

HELEN M.G. WATT

ATTITUDES TO THE USE OF ALTERNATIVE ASSESSMENT
METHODS IN MATHEMATICS: A STUDY WITH SECONDARY
MATHEMATICS TEACHERS IN SYDNEY, AUSTRALIA

ABSTRACT. Given issues related to differences in learner characteristics, effective sampling across the content domain, and recent emphases on assessing meaningfully contextualised abilities and higher-order cognitive processes, the 'traditional' mathematics test arguably does not provide a valid measure of student ability. Consequently, there is a need to incorporate alternative methods of assessment that are able to effectively assess the range of students' mathematical abilities. The present study investigated methods of assessment used by 60 mathematics teachers from 11 secondary schools in metropolitan Sydney, as well as their attitudes to a range of alternative assessment methods, together with reasons why they would or would not implement these. Results showed that teachers were satisfied with traditional tests as valid measures of student ability, particularly for senior school years. Teachers generally did not favour implementing alternative assessment methods, although those with the least years' teaching experience reported more positive attitudes. A major concern raised by teachers about the use of alternative assessment methods related to their perceived subjectivity. Explanations for these findings are advanced for teachers who have varying lengths of teaching experience.

KEY WORDS: alternative assessment, attitudes, Sydney Australia, mathematics teachers, secondary school

1. REVIEW OF LITERATURE

The purpose of this study was to examine practising teachers' attitudes to alternative assessment methods in secondary school mathematics in Sydney, Australia. Explicit research questions relate to (1) teachers' use of alternative assessment methods in maths, (2) attitudes about using alternative assessment methods, and (3) perceived impediments to using alternative methods.

Assessment in mathematics has traditionally been *measurement-driven*, using assessment not only to rank students but also to keep accountability of the educational system (Broadfoot, 1996; Niss, 1993a, b). There have been debates particularly in the USA about the benefits and problems accompanying measurement-driven assessment (see Corbett and Wilson, 1991; Popham, 1987). Instructional practices of teacher-led recitation of mathematical procedures, where teachers demonstrate and students are expected to reproduce a broad range of facts and mathematical operations

in timed pencil-and-paper tests (e.g., Goodlad, 1984; Powell et al., 1985), have likely been responsible for the continuing association of this type of assessment with mathematics.

Since the influential Cockroft Report (Cockroft, 1982) in Britain, conceptions of mathematics learning have shifted significantly. That report provided a focus for developing ideas of mathematics educators around the world, including in Australia, with its arguments used to support curriculum changes in the UK and elsewhere, particularly in relation to assessment (Galbraith, 1995a). Similarly, in the USA, the National Council of Teachers of Mathematics standards (NCTM, 1989) endorsed 'recognition of mathematics as more than a collection of concepts and skills to be mastered; it includes methods of investigating and reasoning, means of communication, and notions of context' (p. 5). Disappointingly, an international study comparing teaching practices and assessment forms throughout the USA, England and Wales (Firestone et al., 2000), reported strikingly similar mathematical instruction across countries, where teachers explain procedures without targeting deep conceptual learning, and students replicate procedures using sets of small problems. Prior research suggests that assessment techniques are not being aligned with the new mathematical goals (e.g., Niss, 1993b; Schoen, 1989).

In parallel with overseas developments, curriculum reform and changing societal demands for a technologically savvy workforce (e.g., Mathematical Sciences Education Board, 1993), we have seen an emphasis on problem-solving skills in Australian secondary school mathematics curricula. In Australia, developments have been generated locally, being in sympathy with, but not derivative from initiatives overseas (Galbraith, 1995a). Assessment in mathematics has become a major contemporary focus amongst Australian educators, largely through the impetus of recent curricular initiatives at both national and state levels (see Stephens and Money, 1993 for a review). These developments have paralleled initiatives also occurring in the UK, the USA, the Netherlands, Portugal, South Africa and the Western Pacific Rim (Clarke, 1996). Australian issues therefore relate to the international scene more broadly where similar questions are being addressed, although it is important to note that such trends are not universal, with Hong Kong for example retaining traditional time-restricted examinations, these being culturally consistent with community values based on Chinese philosophy (see Leung, 1995; cited in Clarke, 1996). In the Australian context, given decreased demand for computational skills, syllabi now emphasise mathematical process as distinct from product. Rather than someone who is able to neatly replicate a learned procedure to a routine task in a familiar context, a successful mathematics student has been reconceptualised as one who is able to devise

problem-solving strategies, identify conceptual similarities in different situations, assess the relevance of different procedures to applied contexts, and work productively with others (Clarke, 1987). It has been argued that newer approaches such as performance-based assessment have strong potential to impact mathematics teaching (Rothman, 1995), through emphasising challenging material for all learners (Smith and O'Day, 1991). In this context, we might hope mathematics teachers would utilise newer forms of assessment.

School education in Australia is State run, which is the source of much of the diversity in practices between the different states. The State of New South Wales (NSW), in which the present study is located, is widely regarded as the most conservative in relation to curriculum and assessment change. The focus on Sydney teachers may therefore well be sampling the views of the most conservative teachers with respect to assessment. It will be of particular interest to explore these teachers' views, to identify the range of assessment practices they use, and to examine what encourages or constrains their practices. Such examination can contribute to better understanding the major deciding factors for teachers using alternative assessment methods.

Should the traditional mathematics test predominate in mathematics secondary school teaching in this study, there are several reasons why we should be concerned. First, differences in learner characteristics imply that over-reliance on one form of assessment disadvantages students who are able to display their knowledge, skills or abilities more effectively through other methods (e.g., Leder et al., 1999). To encompass a range of learning styles and goals, there needs to be a wide range of methods for gathering assessment information (Niss, 1999; Stephens, 1987). Second, effective sampling across the content domain of educational mathematics objectives may not be achieved using the one modality, with some instructional goals likely to be emphasised and others de-emphasised. Resultant content bias will advantage some students and disadvantage others, and not provide a rounded picture of students' mathematical abilities (Stephens, 1988). Assessment information on a non-representative subset of objectives cannot be extrapolated to the whole range of objectives, and so student performance on the traditional mathematics test cannot be used to infer more general mathematical ability (QBSSS, 1992; Stephens, 1988).

Related to these two issues is the concern that currently emphasised meaningfully contextualised mathematical abilities and higher-order cognitive processes are less effectively assessed via the traditional mathematics test than alternative means, such as portfolio assessment (Simon and Forgette-Giroux, 2000). While it is not necessarily true that written tests are

restricted to computation and routine skills, and are capable of assessing a wide range of mathematical capability if set appropriately, unfortunately many of these tests are not well-written, and the traditional mathematics test typically focuses on repetition of learned procedures using small sets of problems (e.g., Firestone et al., 2000).

Even though traditional mathematics tests effectively assess aspects of mathematics which can be tested in an unambiguous and straightforward way, through students' performance on routine skills and algorithms (Clarke and Lovitt, 1987; Grimison, 1992; Stephens, 1988), there is a need to explore alternative assessment methods to assess other instructional goals. To date, reliance on the traditional mathematics test has been justified on the grounds of maximising reliability and ensuring comparability, but this has often been at the expense of validity (Clarke, 1996; Lacey and Lawton, 1981). All States in Australia have made significant moves away from external examinations (Stephens, 1988), in order to supplement that information with assessments made at the school level. The intention here was to combine the independent and 'objective' estimate of a student's performance in the external examination, with broader evidence of abilities in areas of the course not necessarily reflected in the examination result (Karmelita, 1987). School-based assessment was intended to broaden the assessment base, and include tests and activities which assess some objectives more validly than the traditional mathematics test (Karmelita, 1987). Rather than implying *more* assessment, it implied the use of newer *forms* of performance-based assessment such as oral assessment, investigations, problem-solving, projects and assignments (see Karmelita, 1987; NSW Board of Studies, 1996).

Although much responsibility for assessment has been given over to the school in order to allow for the use of such alternative testing methods, this gesture seems to have been misinterpreted, with the result being a greater *frequency* of traditional mathematics tests given by the school. One suggested explanation is that in order to protect the high status accorded to mathematics in our society, mathematics teachers oppose the idea of using alternative assessment techniques (Clarke, 1987). This explanation may also relate to the extent to which mathematics is perceived as objectively assessed, with an implicit assumption that numerical ratings are more defensible than other forms of assessment, despite the fact that such scores or ratings may subsume elaborate judgements that remain invisible (Delandshere and Petrosky, 1998). Other possible explanations relate to teacher indifference, given the lack of any imperative to change their assessment practices; or to deeply embedded knowledges and beliefs about teaching mathematics (Firestone et al., 2000).

The most influential reason prior research suggests for teachers' resistance or indifference to alternative forms of assessment, is that although informally such assessment information shapes teachers' opinions of students' competence, it is regarded as highly subjective (e.g., Watson, 2000). However, teachers' rather definite opinions of their students formulated in this way tend to be quite accurate, and traditional assessment often does little more than legitimise and quantify the assessment made through alternative means in extended classroom contact (Clarke, 1987). One article (Hoge and Coladarci, 1989), which reviewed findings from 16 studies to assess the correspondence between teacher judgments and students' test scores, found there was a strong association between these, which is in contrast to the common concern that teacher judgments cannot be trusted. It appears therefore that there is little basis for this concern with unreliability. Where discrepancies between traditional and alternative assessment occur, they can frequently be attributed to atypical student performance, and teachers often adjust pass marks, test scores and grades to accord with their informal assessment of students (Clarke, 1987). It is the lack of structure in such informal assessments, and the fact that they are not systematically recorded, that likely results in them lacking the status accorded to a traditional test score (Clarke, 1987; Niss, 1993a). The measurement-driven assignment of numbers to test responses to measure the extent of individual achievement has dominated educational assessment through most of the twentieth century, with resulting scores used to make judgements about the quality of performances. Numerical ratings are of course also easier to rank students by, a preferred process for selection purposes to tertiary institutions and the workforce (Findlay, 1987), which may be another reason for maintaining the status quo.

The present study investigated examples of assessment practice that can be identified in a sample of Sydney secondary mathematics teachers, what encourages and sustains those practices and what can be done to encourage other teachers to broaden their own assessment practice. Although it was anticipated that the traditional test would be the main assessment method used, teachers' use of alternative assessment methods was canvassed for each year level, as it was expected teachers' behaviours and attitudes would differ according to which school years were being discussed. In the state of NSW Australia, mathematics syllabi are written for each of grouped years 7–8, 9–10 and 11–12. The years 7–8 syllabus is focused mainly on consolidation of learned material from primary school (years 3–6), and particularly emphasising problem solving. The syllabi for years 9–10 and 11–12 become increasingly crowded, allowing little time to explore new avenues of teaching and assessing, particularly in senior secondary years. The increasingly externally-driven assessment focus in senior years also

makes it possible that less use would be made of alternative assessment methods, with the NSW state-wide School Certificate occurring in year 10, and the Higher School Certificate over years 11 and 12, and likely consequent backwash effects on teachers. These considerations may imply that the use of alternative assessment methods would be more frequent with younger years. It was also anticipated that length of teaching experience would relate to use of and attitudes towards alternative assessment as a result of differences in initial teacher education, with the consideration that recently graduated teachers would be more likely to have been exposed to alternative assessment methods and therefore be more receptive to their use.

The alternative assessment methods targeted in this study included those suggested in the New South Wales mathematics syllabi as well as the curriculum literature in mathematics education, of (1) oral tasks where students give short answers, seminar presentations and debates; (2) practical tasks with students using instruments to apply or deduce mathematical principles; (3) teacher observation of students in structured or unstructured activities and evaluation of the quality of student task engagement; (4) student journals where students keep reflective accounts of their mathematics learning and processes of understanding, from which the quality of their task engagement and development may be explored by the assessor; (5) student self-assessment with students judging the quality of their own and their peers' mathematical understanding and progress; and (6) involving parents in the assessment process, asking them to observe, reflect on and evaluate their child's mathematical understanding and progress.

2. METHOD

2.1. *Participants*

The mathematics staff of 11 Sydney metropolitan schools were invited to participate in the study. Eight schools were government schools (one all-girls in the Western Suburbs; seven coeducational with two from the Blue Mountains, two in Sydney's south, two in Sydney's south-west, and one in the north-west of Sydney), and three were private independent schools (one all-boys and one all-girls from Sydney's upper-north Shore and located within five kilometres of each other, the other all-girls in Sydney's eastern suburbs). Although this was a convenience sample, with schools selected on the basis of their accessibility, all were located in Sydney, and represent a mix of government, private independent, coeducational and single-sex schools. The return rate of self-report surveys, which were delivered and collected in person, was 68% (a total of 60 of the 88 surveys).

3. MATERIALS

The survey developed for this study consisted of eight questions, of which three were quantitative (questions 1–3) and five were open-ended (questions 4–8). Question 1 asked teachers to indicate whether they had been teaching for 0–9, 10–19 or 20+ years. Question 2 asked teachers to rate how well they considered traditional mathematics tests (where students are required to reproduce learned mathematical procedures on small problem sets in pencil-and-paper tests) assess mathematical ability, on a 5-point scale ranging from 0 (not well at all) through 2 (adequately) to 4 (very well), separately for each of years 7–8, 9–10 and 11–12. Question 3 asked teachers to tick which out of a provided list of six alternative assessment methods they used (oral tasks, practical tasks, observation, student journals, student self-assessment and parental assessment). These listed alternative methods were based on suggested methods in the New South Wales mathematics syllabi as well as alternative assessment methods that have been suggested and described through the curriculum literature in mathematics education (e.g., Clarke, 1988; Galbraith, 1995b; Grimison, 1992; Kulm, 1994). The reader is referred to these accounts for more elaborate descriptions of these alternative assessment formats, which were overviewed in Section 1.

Question 4 asked teachers to explain in an open-ended fashion, reasons why they would or would not use these listed alternative assessment methods. Question 5 asked teachers to list the most common methods of assessment they used for each of years 7–8, 9–10 and 11–12. Question 6 was an open-ended question which asked whether the curriculum for each of years 7–8, 9–10 and 11–12 allowed for use of alternative assessment methods; following on from which open-ended question 7 asked how the curriculum could be adapted to allow for alternative methods of assessment. Finally, open-ended question 8 asked what other forms of assessment teachers thought could be used to assess students in mathematics.

4. ANALYSES

For quantitative responses to questions 1–3, descriptive statistics summarised data. Frequency counts summarised length of teaching experience in question 1, while question 2 ratings were summarised by group means for teachers with varying lengths of teaching experience (0–9, 10–19 and 20+ years), separately for each of years 7–8, 9–10, and 11–12. For question 3, proportions of teachers using alternative assessment methods were recorded, according to their length of teaching experience.

Open-ended responses to questions 4–8 were grouped according to emergent themes detailed in the following Results section. This recording of responses into themes was intended to give structure and coherence to open-ended data, allowing for greater ease of interpretation. First, coding summarised data into smaller units for subsequent analysis; and second, ‘pattern coding’ procedures for data analysis outlined by Miles and Huberman (1994) were followed. Pattern coding refers to grouping data summaries formed in the first phase of data coding into a smaller number of sets or themes. This two-stage process of initial coding followed by subsequent pattern coding is a commonly recommended and implemented method for qualitative analysis (e.g., Krippendorff, 1980; Miles and Huberman, 1994; Strauss and Corbin, 1990, 1998; Weber, 1990). Examples from surveys were selected for illustrative purposes as typifying certain themes, providing expression of participants’ voices (Geertz, 1993). Following derivation of emergent themes for questions 4–8, proportions of teachers within each category of teaching experience were recorded for each theme.

5. RESULTS

5.1. *Teacher satisfaction with traditional mathematics tests*

Teachers were relatively satisfied with traditional mathematics tests as a measure of students’ mathematical ability for years 11–12. Teachers appeared less satisfied with traditional tests for measuring mathematical ability with years 9–10, and less satisfied again with these tests’ capability to measure ability for years 7–8 (see Figure 1). For years 11–12 and 9–10, teachers with less teaching experience appeared less satisfied with traditional tests than teachers who had been teaching for longer, while level of satisfaction with the traditional test for years 7–8 appeared unrelated to length of teaching experience. In all cases, mean group ratings were above the scale midpoint, which was anchored at ‘adequately’.

5.2. *Reported use of alternative assessment methods*

Among least experienced teachers, the most common alternative assessment method used was observation (used by 71% of teachers), followed closely by oral and practical tasks (both 64%). Oral tasks were the most common alternative method employed by more experienced teachers (83% for 10–19 years experience, 77% for 20+ years), followed by observation and practical tasks (observation: 79% for 10–19 years, 64% for 20+ years, practical: 67% for 10–19 years, 64% for 20+ years). In general, oral and

TABLE I

Proportions of teachers using various alternative assessment methods according to length of teaching experience

	Years of teaching experience		
	0–9 years: <i>n</i> = 14 (%)	10–19 years: <i>n</i> = 24 (%)	20+ years: <i>n</i> = 22 (%)
Oral tasks	64	83	77
Practical tasks	64	67	64
Observation	71	79	64
Student journals	29	17	18
Student self-assessment	29	29	41
Parental assessment	7	13	9
Overall use of alternative assessment techniques	44	48	46

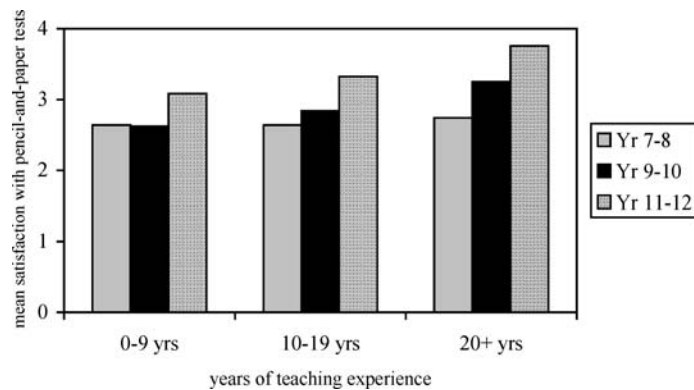


Figure 1. Teacher satisfaction with the written test as a measure of students' mathematical ability, according to length of teaching experience.

practical tasks as well as observation were employed by more teachers, with student journals, student self-assessment and parental assessment infrequently used (see Table I). Parental assessment was the least common method employed for all groups.

5.3. Teachers' reasons for and against using alternative assessment methods

Themes that emerged from analysis of open-ended responses to why teachers would or would not employ alternative assessment methods in

mathematics (question 4) fell into six major categories. Those which related to not employing alternative assessment methods were (1) insufficient time for implementation, (2) unstructured nature, (3) unsuitable, (4) unreliable/subjective, and (5) insufficient resources at hand to permit implementation. The final category, (6) suitable and beneficial, related to employment of alternative assessment methods.

Theme 1 (insufficient time for implementation) included comments such as 'time consuming', 'time factor' and 'can take up too much valuable time'. Theme 2 (unstructured nature) contained responses such as 'difficult to set up', 'keeping other students out of earshot (for oral tasks) is difficult' and other references to organisational problems. Theme 3 (unsuitable) consisted of comments such as 'can't see how to incorporate into maths', 'does not seem appropriate to maths', 'not suitable' and comments such as 'parents rely on teachers to assess, so of course should not be involved in the assessment process'. Theme 4 (unreliable/subjective) included comments such as 'not valid', 'biased', 'inequitable', and 'students are too hard on themselves' which implied invalidity. Theme 5 (insufficient resources) contained responses such as 'costly' and 'not easy to find example tasks, especially after year 7'. Theme 6 (suitable and beneficial) was derived because there were some positive responses related to the value of oral and practical assessment techniques. Proportions of teachers mentioning each theme were tabulated for each identified alternative assessment method, as well as across methods, according to length of teaching experience.

The most common reason teachers would not use alternative assessment methods across the set of alternative methods was that they regarded them as too subjective (18% of 0–9 years' teaching group, 14% for 10–19 years, 19% for 20+ years). Factors not perceived as impediments for these teachers were insufficient resources and alternative assessment methods being too unstructured, with low proportions of teachers citing these reasons. Lack of time was also rarely cited, with the exception of teachers having 20+ years' teaching experience, where 10% raised this as an issue. Unsuitability of alternative assessment methods to mathematics was also rarely mentioned, except in the case of teachers having 10–19 years' experience, 14% of whom identified this as a concern. Table II tabulates proportions of teachers, calculated within each group, who mentioned each of the five perceived impediments to implementing alternative assessment methods. This table also shows that 12% of the least experienced teachers described alternative assessment methods as suitable and useful in measuring students' mathematical ability.

Table III tabulates proportions of teachers from each group raising each of these reasons for and against using these alternative assessment methods, separately for each method. The most frequently raised concern was the

TABLE II
Teachers' attitudes to using alternative assessment techniques across the set of listed methods, according to length of teaching experience

	Years of teaching experience		
	0–9 years (%)	10–19 years (%)	20+ years (%)
Insufficient time	1	2	10
Too unstructured	7	3	2
Unsuitable to maths	3	14	4
Too subjective	18	14	19
Insufficient resources	0	3	2
Suitable and useful	12	0	0

subjectivity of alternative assessment methods, with teachers across all lengths of teaching experience mentioning this for parental assessment (64% 0–9 years, 54% 10–19 years, 59% 20+ years). Other methods where subjectivity was raised as a concern, using the somewhat arbitrary cutoff of 10% to denote substantive significance, were oral tasks (for 10–19 and 20+ years' teaching), observation (10–19 and 20+ years) and student self-assessment (0–9, 20+ years). Practical tasks and student journals were the only methods not criticised by any teacher group as being subjective.

Student journals and self-assessment were criticised for being unsuitable to maths (0–9, 10–19 years for student journals; and 10–19 years for student self-assessment). Perceiving alternative assessment techniques as too unstructured to be effective was mentioned for observation (14% of 0–9 years) and student journals (17% of 10–19 years). Insufficient time was raised as a deterrent to using oral and practical tasks and student journals, but only by teachers with the most teaching experience. Teachers with 20+ years' experience raised lack of time as an issue for oral (18%) and practical tasks (14%) as well as student journals (14%). Inadequate resourcing was only raised as an impediment for practical tasks by 16% of teachers with 10–19 years' experience. Among the least experienced teachers, positive comments were expressed in relation to both oral (29%) and practical tasks (36%), with comments about them being both suitable and useful assessment methods in mathematics.

5.4. *Main methods of assessment employed by teachers*

In general, the main assessment method teachers reported using was the traditional mathematics test. Other methods they listed were assignments and bookwork/homework (which likely overlap each other), observation,

TABLE III
Teacher attitudes to using various alternative assessment techniques according to length of teaching experience

	Years of teaching experience		
	0–9 years (%)	10–19 years (%)	20+ years (%)
Insufficient time			
Oral tasks	7	0	18
Practical tasks	0	5	14
Observation	0	0	4
Student journals	0	0	14
Student self-assessment	0	0	5
Parental assessment	0	8	0
Too unstructured			
Oral tasks	7	0	5
Practical tasks	7	0	4
Observation	14	0	4
Student journals	7	17	0
Student self-assessment	0	0	0
Parental assessment	7	0	0
Unsuitable to maths			
Oral tasks	7	5	4
Practical tasks	0	5	4
Observation	0	0	0
Student journals	12	38	9
Student self-assessment	0	33	0
Parental assessment	0	0	4
Too subjective			
Oral tasks	0	11	14
Practical tasks	0	0	4
Observation	0	17	14
Student journals	0	0	9
Student self-assessment	43	4	14
Parental assessment	64	54	59
Insufficient resources			
Oral tasks	0	0	0
Practical tasks	0	16	0
Observation	0	0	0
Student journals	0	0	0
Student self-assessment	0	0	0
Parental assessment	0	0	0

(continued on next page.)

TABLE III
(Continued)

	Years of teaching experience		
	0–9 years (%)	10–19 years (%)	20+ years (%)
Suitable and useful			
Oral tasks	29	0	0
Practical tasks	36	0	0
Observation	6	0	0
Student journals	0	0	0
Student self-assessment	0	0	0
Parental assessment	0	0	0

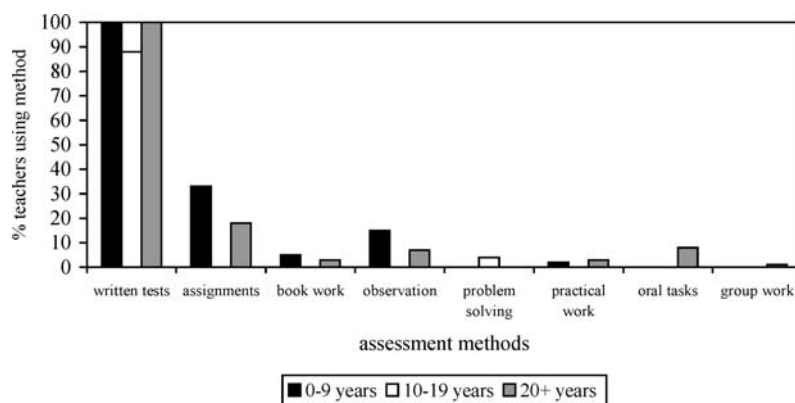


Figure 2. Assessment methods employed by teachers, according to length of teaching experience.

problem solving, practical work, oral work, and group work. Of alternative assessment methods, the most common form used was assignments, followed by observation, with other methods rarely employed. Proportions of teachers using various assessment methods across combined years are represented in Figure 2, according to length of teaching experience. Teachers with the least teaching experience (0–9 years) appeared to use alternative assessment methods more frequently than other teachers, with notable proportions of all three teacher groups using both assignments and observation.

Table IV shows assessment methods used by the teachers broken down for each of years 7–8, 9–10, and 11–12. The traditional written test was overwhelmingly the most common method used for all of years 7–8, 9–10 and 11–12. Teachers having the least years’ experience appeared to use

TABLE IV
Assessment techniques used by teachers for years 7–12, according to length of teaching experience

	Years of teaching experience		
	0–9 years (%)	10–19 years (%)	20+ years (%)
Years 7–8			
Written tests/exams	100	88	100
Assignments	43	0	18
Bookwork/homework	7	0	4
Observation	25	0	9
Problem solving	0	4	0
Practical work	0	0	4
Oral tasks	0	0	9
Group work	0	0	0
Years 9–10			
Written tests/exams	100	88	100
Assignments	36	0	18
Bookwork/homework	7	0	4
Observation	7	0	9
Problem solving	0	4	0
Practical work	0	0	4
Oral tasks	0	0	9
Group work	0	0	0
Years 11–12			
Written tests/exams	100	88	100
Assignments	21	0	18
Bookwork/homework	0	0	0
Observation	14	0	4
Problem solving	0	4	0
Practical work	7	0	0
Oral tasks	0	0	4
Group work	0	0	4

assignments more for younger and less for older students (teachers with 20+ years' experience reported similar, but lower, use for all years, while teachers with 10–19 years' experience did not report use of assignments at any year level). A similar trend was identified for observation, where teachers with 0–9 years' experience appeared to use observation more with years 7–8 and least with years 9–10, while teachers with 20+ years' experience reported using observation similarly with years 7–10 and infrequently for

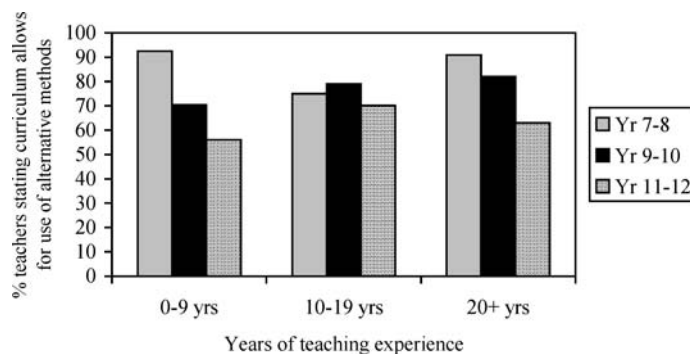


Figure 3. Proportions of teachers perceiving the curriculum to provide for use of alternative assessment methods within years 7–8, 9–10 and 11–12, according to length of teaching experience.

senior years 11–12. Other methods were used too infrequently for patterns to be identifiable across either year level or length of teaching experience.

5.5. Curricular constraints on teacher use of alternative assessment methods

Teachers generally perceived the year 7–8 mathematics curriculum as permitting utilisation of alternative assessment methods, but lower proportions of teachers believed the curriculum for later years allowed for such strategies (except those with 10–19 years' experience reported relatively similar perceptions across years). Figure 3 shows teachers' feelings regarding whether the curriculum permitted them to use alternative assessment methods with each of years 7–8, 9–10, and 11–12, according to length of teaching experience.

5.6. Teacher suggestions for curricular modification and additional alternative assessment methods in mathematics

Suggestions for adaptation of the curriculum and alternative assessment techniques made by teachers are documented in Table V. In relation to curricular adaptation, teachers having the least teaching experience suggested that explicit introduction of instructional strategies such as group work and practical work, may provide opportunities for greater use of alternative assessment formats. One teacher also suggested changing the School Certificate (the external state-wide NSW year 10 assessment), although not how. Teachers with 10–19 years' experience offered no suggestions, with one teacher stating the curriculum could not be changed due to time

TABLE V
Teachers' suggestions for adaptation of the curriculum and suggested alternative assessment methods

Years of teaching experience		
0–9 years	10–19 years	20+ years
How could we adapt the curriculum to use alternative assessment forms?		
<ul style="list-style-type: none"> • More group work • Use of observation • More practical work • Change School Certificate 	<ul style="list-style-type: none"> • Can't change it, due to time factor 	<ul style="list-style-type: none"> • Develop a nation-wide curriculum • Make it less exam/HSC driven • Develop assessment tasks to suit the syllabus • Change teaching methods
What other forms of assessment do you think could be used?		
<ul style="list-style-type: none"> • Open book tests • Assignments • Practical work 	<ul style="list-style-type: none"> • Practical and problem solving questions • Cross-curricular assignments • Oral project work 	<ul style="list-style-type: none"> • Use of novel experiences • Projects • Group projects • Interviews • 'Application' type questions • Mark test/exam with student present • Use of experienced teacher judgements • Research projects

constraints. Teachers with 20+ years' experience suggested development of a nation-wide curriculum, although not why, making the curriculum less examination driven, developing assessment tasks whose content related to syllabus outcomes, and changing existing methods of teaching to permit more opportunities for the use of alternative assessment formats.

Additional forms of assessment that were suggested included a range of specific strategies from teachers with all lengths of teaching experience, with most suggestions coming from the 20+ years group, perhaps supporting the later speculation that these teachers may already have had negative experiences using alternative assessment methods. Some, such as open book tests, were not markedly different from the traditional written test. Additional suggested strategies were assignments and projects, including group and oral projects, practical and problem-solving tasks,

novel experiences, interviews, cross-curricular assignments and the use of experienced teacher judgements.

6. DISCUSSION

6.1. *What is the state of teacher use of alternative assessment methods in maths?*

As anticipated, teachers reported using the traditional mathematics test as their main assessment method, with most frequently used forms of alternative assessment being assignments and observation. Analyses of behaviours and attitudes by length of teaching experience suggest some interesting relationships, although when the group of 60 teachers was sub-divided by length of service the sub-groups were too small to make strong inferences. The response rate of 68% also may have resulted in a biased sample limiting the generalisability of findings. However, results provide interesting insights and directions which could fruitfully be explored with a wider sample.

Results indicate that teachers use alternative assessment methods in similar measures, irrespective of length of teaching experience, with the main alternative forms used including observation and oral and practical tasks. However, it was not made clear in the survey instrument that assessing for performance was meant, so many teachers responded affirmatively if they used the techniques in shaping their personal opinions of students. Future research should clearly distinguish between formal and diagnostic assessment purposes in investigating teachers' use of and attitudes towards alternative assessment methods. In the present study, such a lack of distinction makes it possible that results may be more positive than would otherwise be the case. Similarly, reliance on self-report may have led to an overly positive slant towards responses perceived as desirable. Teachers may also have responded affirmatively even if they did not *effectively* implement alternative assessment practices. Future research should include observational components to overcome these limitations arising from self-report. In the present study, the heavy reliance on traditional mathematics tests may be of more concern in this case.

6.2. *What are teachers' attitudes about using alternative assessment methods?*

Regardless of length of teaching experience, satisfaction with the written test as accurately portraying student capabilities increased with school year. That is, all teachers felt the written test was a good measure of mathematical

abilities for years 11–12, less good for years 9–10, and least good for years 7–8. This finding seems to be at odds with the increasingly higher-order and complex cognitive skills being assessed in later school years, which may be less validly assessed via traditional mathematics tests. Perhaps this reflects the quandary that the more complex abilities a task encompasses, the more difficult it is to interpret its outcome in a reliable way (Niss, 1993a). Alternatively, this could relate to the increasingly crowded nature of the mathematics syllabus with school year (Clarke, 1987), or to a backwash effect brought about by an increasing focus on external examinations which privilege the written test. A trend for teachers to feel the curriculum becomes increasingly less amenable to flexibility in assessing with school year, lends support to the notion that there is *less* provision for flexibility in assessment with senior years. This is likely to be particularly true in an era of increased accountability in mathematical competency testing, both in Australia and overseas (see Levinson, 2000 for a review). This can conflict and create dilemmas for teaching and learning (Webb, 1993), and disempower teachers in impeding expression of their professional responsibility (Galbraith, 1995b). Students may also feel both disempowered and deskilled, when assessment procedures communicate evaluations which discount skills which students possess (Clarke, 1996).

There was a trend for the least experienced teachers to be least satisfied with the traditional test, and most experienced teachers to be the most satisfied, which may well relate to differences in teacher preparation experiences. Newer teachers are more likely to have been exposed to examples of and arguments for alternative assessment methods. Alternatively, those teachers with most experience may have previously had experiences with alternative assessment methods, and based on negative experiences decided that such methods are not suitable or feasible. Across all groups of teaching experience however, no group was particularly satisfied with the traditional test as a measure of mathematical abilities for years 7–8. This is particularly interesting, given the problem solving and creative emphases in the syllabus for these grade levels, less likely to be validly assessed via the traditional test. Over-reliance on the traditional test at these grade levels may therefore be particularly problematic.

The main objection to alternative assessment methods was that they were too subjective, with this concern raised by teachers across all lengths of teaching experience. Resourcing issues and perceiving alternative assessment methods as too unstructured were not seen as impediments to their implementation. Some methods, in particular student journals, were not believed to be suited to mathematics, while time constraints appeared relevant for teachers with longer teaching experience. Implications for ways to encourage all teachers to broaden their assessment base would therefore

seem to be through explication of methods by which likely threats to the objectivity of information derived through alternative assessment modalities may be addressed. For example, *subject profiles*, discussed in the following section, may be a useful approach here. Focusing efforts on aspects such as time allowances would appear to be fruitful only for those teachers with the longest teaching experience, while initiatives related to provision of resources would appear generally non-productive. Only the group of teachers with the least teaching experience made positive comments about alternative assessment methods, providing further support for the speculation that recent graduates would be more in favour of the use of alternative assessment methods.

7. IMPLICATIONS

Teachers' suggestions for changes to the mathematics curriculum included focusing on the match between instructional goals and assessment tasks; using more alternative methods of assessment; and addressing the 'backwash effect' brought about by the requirements of external examinations driving the curriculum, by making the curriculum less driven by final senior high State-wide examination requirements, as well as changing the nature of external examinations. This backwash effect has been frequently identified as constraining teacher use of alternative assessment methods, resulting in alienation of assessment from learning, with educational values becoming compromised by the expectations and accountability of an imposed system of assessment (see Galbraith, 1995a). Teacher suggestions for alternative methods of assessment showed an emphasis on process rather than product, which seems somewhat at odds with their resistance or indifference to alternative assessment techniques.

Given concerns about over-reliance on the traditional written test related to differences in learner characteristics, sampling of the content domain and suitability of this modality for assessing recently emphasised 'real life' and higher-order cognitive skills, the predominance of these teachers' use of traditional tests is of some concern. Since the major objection to the use of alternative assessment methods appeared to be their subjectivity, which has also been found in previous studies (see Galbraith, 1995b; Watson, 2000), it may be the perceived lack of structure and systematisation in informal assessment that leads to lesser status compared with test scores (Clarke, 1987). However, the difficulty of designing appropriate assessment tasks should not be used as a justification for maintaining the current emphases in conventional assessment practices (Thompson and Briars, 1989). A major reason for teachers' avoidance of alternative assessment

methods identified in this study, was their perception of those methods as unreliable. Any efforts to reform assessment practices must therefore address teachers' perceptions of alternative assessment methods as unreliable.

One approach to addressing concerns about the limitations of 'traditional' tests, which also relates to concerns about the potential subjectivity of alternative methods of assessment, has been the development of 'subject profiles' as frameworks for assessing, recording and reporting students' progress. Subject profiles were introduced in 1986 in the Australian State of Victoria, following recommendations for a national focus on assessment and monitoring of standards (Dawkins, 1988). Sadler (1987) also developed a set of operational principles for the use of standards-referenced assessment in the Australian State of Queensland. In the mathematics domain specifically, rubrics for criteria and standards across a range of mathematical skills have been developed (e.g., Galbraith and Clatworthy, 1990; WSIHE, 1989), with such initiatives leading to the Australian Mathematics Profiles document in 1994 (AEC, 1994). This framework is available to teachers throughout Australia to monitor student achievement, although its use is not mandated, since Australian education is controlled at the state and not the national level.

Subject profiles, or 'developmental achievement maps' (Griffin, 1990), provide an explicit identification of outcomes and a framework against which an individual's progress can be traced. These outcomes form a growth continuum for the skill or ability being assessed, typically marked with descriptive indicators for varying stages of developing competence. Subject profiles function as a shared framework, allowing teachers, schools and school systems to communicate about student progress and achievement using a language and standards which are consistent across classrooms, schools and school systems. This approach therefore addresses issues relating to comparability, and provides sets of conditions and marking guidelines which provide for multiple criteria, allowing teachers to use a variety of formal and informal techniques to evaluate student performance and progress, based on a valuation of the validity and professionalism of teacher judgements (Galbraith, 1995b). Utilisation of such an approach should also facilitate a reduction in the amount of formal testing, which often disrupts rather than serves learning sequences.

Although subject profiles focus on the *reporting* of assessment information, they do provide a quality assurance framework within which alternative assessment methods may be used. Subject profiles draw on the professional ability of competent teachers to make sound qualitative judgments of the kind they make constantly in teaching (Sadler, 1987), and addresses the challenge of combining assessment information from a

variety of sources and measures. In making use of the profiles it is assumed a wide range of both formal and informal assessments could be used to arrive at professional judgements regarding students' achievement levels, results of which could then be reported using the common language of the national profiles (Hill, 1994, p. 38). A student's location on a subject profile shows his or her progression through its ordered sequence of competence levels, whereby what is meant by progression in a subject is made explicit, providing a framework against which the development of an individual can be charted (Rowe and Hill, 1996, p. 318). Aside from advantages such as these, subject profiles also address concerns about the subjectivity often associated with alternative assessment techniques. While traditional tests have often sacrificed validity through addressing reliability issues by focusing on consistency in testing, subject profiles permit use of the full range of assessment methods within a quality assurance framework, without sacrificing validity (Rowe and Hill, 1996, p. 340). Here is an approach then that addresses some of the major criticisms to the use of alternative assessment techniques. Confidence in judgements about student capabilities would also be increased if students performed in consistent ways on a variety of tasks requiring a range of mathematical thought (Thompson and Briars, 1989).

Although alternative assessments such as observation, interviews, demonstrations and practical investigations have been suggested in New South Wales syllabus documents for the last two decades or more, there is no imperative for teachers to utilise such methods. A useful policy approach could therefore be to explicitly require teachers to implement a range of alternative assessment methods in mathematics. A major contribution of the present study is to recognise that any such requirement must address teachers' perceptions of alternative assessment methods as unreliable, or else likely fail. Positive attitudes amongst educators should be fostered, so that teachers will be more likely to use a range of assessment methods in addition to the traditional mathematics test, providing a more rounded and accurate portrayal of student capabilities.

ACKNOWLEDGMENT

I would like to acknowledge the guidance and feedback of Mr. Lindsay Grimison from the Faculty of Education, University of Sydney in the beginning stages of this project conducted during my Honours degree. Thank you to Dr. Paul Richardson from the Faculty of Education, Monash University, as well as three anonymous reviewers for feedback on the final draft.

REFERENCES

- Australian Education Council [AEC]: 1994, *Mathematics – A Curriculum Profile for Australian Schools*, Curriculum Corporation, Carlton, Victoria.
- Broadfoot, P.: 1996, *Education, Assessment and Society*, Buckingham, Open University Press.
- Clarke, D.: 1998, *Mathematics Curriculum and Teaching Program: Professional Development Package: Assessment Alternatives in Mathematics*, Curriculum Development Centre, Canberra.
- Clarke, D.: 1987, 'A rationale for assessment alternatives in mathematics', *Australian Mathematics Teacher* 43(3), 8–10.
- Clarke, D.: 1996, 'Assessment', in A.J. Bishop, K. Clements, C. Keitel, J. Kilpatrick and C. Laborde (eds.), *International Handbook of Mathematics Education*, Kluwer, Dordrecht, pp. 327–370.
- Clarke, D. and Lovitt, C.: 1987, 'MCTP assessment alternatives in mathematics', *Australian Mathematics Teacher* 43(3), 11–12.
- Cockroft, W.H.: 1982, *Mathematics Counts*, HMSO, London.
- Corbett, H.D. and Wilson, B.L.: 1991, *Testing, Reform, and Rebellion*, Ablex, Norwood NJ.
- Dawkins, J.S.: 1988, *Strengthening Australia's Schools: A Consideration of the Focus on Content of Schooling*, Australian Government Publishing Service, Canberra, ACT.
- Delandshere, G. and Petrosky, A.R.: 1998, 'Assessment of complex performances: Limitations of key measurement assumptions', *Educational Researcher* 27(2), 14–24.
- Findlay, J.: 1987, 'Criteria-based assessment in Queensland', *Australian Mathematics Teacher* 43(3), 4–6.
- Firestone, W.A., Winter, J. and Fitz, J.: 2000, 'Different assessments, common practice? Mathematics testing and teaching in the USA and England and Wales', *Assessment in Education* 7(1), 13–37.
- Galbraith, P.: 1995a, 'Assessment in mathematics: Purposes and traditions', in L. Grimison and J. Pegg (eds.), *Teaching Secondary School Mathematics: Theory into Practice*, Harcourt Brace, Sydney, pp. 271–288.
- Galbraith, P.: 1995b, 'Assessment in mathematics: Developments, innovations and challenges', in L. Grimison and J. Pegg (eds.), *Teaching Secondary School Mathematics: Theory into Practice*, Harcourt Brace, Sydney, pp. 289–314.
- Galbraith, P.L. and Clatworthy, N.J.: 1990, 'Beyond standard models: Meeting the challenge of modelling', *Educational Studies in Mathematics* 21(2), 137–163.
- Geertz, C.: 1993, *The Interpretation of Cultures: Selected Essays*, Fontana, London.
- Goodlad, J.A.: 1984, *A Place Called School*, McGraw-Hill, New York.
- Griffin, P.E.: 1990, 'Profiling literacy development: Monitoring the accumulation of reading skills', *Australian Journal of Education* 34, 290–311.
- Grimison, L.: 1992, *Assessment in Mathematics – Some Alternatives*, Mathematics Education Research Group of Australasia 15th Annual Conference, University of Western Sydney.
- Hill, P.W.: 1994, 'Putting the national profiles to use', *Unicorn* 20, 36–42.
- Hoge, R.D. and Coladarci, T.: 1989, 'Teacher-based judgments of academic achievement: A review of literature', *Review of Educational Research* 59(3), 297–313.
- Karmelita, W.: 1987, 'The assessment tail should not be wagging the pedagogical dog', *Australian Mathematics Teacher* 43(3), 25–26.

- Krippendorff, K.: 1980, *Content Analysis: An Introduction to its Methodology*, Sage, Beverly Hills, CA.
- Kulm, G.: 1994, *Mathematics Assessment: What Works in the Classroom*, Jossey-Bass Publishers, San Francisco.
- Lacey, C. and Lawton, D.: 1981, *Issues in Evaluation and Accountability*, Metheun, London.
- Leder, G.C., Brew, C. and Rowley, G.: 1999, 'Gender differences in mathematics achievement – Here today and gone tomorrow?', in G. Kaiser, E. Luna and I Huntley (eds.), *International Comparisons in Mathematics Education*, Falmer Press, London, pp. 213–224.
- Leung, F.: 1995, *Mathematics Assessment in Hong Kong*, paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Levinson, C.Y.: 2000, 'Student assessment in eight countries', *Educational Leadership* 57(5), 58–61.
- Mathematical Sciences Education Board: 1993, *Measuring Counts: A Conceptual Guide for Mathematics Assessment*, National Academy Press, Washington, DC.
- Miles, M.B. and Huberman, A.M.: 1994, *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed., Sage, London.
- NCTM: 1989, *Curriculum and Evaluation Standards in School Mathematics*, NCTM, Reston VA.
- New South Wales Board Of Studies: 1996, *Assessing and Reporting Using Stage Outcomes. Part 1: Assessing*, Author, Sydney.
- Niss, M.: 1993a, 'Assessment in mathematics education and its effects: An introduction', in M. Niss (ed.), *Investigations into Assessment in Mathematics Education: An ICMI Study*, Kluwer, Dordrecht, pp. 1–30.
- Niss, M.: 1993b, 'Introduction', in M. Niss (ed.), *Icases of Assessment in Mathematics Education: An ICMI Study*, Kluwer, Dordrecht, pp. 1–8.
- Niss, M.: 1999, Aspects of the nature and state of research in mathematics education. *Educational Studies in Mathematics* 40, 1–24.
- Popham, J.: 1987, 'The merits of measurement driven instruction', *Phi Delta Kappan* 68, 679–682.
- Powell, A.G., Farrar, E. and Cohen, D.K.: 1985, *The Shopping Mall High School*, Houghton Mifflin, Boston MA.
- Queensland Board of Senior Secondary School Studies [QBSSS]: 1992, *Syllabus in Senior Mathematics*, QBSSS, Brisbane.
- Rothman, R.: 1995, *Measuring Up: Standards, Assessment and School Reform*, Jossey Bass, San Francisco CA.
- Rowe, K.J. and Hill, P.W.: 1996, 'Assessing, recording and reporting students' educational progress: The case for 'subject profiles"', *Assessment in Education* 3(3), 309–352.
- Sadler, D.R.: 1987, 'Specifying and promulgating achievement standards', *Oxford Review of Education* 13(2), 191–209.
- Schoen, H.L.: 1989, 'Implementing the Standards: Beginning to Implement the 'Standards' in Grades 7–12', *Mathematics Teacher* 82(6), 427–430.
- Simon, M. and Forgette-Giroux, R.: 2000, 'Impact of a content selection framework on portfolio assessment at the classroom level', *Assessment in Education* 7(1), 83–101.
- Smith, M. and O'Day, J.: 1991, 'Put to the test: The effects of external testing on students', *Educational Researcher* 20(5), 8–12.
- Stephens, M.: 1987, 'Towards an AAMT policy on assessment and reporting in school mathematics', *Australian Mathematics Teacher* 43(3), 2–3.

- Stephens, M.: 1988, 'AAMT discussion paper on assessment and reporting in school mathematics', *The Australian Mathematics Teacher* 44(1), 16a–16c.
- Stephens, M. and Money, R.: 1993, 'New developments in senior secondary assessment in Australia', in M. Niss (ed.), *Investigations into Assessment in Mathematics Education: An ICMI Study*, Kluwer, Dordrecht, pp. 155–172.
- Strauss, A. and Corbin, J.: 1990, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage, Newbury Park, CA.
- Strauss, A.L. and Corbin, J.: 1998, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd ed., Sage, Thousand Oaks, CA.
- Thompson, A. and Briars, D.: 1989, 'Assessing students' learning to inform teaching: The message in NCTM'S evaluation standards', *Arithmetic Teacher* 37(4), 22–26.
- Watson, A.: 2000, 'Mathematics teachers acting as informal assessors: Practices, problems and recommendations', *Educational Studies in Mathematics* 41, 69–91.
- Webb, N.: 1993, 'Visualising a theory of the assessment of students' knowledge of mathematics', in National Council of Teachers of Mathematics [NCTM], *Investigations into Assessment in Mathematics Education: An ICME Study*, NCTM, Reston, Virginia.
- Weber, R.P.: 1990, *Basic Content Analysis*, 2nd ed., Sage, Newbury Park, CA.
- West Sussex Institute of Higher Education [WSIHE].: 1989, *Practical Suggestions for Developing and Assessing Mathematics Coursework*, WSIHE, London.

HELEN M.G. WATT

Gender and Achievement Research Program
University of Michigan
610 East University # 4204D
Ann Arbor MI 48109-1259, USA