment and end-of-year examinations. Both components contribute significantly to students' final results. In both states, teachers use sample questions developed by examining boards and by teachers' groups, as well as examination papers from previous years, to interpret the syllabus, to guide their teaching, and to prepare students. The examinations require students to use a mixture of extended and short answers to questions. Experienced graders are employed to score student responses and the grading schemes are subsequently reported to teachers. There is very little difference in the style and content of examination papers taken by students in Victoria and NSW at the end of Year 12.

Where teachers are left free to determine the forms of school-based assessment, they may be free, in principle, to utilize investigations and other extended tasks, but in practice in NSW the forms of their school-based assessments seem to mirror closely the forms of the end-of-year examinations. By contrast, in Victoria, where the forms of school-based assessments are themselves mandated, these assessments exercise a significant leverage on teaching and forms of assessment, especially where the assessment tasks are set by the examining body with extensive guidance for teachers in applying criterion-based assessment. By mandating the forms of school-based assessment in the senior secondary years and by providing extensive advice to teachers on how these assessments are to be implemented, a powerful message is conveyed to teachers that these forms of assessment are to be taken seriously in the preceding years of secondary school. Where such alternative forms are simply encouraged or permitted, there is little inducement for teachers to depart from a prudent decision to focus school-based assessments along similar lines to those required in traditional end-of year examinations.

In the absence of mandated alternative forms of assessment in the earlier years of schooling, there will be little perturbation to existing teaching beliefs about what constitutes good assessment and what is needed to prepare students for subsequent work in the later years of secondary school. Alternative assessment, indeed, may be permitted with those groups of students who are perceived as unlikely to go on to the most demanding courses in later years. Such freeing-up of forms of assessment for these lower-achieving groups will be seen to be of no direct concern for those who are destined for more demanding and rigorous courses.

The emphasis in this paper is on a 'curricular alignment', by which assessment matches curricular goals and instructional practice and, by this correspondence, serves as a model for both. The importance of such alignment, as demonstrated in this study, should not be seen as support for an assessment-based accountability system. Systems that reward or punish teachers on the basis of the assessment of their students' performance appeal to a philosophical framework (and a model of teacher professionalism) entirely different from the rationale of congruence between curricular policy, instruction, and assessment.

The principle of alignment needs to be interpreted as the degree of congruence between the expectations of a school system for students' performance and the various elements of the system's assessment arrangements. A corollary of this principle might be that expectations and assessment should be aligned in a clear, efficient, and economical manner. It appears from our research that teachers in Victoria recognize that complex mathematical performances of the kind valued within the VCE require careful groundwork. It is essential, at this point, to acknowledge the importance of providing teachers with documentary support and appropriate pre-service and in-service training in the implementation of both the new curriculum and the corresponding new assessment techniques.

The lack of uniformity with which Victorian teachers implemented or chose not to implement various reform initiatives that were not seen by

them as part of the VCE is mirrored by a similar lack of uniformity with which teachers in NSW chose to implement or not implement practices that they did not see as embodied in the forms of assessment which comprised the HSC. There is clear evidence that some schools in NSW currently employ a range of 'reform' practices in Years 7–10. In Victoria, prior to the VCE, such innovative practices were employed unevenly in Years 7–10 and simply disappeared in Years 11 and 12, where teachers knew what needed to be done in order to prepare students for end-of-year examinations. For similar reasons, they are virtually absent from Year 11 and 12 teaching and assessment in NSW. The linguistic and methodological forms of assessment that are embodied in Years 7–10 in Victorian schools 'mimic' their officially mandated correlates in the VCE. It is this congruence between mandated forms of assessment in the VCE and what takes place in the preceding years of schooling that undermines any argument that they could be shaped by general patterns of innovation.

The debate concerning the role of external high-stakes assessment and the desirability of alignment between assessment and instruction is not helped by the rhetorical use of exaggerated caricatures of performance assessment such as that employed by Haertel (1999).[6] We would, however, fully concur with Haertel's suggestion that 'The hoped-for benefits of assessing students using complex, integrative, hands-on tasks that blur the lines between testing and teaching can in fact be realized when teachers themselves make skillful use of sound performance assessments in their own classrooms' (p. 663). On the basis of this study, it is our claim that the VCE provided the impetus for teachers to change their assessment practice in just this way, and also provided a model, and a language for talking about the model, together with widespread professional development that assisted teachers in implementing such forms of assessment in their own classrooms.

In both NSW and Victoria, teachers saw it as imperative to build into their teaching programmes and into their procedures for assessment and reporting a range of practices that progressively initiated students into the types of performance required in Year 12 assessment. Mandated high-stakes assessment in the senior years of school fundamentally shapes teachers' perceptions about what is important to teach and to assess; and about what performances to encourage and how these might be assessed. Where new forms of assessment are mandated in Year 12, these act as a powerful mechanism for initiating change in the preceding years of school. An exclusive focus on examinations in Year 12, on the other hand, serves to filter out, or to severely limit, the use of alternative assessment practices which are not seen to contribute directly to building expertise in exam-taking. As a consequence, performances not easily displayed in examinations are omitted from classroom assessment and instruction. Curricular reform undertaken without corresponding assessment reform appears unlikely to succeed. As a result of the studies reported here, the assumption that changing assessment policy can leverage reform has been given an empirical foundation.

Acknowledgements

We wish to express our appreciation for the helpful comments received from Stephen Lerman and Carne Barnett on an earlier draft of this paper. The comments of three anonymous reviewers are also gratefully acknowledged.

Notes

1. In several countries, developments in assessment can be linked to specific national projects or initiatives. Some of these are illustrated in Clarke (1996) and Burton (1996). Other related issues are discussed in Leder (1992), McGaw (1993) and Stephens and Izard (1992). There is also a growing body of research into issues of equity surrounding assessment. Questions have been raised about the social and cultural biases in many assessment tasks, especially 'realistic' contextualized tasks (Darling-Hammond 1994, Cooper and Dunne 1998), and the validity of teachers' assessment decision-making (see Dunne 1998, Morgan 1998, Watson 1998). Reference to the cultural aspects of assessment raises an issue that has only been hinted at in this paper. It is important to note that enthusiasm for the new assessment approaches is not universal. For example, Leung (1995) argued that the traditional time-restricted examination continues to meet the needs of the school system in Hong Kong, as it has for over 1000 years. He suggested that the conventional examination is culturally consistent with the values of communities that share a Chinese philosophy. There are communities for which the imposition of Western-style assessment creates a fundamental cultural conflict. For example, in some traditional communities of indigenous Australians it is not socially acceptable for a person to ask a question to which he or she already knows the answer. Such a caveat renders any examination system and most Western instructional methods inadmissible. In such a culture, education proceeds by elaborate modelling by skilled members of the community and assessment consists of the learner progressively increasing their participation in the purposeful activities of the community as their expertise and success grows. In an environment in which poor navigation skills can have terminal consequences, performance assessment is inevitable, mandatory, and unforgiving.
2. Introduced in all schools in 1990, the VCE provides a common credential for all students in these last 2 years of school.
3. Educational systems are not static entities. All assessment schemes should be continually evaluated and refined. The VCE continues to be modified to optimize economies of implementation and the demands on student and teacher time. Similarly, additional strategies have been introduced to increase confidence in the authorship of submitted work. Such changes represent refinements of the original assessment model and in no way affect the conclusions drawn in this study.
4. Recent changes have seen the introduction of a new syllabus for Years 9 and 10 in NSW that places much more emphasis on investigative activities, and a change in the role of the School Certificate examination, which is no longer used solely as a statistical moderator. Students now receive a grade in mathematics based on their performance in the examination.
5. A more detailed report of the Victorian study can be found in Clarke and Stephens (1996).
6. In contrast with Haertel's (1999) characterization, high-stakes externally-mandated school-based assessment, as conducted within the VCE, was carried out over an extended period of at least 2 weeks, including periods of work at home, unrestricted access by students to a wide variety of resources (provided these were acknowledged), with the provision of teachers' feedback during task-completion. Tasks were designed to address the *taught curriculum* and students chose from a range of assessment tasks. The products of such assessment activities were scored by the student's own teacher, employing a

centrally-specified rubric. It is not our intention to advocate any of these specific task characteristics, but it is important to recognize the limitations of a critique, such as Haertel's, based on such a restrictive assessment model.

References

Australian Education Council (1991) *A National Statement on Mathematics for Australian Schools* (Melbourne, Australia: Curriculum Corporation).

Barnes, M., Clarke, D. J. and Stephens, W. M. (1995) Links between assessment and the teaching of mathematics in New South Wales secondary schools: preliminary report. In B. Atweh and S. Flavel (eds), *MERGA 18: Galtha. Proceedings of the 18th Annual Conference of the Mathematics Education Research Group of Australasia* (Darwin, Australia: Mathematics Education Research Group of Australasia), 57–65.

Barnes, M., Clarke, D. J. and Stephens, W. M. (1996) The impact of external assessment on teaching practice: constraints on change in the classroom. In P. Clarkson (ed.), *Technology in Mathematics Education: Proceedings of the 19th Annual Conference of the Mathematics Education Research Group of Australasia* (Melbourne, Australia: Mathematics Education Research Group of Australasia), 64–71.

Brown, T. and Ball, S. (1993) *A Report on the VCE Verification Process* (Melbourne, Australia: Victorian Curriculum and Assessment Board).

Burkhardt, H. and Resnick, L. B. (1991) *A Balanced Assessment Package* (Nottingham, England: Shell Centre for Mathematical Education).

Burton, L. (1996) Assessment of mathematics – what is the agenda? In M. Birenbaum and F. Dochy (eds), *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge* (Dordrecht, The Netherlands: Kluwer), 31–62.

California State Department of Education (1985) *Mathematics Framework for California Public Schools: Kindergarten through Grade Eight* (Sacramento, CA: California State Department of Education).

California State Department of Education (1987) *Mathematics Model Curriculum Guide: Kindergarten through Grade Eight* (Sacramento, CA: California State Department of Education).

California State Department of Education (1989) *A Question of Thinking: A First Look at Students' Performance on Open-ended Questions in Mathematics* (Sacramento, CA: California State Department of Education).

Clarke, D. J. (1996) Assessment. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick and C. Laborde (eds), *International Handbook of Mathematics Education* (Dordrecht, The Netherlands: Kluwer), 327–370.

Clarke, D. J. and Stephens, W. M. (1996) The ripple effect: the instructional impact of the systemic introduction of performance assessment in mathematics. In M. Birenbaum and F. Dochy (eds), *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge* (Dordrecht, The Netherlands: Kluwer), 63–92.

Close, G. S., Boaler, J., Burt, A., Coan, E. and Symons, K. (1992) *Materials to Support the Assessment of Ma1: Using and Applying Mathematics* (London: School Examinations and Assessment Council).

Cooper, B. and Dunne, M. (1998) Social class, gender, equity and National curriculum tests in mathematics. In P. Gates (ed.), *Proceedings of the 1st International Mathematics Education and Society Conference* (Nottingham, UK: Centre for the Study of Mathematics Education), 132–147.

Darling-Hammond, L. (1994) Performance-based assessment and educational equity. *Harvard Educational Review*, 64 (1), 5–30.

Dunne, M. (1998) Pupil entry for National Curriculum tests in mathematics: the public and private life of teacher assessments. In P. Gates (ed.), *Proceedings of the First International Mathematics Education and Society Conference* (Nottingham, UK: Centre for the Study of Mathematics Education), 148–157.

Haertel, E. H. (1999) Performance assessment and education reform. *Phi Delta Kappan*, 80 (9), 662–666.

LEAL, L. C. and ABRANTES, P. (1993) Assessment in an innovative curriculum project for mathematics in grades 7–9 in Portugal. In M. Niss (ed.), *Cases of Assessment in Mathematics Education* (Dordrecht, The Netherlands: Kluwer), 173–182.

LEDER, G. C. (ed.) (1992) *Assessment and Learning of Mathematics* (Melbourne, Australia: Australian Council for Educational Research).

LEUNG, F. (1995) Mathematics assessment in Hong Kong. Paper presented to the videolink symposium International Perspectives on Mathematics Assessment as part of the annual meeting of the American Educational Research Association, San Francisco.

MCGAW, B. (1993) School-based and external assessment: international, interstate and local issues. Paper presented at the forum School-based and external assessment: Uses and issues in post-compulsory years, conducted by the Incorporated Association of Registered Teachers of Victoria (IARTV), Melbourne, Australia.

MINISTRY OF EDUCATION (1986) *Future Directions of Postcompulsory Schooling* (Melbourne, Australia: Ministry of Education).

MORGAN, C. (1998) Assessment of mathematical behaviour: a social perspective. In P. Gates (ed.), *Proceedings of the 1st International Mathematics Education and Society Conference* (Nottingham, UK: Centre for the Study of Mathematics Education), 277–283.

NATIONAL AGENCY FOR EDUCATION (1995) *National Test in Mathematics Course A* (Stockholm, Sweden: Swedish National Agency for Education).

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS (1989) *Curriculum and Evaluation Standards for School Mathematics* (Reston, VA: National Council of Teachers of Mathematics).

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS (1995) *Assessment Standards for School Mathematics* (Reston, VA: National Council of Teachers of Mathematics).

ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT: PROGRAM FOR INTERNATIONAL STUDENT ASSESSMENT (OECD PISA) (1998) Framework for Assessing Mathematics Literacy. Draft document prepared by the Mathematics Functional Expert Group for the National Project Managers' Meeting, Arnhem, The Netherlands, 26–30 May.

PACIFIC RESOURCES FOR EDUCATION AND LEARNING (1997) Standards for Pacific education. *Pacific Education Updates*, 9(2), 16.

REARDON, S.F., SCOTT, K. and VERRE, J. (1994) Symposium: equity in educational assessment: editors' introduction. *Harvard Educational Review*, 64(1), 1–4.

SCHOENFELD, A. H. (1994) Learning to think mathematically: problem-solving, metacognition, and sense-making in mathematics. In D. Grouws (ed.), *Handbook of Research on Mathematics Teaching and Learning* (New York: Macmillan), 334–370.

STANLEY, G. (1998) Letter to the editor, *Sydney Morning Herald*, 8 June.

STEPHENS, W. M. and IZARD, J. (eds) (1992) Reshaping assessment practices: assessment in the mathematical sciences under challenge. *Proceedings from the 1st National Conference on Assessment in the Mathematical Sciences* (Melbourne, Australia: Australian Council for Educational Research).

STEPHENS, W. M., CLARKE, D. J. and PAVLOU, M. (1994) Policy to practice: high-stakes assessment as a catalyst for classroom change. In G. Bell, B. Wright, N. Leeson and J. Geake (eds), *Challenges in Mathematics Education: Constraints on Construction Proceedings of the 17th Annual Conference of the Mathematics Education Research Group of Australasia*, Vol. 2 (Lismore, Australia: Mathematics Education Research Group of Australasia), 571–580.

SUPOVITZ, J. A. and BRENNAN, R. T. (1997) Mirror, mirror on the wall, which is the fairest test of all? an examination of the equitability of portfolio assessment relative to standardized tests. *Harvard Educational Review*, 67 (3), 472–502.

VAN DEN HEUVEL-PANHUIZEN, M. (1996) *Assessment and Realistic Mathematics Education* (Utrecht, The Netherlands: CD-b Press).

VICTORIAN BOARD OF STUDIES (1995) *Mathematics Assessment Activities Using Number, Tools and Procedures* (Melbourne, Australia: Victorian Board of Studies).

WATSON, A. (1998) Potential sources of inequity in teachers' informal judgements about pupils' mathematics. In P. Gates (ed.), *Proceedings of the 1st International Mathematics Education and Society Conference* (Nottingham, UK: Centre for the Study of Mathematics Education), 337–344.

WEBB, N. L. (1997) Determining alignment of expectations and assessments in mathematics and science education. *NISE Brief*, 1(2), 1–8.

WU, H. (1998) Some observations on the 1997 battle of the two Standards in the California math war. An expanded version of a colloquium lecture at the California State University at Sacramento, February 12, 1998. http://mathematicallycorrect.com/hwu.htm. [Visited 21 January 2000.]