Numeracy assessment and associated issues

Dr Jan Lokan, ACER
Mr Brian Doig, ACER
Ms Catherine Underwood, ACER

Acknowledgement

This paper was commissioned by the Australian Association of Mathematics Teachers Inc. and funded by the Commonwealth Department of Education, Training and Youth Affairs.

Note

This paper was prepared in 1998/1999 and some information in the paper may not reflect more recent developments.

Disclaimer

The views expressed in this paper do not necessarily represent the views of the Commonwealth Department of Education, Training and Youth Affairs or the Australian Association of Mathematics Teachers Inc.

Copyright

Contents

Background

The 1997 National Literacy and Numeracy Plan provides impetus for a fresh look at numeracy-related activities in Australian education systems. Much more has been written about teaching and learning activities designed to encourage numeracy development than has been written about assessment. This paper considers the challenges for assessment of current ideas of what constitutes a 'numerate citizen', with a focus on both national and international perspectives.

Since its introduction by Crowther in the UK in the late 1950s, the term 'numeracy' has been redefined so many times, usually meaning different things in different countries, that we need to establish the meaning we are assigning to it before we proceed with further discussion of assessment issues. A consequence of this variety of meanings is that caution must be exercised in dealing with research or other writings that address 'numeracy'. Unless there is a clear definition of the way the term is being used, careful reading is needed to ascertain what the writers mean by 'numeracy' (and sometimes this is not possible to determine).

Definitions

Within Australia, each State and Territory education system has either adopted or is developing its own definition, and consequent assessment policy, for numeracy. Rather than examine these individually we shall take the 'common understandings' delineated by the Australian Association of Mathematics Teachers (AAMT) (1997, p. 11) as a basis for the discussion in this paper.

The key elements of these common understandings are that numeracy is essentially the effective use of mathematics 'to meet the general demands of life at home, in paid work, and for participation in community and civic life'. It is deemed to be a responsibility of schools, and is:

- distinct from literacy
- more than number sense
- not only school mathematics, and
- cross-curricular, (AAMT, 1997 p. 39).

This definition, accepted by a wide range of mathematics and other educators in Australia, is much broader, for example, than the definition of numeracy recently advocated in the United Kingdom: that 'numeracy is the ability to process,

communicate and interpret numerical information in a variety of contexts', (Askew, Brown, Rhodes, Johnson & William, 1997, p. 4). Under this definition there is a clear focus on skills with numbers. An even narrower definition is provided by a researcher from the Educational Testing Service in the USA: numeracy is a person's 'proficiency in understanding quantitative phenomena that are presented in a tabular way' (Wainer, 1997). Other areas of mathematics and mathematical ideas do not enter the picture in either of these definitions.

It is worth commenting that the 1997 forum of Australian mathematics educators who defined the boundaries of numeracy referred to above, believe strongly that numeracy should not be viewed as encompassed by literacy. Some aspects of 'literacy', in terms of reading, writing and speaking, impinge on numeracy skills, in that a person's use of mathematics to meet the demands of everyday life needs reading skills (to interpret contexts) and communication skills as well as some kinds of mathematical skills. If the term 'literacy' is used in the very broad sense of meaning a 'competency', then 'numeracy' *is* a 'literacy'. To an extent this is a matter of semantics, but those who argue that 'quantitative literacy' should be subsumed in a general definition of 'literacy' (for example, the International Adult Literacy Survey [IALS]) could be perceived as doing 'numeracy' a disservice. In the Australian view, 'numeracy' is more than skills with numbers and the ability to read material presented in tables and graphs, which, from the IALS tests, is what is meant by 'quantitative literacy'.

Assessment purposes and practices

Purposes

Assessment needs to be seen in the context of the needs of the major stakeholders in education. Of these, two groups are perhaps the most important and demanding of attention: teachers and education system administrators. (This is not meant to deny the needs of parents or students). The purposes of both of these groups for collecting assessment information are related to an interest in improving student learning, though one much less directly so than the other. In one case the main purpose may be to provide evidence of class level achievement or it could be primarily for diagnosis, with a view to refining teaching programs based on individual students' needs. In the other case the intention is to provide evidence of educational achievement at school or system level. It is useful to think of a continuum of assessment purposes, ranging from the informality of some teacher assessments in day-to-day teaching to the formality of statewide assessments.

Professor Wynne Harlen and her colleagues (1994) have listed five main principles for assessment in education. These are that assessment:

- is a continuous part of teaching and learning
- should improve learning by exerting a positive force on the curriculum (hence an assessment program must reflect the full range of curriculum goals)
- provides a good means of communication with parents to help support their children's learning
- must involve procedures suitable to the purpose
- must be used fairly by schools.

A sixth principle, also stated by Harlen et al., is that citizens have a right to information about the standards being achieved in an educational system. We have separated this point from the other five because it is the only one that clearly relates to the formal end of the assessment purpose spectrum.

The professional value of good assessment programs is clearly recognisable in the above list of principles. The educative value may arise through teachers' exposure to innovative assessment measures, through exposure to well-designed assessment tasks and procedures, and through learning more about their students' learning. The desirable continuing nature of assessment is projected, as is the use of assessment results to diagnose students' weaknesses (or strengths) or to improve a teaching program. Other authors have recognised that assessment, particularly assessment for accountability, is bound to drive the curriculum to some extent. For example, Resnick and Resnick (1992), leaders of the New Standards movement in the USA, put forth a plea for assessments to be made 'worth teaching to'. According to Wiggins, a prominent writer in the field of innovative assessment, 'we should not feel despair about such a view… a school *should* "teach to the test"… The test must offer students a genuine intellectual challenge' (1989, p. 704).

In the USA, there has been a rush to expand or introduce testing programs since the publication of the report, *A Nation at Risk* (Office of Educational Research and Improvement, 1983). This expansion of testing led Petrie (1987) to write that 'it would not be too much of an exaggeration to say that evaluation and testing have become the engine for implementing educational policy' (p. 177). In a review of US statewide assessments of 'standards' carried out in 1989–90, Coley and Goertz (1990) found that only two states carried out minimum competency assessment. Only two states developed assessment materials based on an agreed curriculum. A few states did not have statewide assessment programs at that stage and the others, almost 40 in number, had committees of assessment experts develop tests, by means of which the state standards were defined. No doubt these experts had some concept of what the curricula might be when they developed the tests, but the curricula did not shape

the tests in the way that they do in Australia, as discussed in the next paragraph. In these circumstances in the US, it is quite likely that the tests had considerable influence on curricula, rather than the other way around. By 1998, the situation in the US had reversed: 47 states either had assessments aligned with statements of standards or were planning to do so, rather than having the assessments *determine* the standards (Glidden, 1998).

In Australia, tests arising directly from curricula are developed for statewide assessment. For example, in reporting to the Minister on their Year 6 tests (now Year 5 tests), officers of the Queensland School Curriculum Council wrote that 'the Year 6 test items were developed with the Queensland syllabuses and classroom curriculum contexts very much in consideration…. The items presented in each of the Year 6 tests have the potential to heighten teacher awareness of the syllabus and highlight the linking of assessment to aspects of the Queensland curriculum' (1998, p. 6). In Australia to date, changes in the foci of assessment have lagged behind changes in curriculum content or approaches.

The preceding comments pertain more to tests used for accountability than other purposes. Tests for a teacher's own purposes are likely to be quite different in nature, with more items related to a topic or skill with which students may be having difficulty rather than small numbers of items on a wide range of topics. It is important, as Harlen states, to tailor the nature of the test to the purpose at hand. It is also very important for assessments to change in style as curriculum emphases shift. 'As the learning needed for the future changes, we must adapt assessment to the changes' (ETS Annual Report, 1995).

As we saw above, numeracy education, as viewed in Australia, has partly different emphases from mathematics education. 'Numeracy' goes beyond school mathematics but at the same time is very much less than mathematics. The broader the definition of numeracy that we accept, and the more it is related to life experiences than to (or as well as) school experiences, the more difficult it is to devise ways of assessing the full range of numeracy skills. To assess the full range of numeracy skills, the boundaries of what is known about assessment will need to be 'pushed'. Some promising attempts in this direction are referred to in the next section and also later in this paper. Even so, it is not feasible to think of assessing students' numeracy skills in actual, real life situations. At best, situations have to be simulated. Further, in the reality of day-to-day teaching, such time consuming assessment activities — even if good assessment practice — are a strain on resources and difficult to implement. These points are elaborated in the discussion of particular programs later in the paper.

Overview of practices

Good, rich assessment programs have the potential to yield information for parents on their children's achievements and progress. The move in Australia in recent years to outcomes-based approaches to both teaching and assessment has facilitated the communication of meaningful information, particularly where the defined outcomes clearly follow a developmental pattern (as is the intention with the Australian curriculum profiles and their State variants). The move to descriptive reporting in many of Australia's statewide assessment programs has provided for better communication of children's achievement to parents. These programs use intended levels of outcomes, in combination with item response modelling techniques, to group clusters of tasks in order to define levels of development along the learning continuum being assessed. Student results are both displayed in relation to the whole continuum and described in terms of the skills typically demonstrated by students at that level of achievement.

Statewide assessments

All Australian States and Territories assess mathematics or 'aspects of numeracy' on a statewide basis, some annually, some at two-yearly intervals, some less frequently but on a regular cyclic basis interspersed by assessments in other learning areas. Some States undertake light sample testing in order to monitor standards (for example, Western Australia); most others undertake cohort testing in order to monitor standards *and* report individual students' performance to parents. The introduction of reporting against numeracy benchmarks means that all States and Territories undertake some kind of cohort assessment. Typically, students are first assessed in this way in Year 3, with one or more later assessments. Typical years for later assessments have been the year before the end of primary school (that is, Year 5 or Year 6) and a year in early secondary school. The introduction of numeracy benchmarks will most likely mean that all States and Territories will bring their assessments into line, to focus on Years 3, 5 and 7, extending to Year 9 as benchmarks for that stage become available. A table showing details of system monitoring programs by State and Territory at the end of 1998, is included in Table 1 in the appendix to this paper.

Diagnostic assessment

Statewide assessment programs provide only a single occasion 'snapshot' of

students' achievement levels. They provide the opportunity for all concerned to have a picture of the students' skills against a wider frame of reference. The vast majority

of assessment is carried out by teachers for purposes different from the purposes of statewide assessments. Teachers are the primary assessors of students and they typically assess in the many ways that reflect the goals of teaching and learning, particularly ways described as 'diagnostic'. Teachers are well aware of the necessity for early diagnosis of students who may be 'at risk' of failing to reach adequate levels of proficiency, particularly in the basic areas of literacy and numeracy. They are also aware of the need to monitor student progress as part of their day-to-day teaching practices, and most are aware of the value of obtaining assessment information so that they can adjust their teaching program to the needs of their students.

Diagnostic assessment practices outlined here are those forming part of 'intervention' programs currently in place around Australia, as well as those used by teachers and psychologists working with students requiring or likely to require remedial instruction. For the latter purpose individual professionals sometimes construct their own tests, but there are also commercially available tests. For the former purpose, a range of methods is used to identify the students considered to be 'at risk' and therefore in need of additional support. Diagnostic assessments are most often carried out in the early years of school, though States and Territories with cohort testing programs often use results on those as a further opportunity to flag potential problems. (Cohort testing programs can only be used as tentative 'pointers', because they survey a wide range of topics with at most one or two items per topic, whereas diagnostic tests contain several items per topic). A brief summary of State and Territory diagnostic programs is given in the next two paragraphs. More details are given in the section on 'Specific practices' and in Table 2 in the appendix.

The ACT makes use of results from its cohort testing program to identify the lowest achieving 20 per cent of students. In line with their assessment, focal years are Years 3, 5, 7 and 9. New South Wales has a broad screening process which classroom teachers involved in the *Early School Assessment* project in Years K–3 use. There is a particularly strong focus on assessing students in their first year at school. The locally developed *Schedule of Early Number Assessment* (part of the *Count Me In Too* package) is used for this purpose.

Victoria has an *Early Years Numeracy Program*, a component of which is the identification, through comparison with developmental stages, of 'at risk' students. Victoria has evaluated the NZ *School Entry Assessment* (described later) which is applicable to all students but is also a vehicle for identifying students in need of additional help. Queensland's *Diagnostic Net* is used in that State during the first

three years of compulsory schooling, but mainly at Year 2. By means of a thorough process of observing and mapping progress on developmental continua, teachers identify children requiring additional assistance. In this case the main assessment method is teacher observation combined with careful recording.

South Australia is currently introducing their *School Entry Assessment Program*. Teachers will assess students' numeracy through observation during normal classroom teaching but also through use of some specially developed assessment tools. The program is for all students but it will enable 'at risk' students to be identified and then helped. Western Australia has just begun to implement a *Students at Educational Risk* program, in which teachers develop profiles of students' achievements and use these in relation to typical expectations to identify students who need additional support. The *First Steps* program has recently been expanded to include numeracy as well as literacy. There is also a proposal in WA to develop a *Numeracy Net*, which would focus on the identification of misconceptions and on planning curricula to assist in remedying these.

Tasmania began its *Flying Start* program in 1997. In this program there is a focus on professional development for K–2 teachers, on reducing class sizes and the appointment of resource teachers to assist with early identification and intervention for students deemed to be 'at risk'. While literacy and other areas are also part of *Flying Start*, there is a continuing emphasis on numeracy skills, reinforced by the Government's identification of numeracy as a priority. The intention to strengthen the program from 1999 saw the introduction of an early 'number development' element based on Wright's *Count Me In Too* materials in use in NSW. In the NT three programs were trialled and evaluated in 1998, with a view to recommending one to be adopted from 1999. One of the programs was the NZ *School Entry Assessment package* (see later), used as students move from Transition to Year 1. The NT has also developed its own *Assessment in the Early Years*, a guide for teachers on strategies for identifying students at risk of not achieving at appropriate benchmark levels.

Commercially available diagnostic tests for various aspects of mathematics and various grade levels are primarily the *Booker Profiles in Mathematics* (Booker, 1994); the *Diagnostic Mathematical Tasks* (Schleiger & Gough, 1993); the *Diagnostic Mathematics Profiles with DIAMAP* (ACER, 1991); *KeyMath* (American Guidance Service, 1997); the *Mathematics Topic Pre-Tests* (NZCER, 1990); and *Stop! Look & Lesson* (Palmer, Kays, Smith & Doig, 1994). Of these, *KeyMath* seems, on the basis of sales volume, to be more widely used than the others, but none is used extensively. The *Stop, Look & Lesson* materials are likely to be particularly useful for diagnostic purposes in that they provide a thorough classification of errors made by students in carrying out mathematics assessment tasks, which point to particular misconceptions the students are likely to have.

Assessment of giftedness

All States and Territories currently have some kind of program for academically gifted students in place, usually covering all areas of the curriculum, but including

some with a special emphasis on mathematics or, in Tasmania, numeracy. Other types of giftedness are also catered for in most cases. Students are chosen for the programs in different ways, including student preference, teachers nominating students and students' performance on specially designed selection tests. New South Wales, South Australia and Western Australia each have their own selection test programs.

Instruments for classroom assessment

Most assessment materials used by teachers in the course of their teaching programs are constructed by the teachers themselves. For those who wish to use externally referenced measures, referenced either to other populations of students or to pre-defined standards of some kind, a range of possibilities is available commercially. These include the *Developmental Assessment Resource for Teachers (DART) — Mathematics* (Recht, Forster & Masters, 1998); the *Diagnostic Mathematical Tasks* (Schleiger & Gough, 1993) which, despite their name, are used as regular classroom assessment measures as well as for diagnostic purposes; the *Group Review of Algebra Topics* (Doig, 1991); the *Orchid* Series (ACER, 1997); the *Profiles of Problem Solving* (Stacey, Groves, Bourke & Doig, 1993); and the *Progressive Achievement Tests in Mathematics* (ACER, 1997). Of these, those that yield the most pertinent information for Australian use are the *DART* mathematics materials and the *Orchid Series*, because of the wide range of types of assessment tasks used together with their link to the mathematics curriculum profiles. Close behind are the newly revised *Progressive Achievement Tests in Mathematics (PAT-Maths),* which are now also linked to the mathematics curriculum profiles.

Assistance for teachers in constructing their own assessment tasks is offered in a recent series of magazine-type publications as part of the DETYA-supported *Assessment Resource Kit (ARK)* (Masters & Forster, 1996). These booklets are written with teachers as the intended audience, and describe how to design and use a range of types of assessment tasks: pen and paper, performances, portfolios and projects. The rationale and benefits of developmental assessment are also described in the *ARK* series, as are 'progress maps' and how to use these to record and report students' learning progress. The *ARK* materials have been received enthusiastically by systems and teachers in many parts of Australia, particularly in South Australia where copies have been supplied to schools by the Department of Education, Training and Employment. Workshops for systems officers and teachers on the material covered in the *ARK* magazines have been conducted in many parts of Australia over the past two years, supported by funding from DETYA.

Specific practices

Given the interest expressed in the National Literacy and Numeracy Plan in the early identification of students who are considered to be potentially at risk in terms of their future educational success, some of the assessment strategies to enable early identification of such students are described in a little more detail here. The descriptions are illustrative rather than exhaustive. There is emphasis on programs functioning on a statewide basis, though some of the most promising programs are currently in use on a smaller scale. For some of these there is an intention to adopt them on a wider scale pending the success of pilot trials. Others are mentioned because of their demonstrated 'track record' in improving the numeracy skills of students identified as experiencing difficulties. No doubt there are many other programs, for example programs developed and used by individual teachers or groups of teachers, that could be described. However, the purpose of this paper is to highlight issues and make some suggestions for future practice, not to provide an exhaustive review. Selected details of the wide range of strategies and programs used around Australia are presented in Table 2, in the appendix to this paper.

In the next few paragraphs, some of the most prominent and/or successful programs are briefly described State by State. Evaluative comments are made in the section headed 'Some caveats'.

New South Wales: Count Me In Too

The *Count Me In Too* program is based on the research and practices of Wright, in particular, and is funded by the New South Wales Government. The program is an extension of the *Count Me In* professional development materials. Teachers undertake professional development to learn about counting stages and to examine activities based on these stages. They view video-clips in which the strategies and counting stages are highlighted together with teaching suggestions. The basis of the *Count Me In Too* program is the *Schedule for Early Number Assessment*, a diagnostic interview protocol used by teachers with individual students.

The results of each student's interview are related to a learning framework, based on research on children's number development. After the initial interview, continuing assessment forms part of the teaching-learning process. As part of the program, teachers are responsible for interviewing and videotaping a small number of students in their own class to use as a basis for developing their interviewing and interpretative skills. At present the program is being used intensively in pilot schools, with a view to its use becoming more widespread.

New South Wales: Mathematics Recovery

Mathematics Recovery started in 1992 as a three-year collaborative research project in north-eastern New South Wales, jointly funded through the Australian Research Council, New South Wales regional government and the Catholic school system (Wright, Stanger, Cowper, & Dyson, 1996). The program, for selected first-grade students, is a long-term, individualised teaching program with the aim of advancing the students' arithmetical learning to the point where they may return to the regular classroom. It is thus highly remedial in its focus.

The theoretical framework underpinning the *Mathematics Recovery* approach to 'at risk' students is based on constructivist teaching experiments (see for example, Steffe, von Glaserfeld, Richards and Cobb, 1983). In particular, the program documents children's progression through six 'counting stages' identified by research undertaken in the USA, particularly by Steffe (see Steffe, Cobb, & von Glaserfeld 1988). The six counting stages also form the basis of the teaching tasks of the intervention program. Wright and his colleagues have constructed a large bank of teaching tasks for teachers to use in the program. Selections are made from the bank to ensure that the tasks used are suited to the students' identified needs.

The organisation of the *Mathematics Recovery* materials, as their name suggests, is similar to that used for the well-known *Reading Recovery* program: students devote a full morning per week to the program over an extended period (usually twenty weeks).

Queensland: Year 2 Diagnostic Net

The *Year 2 Diagnostic Net* used in Queensland is based upon two assessment approaches. First, detailed descriptive continua of mathematical development are provided. These continua are focused on number, space and measurement, and are divided into key phases that identify significant milestones in development. Thus they are said to 'map' a child's mathematical development. Teachers in the early years are required to observe their students, and record their observations using a checklist of key indicators.

The second aspect of the *Year 2 Diagnostic Net* is 'validation'. In addition to the observation of students' mathematical development, teachers are also required to use a set of 'validation' tasks provided by the State Department of Education. These assessment tasks are carefully designed to provide a 'validation' of the teacher's judgements that had been based solely on observation.

Children who are deemed to be 'at risk' are then provided with a suitable intervention program. As the *Year 2 Diagnostic Net* developmental continua are inter-

linked with the Queensland Year 1 to 10 mathematics syllabus and resource documents, these provide a basis for any program of intervention that teachers may plan to implement based on a child's performance in the *Year 2 Diagnostic Net* developmental mapping.

Schools report to the parents of each Year 1 and Year 2 child in a standard report format. This report describes the phases of development (in literacy and numeracy) and indicates in what phase the child is operating. The reports are followed by parent-teacher interviews where an individual child's development can be discussed in more detail.

Victoria: School Entry Assessment

Following a preliminary review in 1998, in 1999 Victorian primary schools are trialling the New Zealand developed *School Entry Assessment (SEA)* materials. Subsequent to this trial, all schools will have the opportunity of using the materials, which are described below under 'New Zealand'.

Victoria: Mathematics Intervention

While not used in a large number of schools, the *Mathematics Intervention Program* is included here because it is proving to be one of the most promising programs for assisting students based on diagnosis of their problems. *Mathematics Intervention* aims to identify, then assist, students in the first years of school who are 'at risk' of not coping with mathematics. The program offers students the chance to experience success in mathematics by developing the basic concepts of number upon which they build their understanding of mathematics. *Mathematics Intervention* was developed as a collaborative project between university researchers and the principal and staff of a primary school in the metropolitan area of Melbourne. It has since spread to other schools, and a six-day training program in its use is offered for teachers two or three times a year.

The program is based on recent research about children's early arithmetical learning (see, for example, Steffe, von Glaserfeld, Richards and Cobb, 1983; 1988; Wright 1991; Wright et al., 1996). The initial assessment for the *Mathematics Intervention* program requires teachers to assess the extent of the student's mathematical knowledge by observing and interpreting the actions the student takes while working on a set task. As there was no comprehensive test available that allowed students to talk about their mathematical strategies, an instrument, the *Initial Clinical Assessment Procedure-Mathematics (ICAPM)* — Level AA (Pearn, Merrifield, Mihalic, & Hunting, 1994) was developed. All teachers involved with the *Mathematics Intervention* program have

attended a course in *Clinical Approaches to Mathematics Assessment* (Gibson, Doig & Hunting, 1993; Hunting & Doig, 1992) to develop and refine their observational and interpretative skills. The developers of *Mathematics Intervention* believe that this training is a necessary step for teachers working with students 'at risk' in mathematics.

Students are chosen to participate in the program, based on the results of clinical interviews. They are withdrawn from their classes and work in groups of no more than three, with a clinically-trained teacher, to assist with the development of their mathematical language skills and cooperative strategies.

In *Mathematics Intervention* emphasis is placed on the verbal interaction between teacher and students, and between students. Each session is planned to build on previous understandings as interpreted by the teacher during the session. The *Clinical Approaches to Mathematics Assessment* course ensures that teachers can observe what the child is doing, interpret the child's actions, act on these actions and then reflect on the intervention. Experience with the *Mathematics Intervention* program has highlighted several teaching strategies that will allow students to experience success with mathematics (see Pearn & Merrifield, 1996, for examples).

New Zealand

The New Zealand developed *School Entry Assessment* (SEA) is a series of nationally (New Zealand) standardised performance tasks (New Zealand Ministry of Education, 1997). It is summarised here because some Australian States are using it or have decided to adopt it. The numeracy task, *Check O*ut, is in the form of a shopping game. All children entering school are assessed with the SEA in their first two months at school, and within the context of the regular classroom. This period allows children to settle in to school before the assessment takes place.

Check Out is administered individually by classroom teachers, who then interpret the results in terms of the New Zealand curriculum frameworks. The SEA has both summative and formative aspects, thus providing a wide range of information to teachers. It is expected that information gained from the tasks will be used as a basis for program planning. It is also expected that summary data will be reported to the Ministry of Education.

Some caveats

Some of the programs described immediately above are not intended to be used with all students. Rather, they are more probing assessments that are most appropriately used after coarser-grained assessment methods have indicated that particular

students are likely to be experiencing difficulties. Others are recommended for use with all children entering school. In either case, relief time or other assistance for teachers seems to be desirable because of the difficulty of managing the rest of the class at the same time as carrying out a detailed one-to-one interview (or perhaps one-to-two).

It will be particularly interesting to observe the large-scale use of individual assessment of beginning students in Victoria with the SEA materials, given the time required for each student and the detailed record keeping necessitated by the assessment. In a qualitative study of implementation of the SEA in New Zealand, Williams and Dixon (1998) set out to discover how schools were using the information derived from it to enhance teaching and learning. Teachers reported that it was very difficult to manage the detailed administration of the assessment to individual students at the same time as attending to the needs of the rest of the class. It was common for the teachers interviewed to explain that, while they had been trained in the administration of the assessment, they had not received follow-up training in the effective use of formative information about their students derived from it. It has also been noted that the duality of purpose of the SEA (both formative and summative) causes problems and tensions in the kinds of results and reporting demanded, the use to which results are put and the conditions under which the assessment is administered.

The programs described, except *Mathematics Intervention*, all place a heavier emphasis on identifying students at risk than on what to do about problems once these have been identified. In a meeting of mathematics educators held recently at ACER, the point was made that, even with the supposedly thorough Year 2 *Diagnostic Net* in Queensland, there had not been enough thought given to follow-up strategies for students 'caught' by the Net, nor had there been sufficient training provided to teachers in this respect. At an even more recent meeting of representatives of Education System Chief Executive Officers, it was stated that in Western Australia, even with extensive use of the *First Steps* literacy materials, there are still many children with deficient literacy skills who are not improving in the way that had been hoped. The consensus was that there is no clear strategy that will work for all students in a State. The request was made for research to identify which strategies will work, in what circumstances and with which students. As recognised for literacy, it is unlikely that any one intervention program will be the solution to all numeracy learning problems.

To use any of the numeracy assessment programs described in the previous few paragraphs, teachers would ideally be provided with opportunities for professional training in how to administer the assessment and how to interpret and make use of the results. There is definitely a place for these kinds of assessment tasks, administered in very small groups or with individual students, but they are labour

intensive and therefore costly. Their place is for diagnosis and intervention. It would neither be feasible nor necessary to incorporate such procedures into a cohort-based assessment program. In Australia we should learn from experience in the UK with Standard Assessment Tasks at Year 3. Realistic, thematic assessment tasks were developed, designed to be carried out over several weeks. The record-keeping demands on teachers were not intended to exceed 30 hours, but most teachers reported spending more than that amount of time, some saying that they spent more than 100 hours. The assessment has now been modified to eliminate most of the time-consuming, hands-on tasks, to make it more manageable for teachers.

To ensure the most positive benefit from any assessment, teachers will need to be provided with guidelines and suggested strategies for modifying instructional programs based on the results of the assessment. Opportunities for professional development in this regard would be advantageous.

What do research studies indicate about Australia's mathematics expectations and achievement?

IEA studies

First and Second International Mathematics Studies

Australia participated in 1964 in the First International Mathematics Study (FIMS) sponsored by the International Association for the Evaluation of Educational Achievement (IEA). Only two populations of students were sampled: students aged 13 at the time of testing (August); and students in Year 12. The Australian sample for FIMS was restricted to government schools from five States (excluding South Australia). Almost 3000 13 year-old and 1100 Year 12 students were tested. The Year 12 students had to be studying a mathematics course that would qualify them to study mathematics at university level.

Twelve countries participated in FIMS at upper secondary level (Year 12) and ten at lower secondary level. Australia's results placed us almost exactly at the international average for the 13 year-old population, with Japan, Belgium, Finland and the Netherlands as the highest achievers and Sweden and the USA as the lowest. At Year 12, Australia's results were considerably below the international average, with only the USA sample achieving more poorly than the Australian sample. Israel, England, Belgium, France, the Netherlands and Japan performed particularly well in this senior secondary population.

When the items were divided into various subgroupings, by content strand and by higher and lower mental processes, the profile of the Australian 13 year-old students' performance was found to be close to the international average in all aspects except algebra, in which their performance was slightly above average. The Australian Year 12 students performed well below average in all areas except calculus, where their result was about average (Husen, 1967).

Australia did *not* participate in the full Second International Mathematics Study (SIMS) in 1980, in which largely new tests were used. Instead, testing was carried out in two student populations in 1978, using the FIMS tests from 1964. Samples of about 5300 13 year-olds and 2900 Year 12 students from all sectors and from all areas except the Northern Territory were tested in 1978. International comparisons of the 1978 results cannot be reported, but a very detailed analysis of changes in Australian mathematics achievement between 1964 and 1978 is provided in Rosier (1980). Using 1978 data only from students in government schools in the five Australian States which participated in FIMS, Rosier showed that there was a slight overall decline in the achievement of 13 year-olds except for students in Western Australia. Geometry scores declined more than algebra or arithmetic scores, which Rosier attributed to changes in the geometry curriculum from 1964 to 1978. In terms of the process skills tested, Rosier noted a slight decline in each of computation, knowledge, application and comprehension, none large enough to be significant. He commented:

> There was no evidence to indicate improved performance at the higher levels of this hierarchy of skills, although such improvement was one of the aims behind the changes in the curriculum in the 1960s. At the same time, there was no evidence of a marked decline in skills of computation, (p. 190).

By contrast, the achievement of Year 12 students in all States and in all areas of mathematics improved between 1964 and 1978—a salutary finding given the comparatively poor performance of our students at this level in 1964.

Third International Mathematics and Science Study

The most recent IEA study of mathematics learning, the Third International Mathematics and Science Study (TIMSS) was the most comprehensive study of mathematics learning ever undertaken, both in terms of its country coverage and in terms of the scope of data collected. Australia participated in TIMSS at all three population levels: 9 year-olds, 13 year-olds and students in Year 12. Randomly sampled schools and students from all States and Territories and all education sectors took part. Much larger samples of students, from many more countries, were tested in TIMSS than in the earlier mathematics studies.

At each of the 9 year-old and 13 year-old levels, students were sampled from the two adjacent grades containing the majority of the age group. These grades are referred to in all the TIMSS reports as the 'upper' and 'lower' grades within each sample. In Australia, because of the different ages at which students start school State by State, Years 3 and 4 were sampled in some States and Years 4 and 5 in others to capture the 9 year-olds. Similarly, the 13 year-old sample involved Years 7 and 8 or Years 8 and 9.

Australia's comparative results in mathematics in TIMSS were creditable, though not as good as our comparative results in science. The results summarised below are cited from the three international TIMSS reports (Beaton, Mullis, Martin, Gonzalez, Kelly & Smith, 1996; Mullis, Martin, Beaton, Gonzalez, Kelly & Smith, 1997; 1998).

In mathematics:

- at the upper grade for the 9 year-olds, six of 26 countries achieved significantly higher results than Australia, five countries achieved at the same level and 14 countries (including Canada, Scotland, England and New Zealand) achieved significantly lower results; countries outperforming Australia were Singapore, Korea, Japan, Hong Kong, the Netherlands and the Czech Republic;
- at the lower grade for the 9 year-olds, only the four Asian countries listed above performed better than Australia, eight countries performed at the same level and eleven countries' results were lower than Australia's (including Scotland, England and New Zealand) (only 24 countries tested lower grade students);
- at the upper grade for the 13 year-olds, nine of 41 countries achieved significantly higher results than Australia (including the four Asian countries and five European countries), 13 countries achieved at the same level and 19 countries recorded significantly lower results (including New Zealand, England, the USA and Scotland); and
- at the lower grade for the 13 year-olds, seven of 39 countries outperformed Australia (the four Asian and three European countries), 13 tied with Australia and 18 achieved significantly lower results than Australia (including the above four English speaking countries).

In advanced mathematics at Year 12, Australia was in the top group of countries, both in an analysis based on all students tested and in an analysis based on the highest scoring five per cent. In the former analysis six (of 16) countries recorded significantly lower results than Australia, while in the latter analysis ten countries were placed below Australia.

In terms of content areas *within* mathematics:

- the upper grade 9 year-old sample scored relatively better in 'geometry' and relatively worse in 'whole numbers' and 'fractions' than in 'measurement,

estimation and number sense', 'patterns, relations and functions' and 'data representation and analysis';

- the lower grade 9 year-old sample's relative scores were similar to those of the upper grade 9 year-old sample, except that performance in 'fractions' was at an average level;
- the upper grade 13 year-old sample scored relatively better in 'data representation and analysis' and relatively worse in 'geometry' than they did in 'fractions and number sense', 'algebra' and 'measurement';
- the lower grade 13 year-old sample scored relatively better in 'fractions and number sense' and 'measurement and relatively worse in 'geometry' than they did in the two other areas; and
- the year 12 sample scored higher in 'numbers and equations' and 'calculus' than they did in 'geometry'.

Of most interest in these results are:

- the relatively poor performance of the primary age students in 'whole numbers', suggesting a lack of computation skills that was borne out by an examination of performance on individual items (further discussion of this point is provided by Stacey (1997) in the paper presented to the ACER Conference on TIMSS results);
- the better relative performance in 'data representation and analysis' at lower secondary level than at mid-primary level; and
- the change in relative performance on 'geometry' between the 9 year-old sample and the 13 year-old sample (apparently we place more stress on spatial concepts in primary school than we do on these concepts and properties of two-dimensional figures, angles, lines, etc. in lower secondary school and on formal geometry as part of advanced mathematics).

Comparisons of international expectations

In late 1997 and early 1998, ACER staff undertook some analyses of expectations or objectives for mathematics at Years 3 and 5 in several countries, including Singapore, Hong Kong and Japan (three of the four highest achieving TIMSS countries), several states in the USA, the Canadian provinces of British Columbia and Alberta, Hungary and Ireland. These countries' statements of expectations were compared with the October 1997 draft of the Australian numeracy benchmarks, in order to inform the benchmark development and revision.

The comparisons undertaken showed that other countries mostly expected the concepts and skills featured in the Australian benchmarks to be addressed at an earlier grade level (Lokan & Ainley, 1998). Apart from the document from Hungary, the other documents examined did not specify what were considered to be minimum achievement levels and they tended to be more comprehensive than a set of basic benchmarks would be. It also needs to be remembered that the Australian benchmarks are not meant to define the curriculum in any sense – rather, they provide indicators of minimum levels of achievement that virtually all Year 3 and Year 5 students are expected to have attained at those grade levels. Extensions of this work continued in 1998 and 1999.

'Standards' over time within Australia

The Longitudinal Studies of Australian Youth (LSAY) provide linked assessments of numeracy achievement over a period of approximately 20 years for samples of 14 year-old students. The tests used consisted of mostly multiple choice items, but they did make an attempt to cover applications of mathematics in every day situations. Random samples of students were drawn and were tested first in 1975 with items designed to assess minimum competency, but later tests provided information on a wider range of items, covering wider levels of difficulty. Analyses over time are presented in terms of mean scores on a common scale, determined by making use of the common items to link all results by means of item response modelling, and also in terms of the percentages of students attaining a pre-defined level of mastery on the tests. Percentages correct on common items used in 1975, 1980, 1989 and 1995 are also reported.

Among 14 year-olds, the mean scores on the numeracy scale showed that there was no overall change in performance between 1975 and 1995. There was a small improvement over that time interval in the percentage of students attaining mastery. The test items were classified as 'computational', 'practical' (strongly relating to everyday contexts, e.g. hours of opening of a chemist shop) and 'conceptual'. A small decline was noted on the computational items, no consistent change was evident on the practical items and there was a slight improvement on the conceptual items (Marks & Ainley, 1997).

Changes in mathematics achievement over the four year period from 1994 to 1998 will shortly be able to be reported both in terms of two cohorts of the same age (13 year-old students) and in terms of the same cohort progressing from age 9 to age 13. These comparisons will be made possible through the repeat testing of lower secondary level students as part of the worldwide repeat of TIMSS.

Trend data are beginning to be reported from several of the statewide testing programs. Tasmania, with its sample testing of 10 year-old and 14 year-old students every four years, using items from the 1975 Australian Study of School Performance, has reported comparative data on a regular basis. However, as the numbers of test items common to two or more occasions became less and less, it was realised that the comparisons would most likely not be valid. Some recent examples are the report on the 1996 monitoring of mathematics achievement at Years 3, 7 and 10 in the WA Monitoring Standards in Education (MSE) program (van Wyke, 1998), the report of the NSW Basic Skills Testing at Years 3 and 5 (New South Wales Department of School Education, 1998) and the report of statewide performance of students in Queensland on their statewide assessment programs in 1995, 1996 and 1997 (Queensland School Curriculum Council, 1998).

The WA report showed that there was a small improvement in the achievement of students tested in 1996 and in 1992. At Year 3, the difference was not significantly different from zero, but at Years 7 and 10 significantly higher means were recorded in 1996. The differences were not large enough to be of practical significance. By contrast, the NSW report showed virtually no change in average 'aspects of numeracy' achievement at both Years 3 and 5 from 1994 to 1997, apart from an unusually high achievement by the Year 3 students in 1996. (No explanation for this anomalous result was offered in the report). The Queensland report showed increases in all three aspects of numeracy assessed (Number, Measurement and Space) from 1995 to 1996, and a generally smaller increase, or for some subgroups a decrease, from 1996 to 1997. For all groups the 1997 mean score was higher than the 1995 mean score. Results were analysed for several groups of students and all followed the same pattern except for indigenous students whose result in Measurement did not increase from 1995 to 1996.

What do statewide assessments and international studies indicate about equity issues?

Relationships with socioeconomic status

Educational achievement is typically correlated at about 0.3 with the socioeconomic or socio-educational background of the student (Ainley, Graetz, Long & Batten, 1995). The international IEA studies were no exception. In FIMS, correlations of the order of .22 were reported between achievement and status of father's occupation (Husen, 1967). In the Australian version of SIMS, correlations between these two variables by State ranged from .15 to .34, with a median value of .26 (Rosier, 1980). In

TIMSS, Australia was the only country that collected data on parents' occupations. The median correlation between total mathematics achievement and family occupational status, indicated by the higher of mother's and father's occupations, was .27 for the 9 year-olds and .30 for the 13 year-olds (Lokan, Ford & Greenwood, 1996; 1997).

The TIMSS assessment contained a mixture of multiple choice, short answer and extended response items, as well as some practical tasks involving work with apparatus (shapes, a calculator, plasticine, dice, etc.). Correlations were computed for several student background variables in relation to achievement on each type of task. It was expected that socioeconomic status would be more highly correlated with the extended response scores than with either multiple choice or short answer scores, but this did not turn out to be the case. Further, socioeconomic status was less important as a predictor of achievement on the practical tasks than it was on the written tests (Lokan, Adams & Doig, 1999).

These TIMSS results are encouraging from the point of view of equity in assessment, given that research exists which shows that some of the forms of 'authentic' assessment currently used extensively in schools, such as extended projects and portfolios of work, can lead to greater inequities for disadvantaged students. Two reasons for this are usually proposed: that the students may have fewer resources at home to draw on; and that they may be limited in the writing skills needed in most kinds of project work. There are also many influential writers who have expressed concern that wider ranges of assessment tasks may well lead to greater group differences in achievement – for example:

> It would be a mistake to assume that shifting from fixed-response standardized tests to performance-based assessments will obviate concerns about biases against racial/ethnic minorities or that such a shift would necessarily lead to equality of performance (Linn, Baker & Dunbar, 1991, pp. 17–18).

This message was reinforced in an ETS Research Report two years later (Zwick, Donoghue & Grima, 1993):

> Although the belief has been expressed that performance assessment provides a more equitable approach to testing than multiple-choice measures, some forms of performance assessment may in fact be more likely than conventional tests to tap construct-irrelevant factors (p. 3).

Instances where adding a performance component to a testing program led to larger mean differences among minority groups are cited in this ETS report.

Some of the strongest words of caution appeared in an article by Baker and O'Neill (1994). These authors indicated their belief that 'performance assessment is in for a rough time on the equity issue' (p. 24). All of these cautions need to be kept in mind when assessments are being designed, but the TIMSS results have demonstrated that

wider ranges of tasks seem to be no less equitable, in terms of scores gained, than more traditional tasks.

Gender differences

Until quite recently, it has been common to find studies reporting results that show males performing at a significantly higher level in mathematics than females. In FIMS, Husen (1967) reported that 'sex was related to mathematics achievement in almost all countries, the boys scoring higher than the girls at all levels' (p. 39). In Australia, such gender differences had disappeared by the time of SIMS, though they persisted in many other countries. By the time of TIMSS, significantly different performance by gender, always in favour of males, occurred in only six of 24 countries for the 9 year-olds, in only seven of 41 countries for the 13 year-olds, but in 11 of 16 countries in advanced mathematics at final year secondary level. Australia was one of a very small number of countries where no gender difference in mathematics performance at any of these levels was found (Lokan, Ford & Greenwood, 1996; 1997; Lokan & Greenwood, 1999).

Published results from statewide assessments of 'aspects of numeracy' in New South Wales show that females performed slightly better than males in 'Number' at Years 3 and 6, at about the same level in 'Measurement' but slightly lower in 'Space' (Masters, Lokan, Doig, Khoo, Lindsey, Robinson & Zammit, 1990). More recently, the public report of the BST results from 1997 discusses gender differences in literacy, but makes no mention of any such difference in numeracy (New South Wales Department of School Education, 1998). In Western Australia in 1992, females achieved at the same level as males in all areas of mathematics at Years 3 and 7, except for Space at Year 3 where the females performed better. Males significantly outperformed females in all areas *except* space – and algebra – at Year 10. In both of these areas there was no difference in performance between the Year 10 gender groups (Titmanis, Murphy, Cook, Brady & Brown, 1993). In 1996, Year 3 females again performed better than males in Space and, in Years 7 and 10, males performed significantly better than females in Measurement. In other strands at Year 10 there was no gender difference (van Wyke, 1998). The 1998 Queensland report shows that the 1997 performances of males and females were similar in Number and Measurement, while in Space the performance of males was slightly above that of females. The report of the 1997 Multilevel Assessment Program in the Northern Territory (NT Board of Studies, 1998) showed that males' performance at Year 6 and Year 4 was very slightly higher than females' in mathematics (three scale score points and one scale point in more than 700, respectively) but the differences were not significant.

The overall picture from these Australian results is that efforts to eliminate gender inequities in mathematics classrooms appear to have largely succeeded. Any differences found were small, and considered to be of no practical importance.

Differences for Indigenous and NESB students

A particularly pertinent and thorough review of the role language plays in mathematics is to be found in Dawe and Mulligan's (1997) chapter on language factors in mathematics learning and assessment. This review illustrates the major features of language in mathematics with examples drawn from the New South Wales Basic Skills Testing Program. The current emphasis on mathematics in context (not as 'naked numbers') means that inevitably there is a need for students to have a reasonable command of standard written English. Correlations between a measure of word knowledge (as a surrogate for verbal ability) and achievement in TIMSS mathematics of .60 and .47, for 9 year-olds and 13 year-olds, respectively, support this conclusion. In terms of the students' home language background, correlations with TIMSS mathematics achievement were .11 and .10, for 9 year-olds and 13 year-olds, respectively (the analogous correlations were a little higher for science).

States and Territories typically report results from their statewide assessment programs for students according to their non-English speaking background status. Generally, the achievement of students from non-English speaking backgrounds was lower than the achievement of those from English speaking backgrounds. An exception to this pattern was reported in Queensland, where the NESB students performed at a similar level to the ESB students in Number, Measurement and Space.

The achievement of Indigenous Australians, as reported in the State and Territory statewide assessments and also in TIMSS, lags considerably behind that of their non-Indigenous counterparts in all areas of mathematics. The report of the 1996 Western Australian MSE testing comments that 'the performance of Aboriginal and Torres Strait Islander students continues to be a concern. In general terms, their performance at each year level was almost a full outcome level lower than the performance of the rest of the population' (p. 6). The report of the NSW 1997 testing showed that the of Aboriginal and Torres Strait Islander group demonstrated more growth in numeracy from Year 3 to Year 5, using longitudinal data, than any other group. The report of the 1995 to 1997 testing in Queensland commented that the performance of Aboriginal and Torres Strait Islanders was 'more than extremely below' that of the rest of the population (Queensland School Curriculum Council, 1998, p. 18).

To illustrate a matter of concern in relation to issues of fairness in assessment, the performances of Indigenous and non-Indigenous students in TIMSS are shown by

sample level and type of item in Table 1. A similar pattern emerged in Australia for students whose first language was not English.

Table 1 Performance of Indigenous and non-Indigenous students in Mathematics, TIMSS

Item type	Average % correct			
	9 year-old Indigenous	9 year-old non-Indigenous	13 year-old Indigenous	13 year-old non-Indigenous
Multiple choice	49	59	45	62
Short answer	45	60	20	47
Extended response	38	49	8	31
	Average % omit			
Multiple choice	4	3	4	3
Short answer	14	9	43	14
Extended response	16	10	55	22

The results in Table 1 are a matter of concern in relation to the fairness of different types of assessment. In the analyses overall, it was shown that correlations between language background, gender, word knowledge and socioeconomic status and achievement were similar for each type of item. The numbers of Indigenous students per sample were quite small (about 400, in overall samples of several thousand). Effects for this relatively small group were masked by the results for the groups as wholes. The gaps in achievement were more marked for the older students, which has been noted in other studies (e.g. Masters et al, 1990). Of most concern are the percentages of Indigenous students who did not attempt to respond to the items which required answers to be written rather than selected. The results may suggest that some kinds of assessment are culturally specific and that other forms of assessment should be used to ascertain what these students know and can do.

Can we identify what we would have to do to achieve comparably in mathematics with the Asian countries?

While Australian performance in mathematics was generally above the international average, we are some distance away from the high-scoring Asian countries, as much as half to one standard deviation. At one third to half of a standard deviation per year level (as demonstrated by the differences between the upper and lower grades sampled in TIMSS), these differences indicate that the Asian students are one to two year levels ahead of our students in terms of the concepts and skills assessed in TIMSS. These differences in achievement agree with the analysis of expectations for achievement of outcomes reported by Lokan and Ainley (1998) discussed above.

The evidence suggests that we are inclined to expect less of our students than many other countries expect of theirs; we could consider challenging our students more by introducing some topics earlier, for example. The situation is much more complex than that, of course. The TIMSS data show that students from Asian countries generally spend more time in school, do more homework, have teachers who spend a lot of time marking homework and providing feedback to the students about it, than Australian students do or have. For example, 94 per cent of the Singapore teachers of 13 year-olds said they always mark and return their students' homework, compared with 38 per cent of the Australian teachers. Educators from the Asian countries attribute part of their success to structured programs of instruction, but say that the overwhelming factor is the value attached to educational qualifications by their societies and the consequently strong home support for their children's learning.

In the case of Japan, clues have emerged from the video study undertaken in Japan, Germany and the USA as part of TIMSS. In this study approximately 100 randomly selected mathematics classes were taped in each country. Japanese classes were found, on average, to have a substantially higher proportion of 'time on task' than classes in the other countries, higher proportions of challenging problem situations presented to the students and student involvement in whole-class attempts to suggest solutions to the problems. By contrast, almost half of the time in the American classes was taken up by reviewing homework or starting the next day's homework, and much of the rest of the time was spent in having the students work from worksheets. German classes were somewhere between these extremes. The mathematical content presented to the students was also analysed, and was found in Japan to be at least two years ahead of the content being studied in the USA (Stigler & Hiebert, 1997). The video study is being repeated in 1999, this time including Australia, which may give us some insight into differences in teaching practices between Australia and the other seven participating countries.

There is no clear message from TIMSS about the value of different types of assessment in mathematics. For example, teachers in Singapore and Hong Kong reported very high usage of assessment to diagnose students' learning problems, but in Japan and Korea considerably lower levels of usage for this purpose were reported. The same was true for assessments used to provide feedback to students. Australia, at 76 per cent, had a much higher reported incidence of assessment information being used to report to parents, in comparison with the Asian countries at between 9 per cent (Japan) and 39 per cent (Singapore). External standardised tests were not used at all for mathematics assessment in Singapore and had low reported use in Australia and Japan, but slightly higher use (by about 30 per cent of teachers) in Hong Kong and Korea.

What do we know about numeracy rather than mathematics?

Most of the research findings reported above pertain to assessments of school mathematics. To some extent the LSAY results are different, in that the tests used claim to be tests of numeracy. As such, they contain several tasks set in real-life contexts. Two other assessments that come closer to the ideal of a numeracy assessment are the 'mathematics and science literacy' (MSL) test used with Year 12 students in TIMSS and the planned 'mathematical literacy' test currently under development for the OECD Programme for International Student Assessment (PISA).

'Mathematics literacy' in TIMSS

The TIMSS MSL test was designed to 'provide information about how prepared all the school leavers in each country are to apply their knowledge in mathematics and science to meet the challenges of life beyond school' (Mullis et al., 1998, p. 31). On the mathematics part of this test, Australian students were outperformed only by students from the Netherlands, were at the same level as students from eleven other countries, and performed better than students from eight countries. The test assessed a mixture of concepts thought to be from a core that all students by Year 10 should be expected to know, plus a range of reasoning tasks set in every day circumstances (e.g. using several pieces of information to work out the rent on two office spaces and to decide which was the better value for money). The high achieving Asian countries at lower secondary level in TIMSS did not participate in the Year 12 testing.

'Mathematical literacy' in PISA

The planned PISA mathematics assessment, which will be administered to random samples of 15 year-old students in about 30 countries in the Year 2000, comes a good deal closer to targeting numeracy rather than school mathematics. It is a stated objective of PISA that the assessments will not be tied to narrowly conceived 'international core curricula'. The PISA definition of mathematical literacy is:

> an individual's ability, in dealing with the world, to identify, to understand, to engage in and to make well-founded judgements about the role that mathematics plays, as needed for that individual's current and future life as a constructive, concerned, and reflective citizen (de Lange, 1999, p. 1).

Tests containing a combination of important concepts and applications of those concepts are currently being prepared. More use is made of open-ended items than in TIMSS, in order to tap a wider range of skills. All tasks are situated within personal, educational or community settings.

For the year 2000, when mathematics is a minor domain in PISA and has only one hour of testing time, items from most school mathematics content areas have been developed within the two major areas ('big ideas') of 'change and growth' and 'space and shape'. Three main clusters of competencies have been defined, containing a wider range of competencies than one would expect to be assessed in a school mathematics test. These include the following skills: mathematical thinking; mathematical argumentation; mathematical modelling; problem solving and posing; representation; communication; decoding and interpretation of formal language; solving equations; and knowing about and being able to use a variety of aids and tools to assist mathematical activity.

Both cross-curricular and extra-curricular aspects of mathematics are recognised as components of mathematical literacy in PISA.

'Aspects of numeracy' in statewide assessments

Some statewide assessments include tests referred to as 'mathematics'. In those claiming to assess 'numeracy' or 'aspects of numeracy', predominantly what is being assessed are aspects of school mathematics. The draft numeracy benchmarks also focus on only a part of what can be learned within school mathematics. There are good reasons for these limitations, in that 'numeracy', as defined at the beginning of this paper and also in the discussion of PISA, needs to be assessed in a context much wider than 'school'. Some compromises are suggested later in the paper.

Influence of the National Plan

The National Literacy and Numeracy Plan for Schools advocates

> comprehensive assessment of all students by their teachers as early as possible in the first years of schooling with the purpose of adequately addressing their numeracy and literacy needs and identifying those students at risk of not making adequate progress towards the national literacy and numeracy goals.

An earlier section of the paper, and Table 2 in the appendix, describe the assessment initiatives currently in place or under development to assist in identifying 'at risk' students in numeracy. While some of these, such as Queensland's Year 2 Diagnostic Net, had been introduced well in advance of the National Plan, others have been commenced more recently. The extent of recognition that intervention at an early stage is desirable, in an attempt to prevent later problems, seems likely to have been enhanced as a result of awareness of the National Plan, though this cannot be said confidently without a survey of key players. Early intervention presupposes early assessment.

A parallel move to assess children early in their schooling has recently occurred in the UK, where 'tests' for school beginners were introduced on a national basis at the beginning of September 1998. The baseline 'tests' are administered by teachers and are designed to enable teachers to identify students' strengths and weaknesses so that the students can be taught according to their needs. The introduction of the tests was accompanied by British newspaper headlines that could have appeared anywhere in Australia: 'SCHOOL FAILURES AT FOUR' (*The Express*, 4/9/98) and 'INFANTS ARE LIKELY TO FAIL FIRST SCHOOL TEST' (*Daily Telegraph*, 4/9/98).

What kind of assessment might be most desirable?

The British experience referred to in the previous paragraph, plus British and US experiences with large-scale assessment involving practical tasks and other forms of performance assessment, together with all the activity currently being undertaken in Australia, gives rise to the question of the kind of numeracy assessment that might be most beneficial. To assess numeracy as defined at the beginning of the paper, a wide range of types of tasks would need to be devised and several measurement methods used to collect data on them. The purpose of the assessment would guide the range of tasks to be developed. As we know from others' experiences, some methods can be expected to be expensive and also require a great deal of teachers' time. Defining 'beneficial' may turn out to be a balancing act between what is manageable in terms of resources and what is indicated as the best assessment method for a given purpose.

It is difficult to specify exactly what might be considered to be 'best practice' in numeracy assessment. 'Best practice' is currently a popular term, but what it might consist of is volatile, changing over time and across diverse cultures. Good assessment will differ according to the age of the students and the reasons why the assessment is being undertaken. This ties back to the main principles of assessment formulated by Harlen (cited above): to be good, an assessment program must enhance learning and must cover the full range of curriculum goals.

Given the emphasis on successful performance of everyday tasks requiring understanding and manipulation of mathematical concepts, assessment of numeracy skills lends itself to being undertaken by classroom teachers. For such assessment to be of most value, compendiums of structured tasks would probably need to be provided for teachers to use. Teachers would benefit from training in assessment procedures, such as in the use of results from 'Mathematics Intervention' assessments, or the workshops run a few times each year in relation to the Assessment Resource Kit materials cited earlier. There is also an urgent need for

more training in assessment issues and strategies to be included in preservice and inservice programs.

Designing assessment tasks will not be easy for many of the objectives of a program intended to foster students' numeracy development. Some objectives will be able to be assessed through paper and pencil tasks. Others will require observation of student behaviour in real-life contexts or judgements about products or performances. Given the prominence assigned to the development of mental processing skills as part of being numerate, techniques for assessing these skills in a standard way need to be researched.

For an assessment that is being used to diagnose problems or potential problems, it is important that there are well-founded underlying developmental continua, on which stages of the assessment are based. These will enable progress to be charted so that the results of interventions can be monitored. It will be highly desirable if assessments used more generally also have this capability.

That a reasonably large-scale assessment can incorporate a broad range of tasks and successfully involve teachers in its implementation was recently demonstrated in the National School English Literacy Survey (NSELS) (Management Committee, 1997). Further support for the value of an assessment program involving a wide range of types of assessment tasks is offered by the recent joint award to the Western Australian Monitoring Standards in Education program and ACER. The award was conferred on the winners by the Assessment Research Centre at the University of Melbourne and presented by the then Minister for Schools, Dr David Kemp, at the inaugural National Assessment Awards for exemplary assessment practices. The materials developed were described as exhibiting a genuinely innovative approach which would enable teachers to make many of the decisions about the assessment, including ascertaining links to the Student Outcome Statements.

Lessons from literacy

Literacy education has been on the public agenda for a longer time than has numeracy education. Literacy is generally seen as much broader than numeracy, which may be the reason why it has received more attention. Certainly, literacy skills underpin all learning areas, but numeracy skills are also needed much more widely than just in mathematics. The cause of numeracy education would benefit if teachers could be brought to realise the relevance of numeracy skills to other areas of the curriculum. A review or mapping study of numeracy applications in other subjects could be useful in this regard.

As in literacy education, it is recognised that no one numeracy intervention program for students with learning difficulties in numeracy will provide the solution. There is

need for a numeracy research program similar to the DETYA funded Children's Literacy Research Program that was well supported over much of the 1990s. Some significant research has already begun in the ACER Longitudinal Literacy and Numeracy Study, in which the literacy and numeracy development of a random sample of 10 students from each of 100 schools is being studied over a period of seven years. It will be beneficial if other research into numeracy development could also be undertaken, as well as evaluative research on the effects of the various identification and intervention programs for 'at risk' students that have recently been put in place. In addition, as always, numeracy teaching and assessment would benefit from the enhanced provision of opportunities for teacher professional development.

References

Ainley, J., Graetz, B., Long, M. & Batten, M. 1995, *Socioeconomic Status and School Education*, Australian Government Publishing Service, Canberra, ACT.

American Guidance Service 1997, *KeyMath – Revised: A Diagnostic Inventory of Essential Mathematics*, Minneapolis, Author.

Askew, M., Brown, M., Rhodes, V., Johnson, D. & William, D. 1997, *Effective Teachers of Numeracy: Final Report*, King's College, London, UK.

Australian Association of Mathematics Teachers (AAMT) Inc. and Education Department of Western Australia 1997, *Numeracy = everyone's business, Report of the numeracy Education Strategy Development Conference*, AAMT Inc., Adelaide.

Australian Council for Educational Research (ACER) 1991, *Diagnostic Mathematics Profiles With DIAMAP*, ACER, Hawthorn, Vic.

Australian Council for Educational Research (ACER) 1997, *Orchid Series: Aspects of Numeracy*, ACER, Camberwell, Vic.

Australian Council for Educational Research (ACER) 1997, *Progressive Achievement Tests in Mathematics* revised, ACER, Camberwell, Vic.

Baker, E. L. & O'Neill, H. F. (Jr) 1994, 'Performance assessment and equity: A view from the USA', *Assessment in Education*, *1* (1), pp. 11–26.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L. & Smith, T. A. 1996, *Mathematics and Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*, Boston College, Chestnut Hill, MA.

Booker, G. 1994, *Booker Profiles in Mathematics*, ACER, Hawthorn, Vic.

Coley, R. J. & Goertz, M. E. 1990, *Educational Standards in the 50 States: 1990*, Research Report RR-90-15, Educational Testing Service, Princeton, NJ.

Dawe, L. & Mulligan, J. 1997, 'Classroom views of language in mathematics', in B. Doig & J. Lokan (eds), *Learning from Children: Mathematics from a Classroom Perspective*, ACER Research Monograph No. 52, ACER, Melbourne.

De Lange, J. 1999, *The PISA Framework for Mathematic,* Organisation for Economic Co-operation and Development, Paris, in press.

Doig, B. 1991, *Group Review of Algebra Topics*, ACER, Hawthorn, Vic.

Educational Testing Service (ETS) 1995, *Learning for Tomorrow*, Annual Report, Princeton, NJ.

Gibson, S. J., Doig, B. A. & Hunting, R. P. 1993, 'Inside their heads — the clinical interview in the classroom', in J. Mousley & M. Rice (eds), *Mathematics: Of primary importance,* Mathematical Association of Victoria, Brunswick, pp. 30–35.

Glidden, H. 1998, *Making Standards Matter: 1998*, American Federation of Teachers Washington, DC.

Harlen, W., Gipps, C., Broadfoot, P. & Nuttall, D. 1994, 'Assessment and the improvement of education', *The Curriculum Journal*, *3* (3), 215–230.

Hunting, R. P. & Doig, B. A. 1992, 'Development of a clinical tool for initial assessment of a student's mathematics learning', in M. Stephens & J. Izard (eds), *Reshaping Assessment Practices: Assessment in the Mathematical Sciences Under Challenge,* ACER, Melbourne, pp. 201–217.

Husen, T. (ed.) 1967, *International Study of Achievement in Mathematics: A Comparison of Twelve Countries, Volume II,* Almqvist & Wiksell, Stockholm (and Wiley, New York).

Linn, R., Baker, E. & Dunbar, S. 1991, 'Complex, performance-based assessment: Expectations and validation criteria', *Educational Researcher*, *20* (8), pp. 15–21.

Lokan, J., Ford, P. & Greenwood, L. 1996, *Maths & Science On the Line: Australian Junior Secondary Students' Performance in the Third International Mathematics and Science Study*. TIMSS Australia Monograph No. 1, ACER, Melbourne.

Lokan, J., Ford, P. & Greenwood, L. 1997, *Maths & Science On the Line: Australian Middle Primary Students' Performance in the Third International Mathematics and Science Study*. TIMSS Australia Monograph No. 2, ACER, Melbourne.

Lokan, J. & Greenwood, L. 1998, *Maths & Science On the Line: Australian Year 12 Students' Performance in the Third International Mathematics and Science Study*. TIMSS Australia Monograph No. 3, ACER, Melbourne, in press.

Lokan, J. & Ainley, J. 1998, 'The Third International Mathematics and Science Study: implications for the development of numeracy benchmarks', *Unicorn*, *24* (2), 97–109.

Lokan, J., Adams, R. & Doig, B. 1999, 'Broadening assessment: improving fairness? Some examples from school science', *Assessment in Education 6* (1), 83–89.

Marks, G. & Ainley, J. 1997, *Reading comprehension and numeracy among junior secondary school students in Australia, LSAY Research Report No. 3*, ACER mimeo, Melbourne.

Masters, G., Lokan, J., Doig, B., Khoo, S.-T., Lindsey, J., Robinson, L. & Zammit, S. 1990, *Profiles of Learning: The Basic Skills Testing Program in New South Wales 1989*, ACER, Melbourne.

Masters, G. & Forster, M. 1996, *Developmental Assessment*, Assessment Resource Kit (ARK) series, ACER, Melbourne.

Management Committee for the National School English Literacy Survey 1997, *Mapping Literacy Achievement: Results of the 1996 National School English Literacy Survey*, Department of Employment, Education, Training and Youth Affairs, Canberra.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L. & Smith, T. A. 1997, *Mathematics and Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study*, Boston College, Chestnut Hill, MA.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L. & Smith, T. A. 1998, *Mathematics and Science Achievement in the Final Year of Secondary School: IEA's Third International Mathematics and Science Study*, Boston College, Chestnut Hill, MA.

New South Wales Department of School Education 1998, *Building the Foundations: The 1997 Report on the Basic Skills Tests*, Sydney, Author.

New Zealand Council for Educational Research (NZCER) 1990, *Mathematics Topic Pre-tests*, Wellington, NZ.

Northern Territory Board of Studies 1998, *Results of the 1997 Multilevel Assessment Program in Urban and Non-Urban Schools*, Report No. 2, Darwin.

Office of Educational Research and Improvement 1983, *A Nation at Risk*, US Department of Education, Washington, DC.

Palmer, D., Kays, M., Smith, A. & Doig, B. 1994, *Stop! Look & Lesson*, ACER, Hawthorn, Vic.

Pearn C. A., Merrifield, M., Mihalic, H., & Hunting, R. P. 1994, *Initial clinical assessment procedure, Mathematics — Level AA Years 1 & 2*, La Trobe University, Bundoora.

Pearn, C. A., & Merrifield, M. 1996, 'Strategies for classroom teachers: A lesson from Mathematics Intervention', in H. Forgasz, A. Jones, G. Leder, J. Lynch, K. Maguire & C. Pearn (eds), *Mathematics: Making connections*, Mathematical Association of Victoria, Brunswick.

Petrie, H. G. 1987, Introduction to 'evaluation and testing', *Educational Policy*, *1*, 175–180.

Queensland School Curriculum Council 1998, *Statewide performance of students in aspects of literacy and numeracy in Queensland 1995, 1996 and 1997*, Brisbane.

Recht, E., Forster, M. & Masters, G. 1998, *Developmental Assessment Resource for Teachers – Mathematics*, ACER, Camberwell, Vic.

Resnick, L. B. & Resnick, D. P. 1992, 'Assessing the thinking curriculum: New tools for educational reform', in B. G. Gophered & M. C. O'Connor (eds), *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction*, Kluwer Academic Publishers, Boston, pp. 37–75.

Rosier, M. J. 1980, *Changes in Secondary School Mathematics in Australia 1964–1978*, ACER Research Monograph No. 8, ACER, Melbourne.

Schleiger, H. E. 1993, *Diagnostic Mathematical Tasks*, revised by J. Gough, Deakin University Press, Geelong.

Stacey, K. 1997, 'Implications of the TIMSS for mathematics education in Australia', presented at ACER National Conference, Raising Standards in Mathematics and Science: Insights from TIMSS, Melbourne.

Stacey, K., Groves, S., Bourke, S. & Doig, B. 1993, *Profiles of Problem Solving*, ACER, Hawthorn, Vic.

Steffe, L. P., von Glaserfeld, E., Richards, J. & Cobb, P. 1983, *Children's counting types: Philosophy, theory and application*, Praeger, New York.

Steffe, L. P., Cobb, P. & von Glaserfeld, E. 1988, *Construction of arithmetical meanings and strategies*, Springer-Verlag, New York.

Stigler, J. & Hiebert, J. 1997, 'Understanding and improving classroom mathematics instruction', *Phi Delta Kappan*, *79* (1), 14–21.

Titmanis, P., Murphy, F., Cook, J., Brady, K. & Brown, M. 1993, *Profiles of Student Achievement: English & Mathematics in Western Australian Government Schools*, Ministry of Education, Perth.

van Wyke, J. 1998, *Student Achievement in Mathematics in Western Australian Government Schools*, *Monitoring Standards in Education '96*, Education Department of Western Australia, Perth.

Wainer, H. 1997, 'Improving tabular displays, with NAEP tables as examples and inspirations', *Journal of Educational and Behavioral Statistics*, *22* (1), 1–30.

Wiggins, G. 1989, 'A true test: Toward more authentic and equitable assessment', *Phi Delta Kappan*, *70* (9), 703–713.

Williams, R. & Dixon, H. 1998, 'School entry assessment: praiseworthy or perilous?', paper presented at the Seventh Australia and New Zealand Conference on the First Years of School, Canberra, January.

Wright, R. J. 1991, 'The role of counting in the numerical development of young children', *Australian Journal of Early Childhood*, *16* (2), 43–48.

Wright, R. J., Stanger, G., Cowper, M. & Dyson, R. 1996, 'First-graders' progress in an experimental Mathematics Recovery program', in J. Mulligan & M. Mitchelmore (eds), *Children's number learning,* AAMT Inc., Adelaide, pp. 55–72.

Zwick, R., Donoghue, J. & Grima, A. 1993, *Assessing Differential Item Functioning in Performance Tests*, Research Report RR-93-14, Educational Testing Service, Princeton, NJ.

Appendix[1]

Table 1: Continuing Programs of SYSTEM-MONITORING in Numeracy or Mathematics (State and Territory Education Departments)

State or Territory	**Program**	**When program began**	**Age/year levels assessed**	**Program operation**	**Groups (sub-groups) reported on**	**Program funding**
ACT	Full population assessment introduced into year 5 in 1998. ACER's 'At the Zoo', DART package used. (Planning underway for year 3, 7 and 9 assessment in 1999)	Started in 1998	Year 5	Apart from providing individual, school and system level achievement data by national profile strand the assessment is designed to give benchmarking data for national reporting.	Full population, gender, ESL, Aboriginal and Torres Strait Islanders – data is confidential – no school comparisons are made.	ACT government initiative.
NSW	Basic Skills Testing (Literacy and Numeracy) for all NSW school students.	Operating since 1989.	Students in year 3 aged 7–9 years. Students in year 5 aged 9–11 years.	A new test is developed yearly which in the case of numeracy tests: Number – how well students can count, add, subtract, multiply and divide. Students also answer questions about fractions, decimals and money. Measurement – skills involved in estimating and measuring length, area, volume, mass, temperature and time. Space – use of graphs, shapes, position and direction to answer questions. Reports are sent to schools and parents. Telling parents what their child can do, how their child's results compare with the rest of the state, and describes the numeracy skills tested. School reports tell the teachers what students can and cannot do, gives teachers information about various student groups and the whole of NSW, helps teachers to identify groups of students who might need help, and helps teachers make decisions about school programs.	Boys and Girls. Students aged ≤ 7 yrs (yr 3) Students aged 8 yrs (yr 3) Students aged ≥ 9 yrs (yr 3) Students aged ≤ 9 years (yr 5) Students aged 10 yrs (yr 5) Students aged ≥ 11 years (yr 5) Aboriginal and Torres Strait Islanders NESB NESB – Students who have lived in Australia for 4 yrs or less and never or only sometimes	Funding from both state and federal funds.

[1] This paper was prepared in 1998/9 and some information may not reflect more recent developments.

					speak English at home.	
QLD	To monitor a student's progress in Numeracy early childhood teachers utilise the Number Developmental Continuum. In the later years of schooling a sample test occurs in year 3, and a census test in year 5 with the introduction of a census year 7 test occurring in August 1999	Year 2 Diagnostic Net since 1995. Sample test year 3 and census test year 5 since 1998 with the year 5 test replacing the year 6 test which was introduced in 1995.	Monitoring of student's progress begins upon the student's entry into the compulsory years of schooling and continues with various degrees of formality throughout the compulsory years of schooling.	The monitoring on the Developmental Continuum occurs in the course of daily teaching and observations are validated during a specific time frame. The validation tasks are designed by the Queensland School Curriculum Council and teachers meet to moderate work samples. The sample testing occurs simultaneously throughout the state at specified times and dates.	Not stated.	Various monitoring devices are funded by the state.
SA	Basic Skills Test	Running for 3 years	Years 3 and 5	A statewide test at a particular point in time for all years 3 and 5 students across the state. They are standardised tests which measure aspects of Literacy and Numeracy. The tests are marked centrally and reports for parents are generated from the results (c.f. NSW BST).		State funding
WA	Sample testing	Testing started in 1990. Tested Mathematics in 1992, 1996 and 1998.	Years 3, 7, and 10. Year 5 testing introduced in 1998.	Random sample drawn to test all strands except 'working mathematically' . Trialling occurs in March/April. Sample testing takes place in the second week of September. As of 1999 program becomes population testing in years 3 and 5. School materials. Released so that schools can monitor their own performance.	Males, females, NESB and Aboriginal and Torres Strait Islanders.	Funded by the state
TAS	Cohort testing given on a 'cyclic basis'	Since 1975	From 1997 testing has been based on grade cohorts (not age group cohorts). From 2000 onward, at 2-year intervals, the intention is to test students in grades 3, 5, 7, 9.	Department now uses consultants (ACER in 1996, ARC at Melbourne University from 1997). Items written locally, trialled interstate. Schools given results on disk. Testing usually done mid year; full cohort.	Sex, Aboriginal and Torres Strait Islanders category socio-economic status, school district, rurality / distance, language background other than English, and age.	Funded by the state.

NT	1. Multi-level Assessment Program (MAP)	Multi-level Assessment Program (MAP) – started early 1980s.	Multi-level Assessment Program (MAP) – Transition – Yr 10. Multi-level Assessment Program (MAP) – Years 3 and 5 in urban and age equivalents in non-urban areas.	Multi-level mathematics tests monitor achievements in numeracy of students in years 3 and 5 who will then be assessed against the benchmarks.	NT Department of Education releases results to the schools. All schools participate (urban, rural, remote).	Funded by the NT Department of Education.
	2. Years 3 and 5 Numeracy Benchmarks have been incorporated into the Multi-level Assessment Program (MAP).	Benchmarks were incorporated for the first time in the 1998 testing program	Years 3 and 5	Using the Multi-level Assessment Program (MAP), data are gathered from year 3 and 5 students in urban schools and age equivalent in non-urban schools those with predominantly Aboriginal enrolment. Data are aggregated by way of gender, ESL, Aboriginal/non-Aboriginal to assist with the identification of the number of children 'at risk'.		
	3. Junior Secondary School Certificate (JSSC)	Junior Secondary School Certificate (JSSC) – started late 1980s.	Junior Secondary School Certificate (JSSC) – Year 10.			

(Catholic Education Offices)

State or Territory	**Program**	**When did program begin**	**Age/year levels programs operate**	**How does program operate**	**Which groups (sub groups) monitoring report**	**Other comments**
ACT/ NSW	Not at this stage.					
SA	In the process of researching and undertaking consultation prior to the implementation of an appropriate system methodology.					
TAS	Cohort testing given on a 'cyclic basis' (Information provided by TAS State Department)	Since 1975	From 1997 testing has been based on grade cohorts (not age group cohorts). From 2000 onwards, at 2-year intervals, we intend to test students in grades 3, 5, 7, 9.	Department now uses consultants (ACER in 1996, ARC at Melbourne University from 1997). Items written locally, trialled interstate. Schools given results on disk. Testing usually done mid year; full cohort).		From 1997, (and from c.1975 to 1992), the Catholic Education Office in Tasmania has used the same tests for their sector.

Table 2: Programs and Initiatives in Place for IDENTIFYING 'AT RISK' Numeracy or Mathematics Learners
(State and Territory Education Departments)

State or Territory	**Program**	**When program began**	**Age/year levels assessed**	**Program operation**	**Program funding**	**Other comments**
ACT	ACT Assessment Program and special assistance.	K–10 LA program – 1994 Funding for Learning Assistance teachers in high schools and Resource Teachers in primary schools has been provided to ACT schools since 1985.	Kindergarten to Year 10	Bottom 20 per cent of students are identified through the ACT Assessment Program administered to students in years 3, 5, 7 , and 9. Program provision includes: • Team teaching • Alternate (LA) classes in high schools • Resource support for teachers and students in the mainstream classroom • Small group withdrawal • Provision of information and support to parents • Provision of support to class teachers including in-service meetings, diagnostic assessments, and teaching resources.	Funded by the ACT Government within the total schools education budget.	
NSW	Count Me In Too Project Additional support provided by formation of a broad screening process for classroom teachers in the *Early School Assessment project.*	Early School Assessment project started in 1997. Earlier work had involved teachers in the use of the early Learning Profiles.	Progressive support across years K–3. Draft material has been developed and trialled for school entry. End of Kindergarten and end of year 2 material is in development stage.	Schools are involved in the development and trialling of materials. Professional development focuses on the identification of students at risk and the modification of teaching strategies to address the identified needs.	ESAP funded as part of the national Literacy and Numeracy Plan — Professional Development Projects.	The project provides teachers with assessment strategies for use in the first year of schooling and intervention strategies to improve numeracy performance.

QLD	The Year 2 Diagnostic Net	The Year 2 Diagnostic Net – started in 1995.	The Year 2 Diagnostic Net operates in the first three years of compulsory schooling.	The Year 2 Diagnostic Net involves a four step process in which Years 1, 2 and 3 teachers: • observe and map children's progress using developmental continua for literacy and numeracy; • involve identified Year 2 children in specifically designed assessment tasks and identify children who require intervention; • provide support to children requiring additional assistance; • report to parents about their child's development in literacy and numeracy. To monitor progress teachers use commonly agreed key indicators of literacy and numeracy development. The indicators are grouped into phases of development.	The year 2 Diagnostic net is state funded with schools receiving funding according to the number of children identified as needing extra assistance.	
	Appraisement	Appraisement – to begin in 1999.	Appraisement operates across all the year levels.	**A**ppraisement process involves: • the collection of information about a child's educational needs; • completion of mandatory appraisement tasks in literacy and numeracy; • completion of a Support Plan which builds on strengths and meets identified needs of each student.	Appraisement is also state funded	

SA	The department is introducing a school entry assessment program which will better enable teachers to provide programs which will build on what individual children know and can do when they enter formal schooling. The program consists of assessment tools for literacy and numeracy and will incorporate information from the pre-school and parents. Children will be assessed within the first term of starting school. The program will direct teachers to resources which will support their teaching. While the program is aimed at all children it will provide assistance to those children at risk. The department also administers Basic Skills Tests in Aspects of Literacy and Numeracy for students in years 3 and 5.	The School Entry Assessment Program is in development and will be introduced into schools in 1999. The Basic Skills Tests program has been running for 3 years.		The School Entry Assessment Program is based on teacher judgement in assessing aspects of literacy and numeracy. Teachers are encouraged to observe children's behaviour and learning in the normal classroom setting, to look for particular indicators for each stage and to use assessment tasks if necessary to identify a child's strengths and areas for development. Teachers will be able to gather data to make a class profile and that information can be aggregated to give a school profile. The Basic Skills Tests are done as a statewide test at a particular point in time for all year 3 and 5 students across the state. They are standardised tests which measure aspects of literacy and numeracy. The tests are marked centrally and reports for parents are generated from the results.	Funded by the State through the provision of grants to schools and centralised support.	
WA	The Curriculum Improvement Program, and Making A Difference – Students at Educational Risk.	Curriculum Improvement program was started in 1998, and is expected to be fully implemented by 2004 Students at Educational Risk stared in 1998.	K – Year 10	Students at Educational Risk strategies enable teachers to develop a profile of students achievements. The strategies will enable, amongst other things 'at risk' numeracy/mathematics learners to be identified.	Funded by the state.	A new initiative, the First Steps in Maths Project will commence in 1999. The professional development which will address students 'at risk' in numeracy learning.

VIC	Additional Assistance and a Structured Classroom Program are major components of the Victorian Early Years Numeracy program.	The component of Additional Assistance and Structured Classroom Program is under development	Prep – Year 4.	The Additional Assistance pathway provides a structured approach for those students who require additional support to succeed. The pathway includes: identification, home-school support group and individual learning improvement plans and review. The Structured Classroom Program identifies developmental stages and approaches for ongoing assessment and monitoring to support teachers to identify students 'at risk'.	State funding.	

(Catholic Education Offices)

State or Territory	**Program**	**When did program begin**	**Age/year levels programs operate**	**How does program operate**	**Program Funding**	**Other comments**
ACT/ NSW	Identifying 'at risk' students is done in the context of scheduled applications for Commonwealth funds (Literacy program). Occurs only in a general sense.					
NT	Identifying 'at risk' students Documents: Assessment in the early years of school and strengthening the early years: Improving outcomes.	Identifying 'At risk' students has been ongoing as part of school practice. Documents previous to these were used and were part of common practice in schools.	Identifying 'at risk' students at each year level but particularly in the early years of schooling.	Identifying 'at risk' students – identification by the teacher, through school based support. Some students identified through the Multi-level Assessment Program (MAP). Secondary schools have more formalised method of assessment for placement in their streamed maths program. Documents – assist early childhood teachers to access numeracy skills in appropriate ways so that they are then able to report accurately and implement intervention strategies if necessary. Allows teachers to renew their knowledge of child development and learning.	School based funding from both territory and Commonwealth. sources.	

WA	Learning Difficulties Professional Development Program for identifying students for special programs in maths.	1997 –1998	K – Year 12	Schools identify 'at risk children', set outcomes, develop strategies to achieve outcomes – monitor program, and evaluate program. Professional development assists schools to develop a team approach. Building collaborative teams of teachers to promote a whole school approach to numeracy. Sample cohorts are chosen and tested in both numeracy and literacy.	State funded.	Not all schools apply for numeracy money, however, there is a growing number who see the importance of concentrating on numeracy.
	Retention and participation project for students at Education Risk.	Began Aug 1998.	Years 8 – 12.		System funded.	Program is in its developmental stage and will be presented to selected schools in November 1998.
	Sample testing of selected schools as part of IESIP program	Since 1996.	Year 3, 7, and 10.		Commonwealth funded	
QLD	Identification of students.	Not stated.	Not stated.	Students are identified as needing additional support through a range of intervention strategies. Strategies aimed at developing learning partnerships with parents and care givers. Focus on professional development for teachers with an emphasis on models of teaching and learning which acknowledge numeracy as a social practice. Assessment is an integral part of the planning, teaching and learning cycle. Schools are encouraged to develop a whole school numeracy plan.		Individual dioceses have developed some creative and innovative approaches to the provision of numeracy and numeracy support.

VIC	No specific name	Not stated	Prep – Year 6.	Children with learning difficulties are referred to the Curriculum Advisers who then test the children through the use of Keymath or psychological assessments. Curriculum advisers or special education advisers test the children and then provide support for the classroom teacher through advice and strategies for the classroom teacher. Sometimes programs are instigated for the children with special needs	A Catholic Education Office initiative.	