Reading Comprehension and Numeracy among Junior Secondary School Students in Australia

LSAY RESEARCH REPORT NO 3

Gary N. Marks

John Ainley

March 1997

ACER

CONTENTS

TABLES

ACKNOWLEDGMENTS

This report rests on a considerable volume of work done by other people. It owes much, not only to the body of knowledge and technical development that underpins almost every piece of research, but also to the detailed work that has resulted in the data on which these analyses were based. This includes the national studies of literacy and numeracy from 1975 and 1980 and the later *Australian Youth Survey*, *Youth in Transition*, and *Longitudinal Surveys of Australian Youth* projects.

The success of these studies was very much dependent on the school principals, teachers and students in schools, officials in State and Commonwealth education authorities, and the research and support staff who worked on the studies. Specific thanks are due to state and territory education Ministers and the education system Chief Executive Officers for their assistance. The cooperation of the governing bodies and directors of Catholic and independent Schools is also gratefully acknowledged. The Commonwealth Department of Employment, Education, Training and Youth Affairs has supported these programs in both a conceptual and financial sense.

Several individuals deserve specific acknowledgment. Sid Bourke and John Keeves directed the 1975 and 1980 studies. They and the staff from those projects provided the basis on which subsequent studies were built. Trevor Williams initiated the idea of a longitudinal survey of transitions between school, further education and work, based on the original national testing program. He and Michael Long ensured that those surveys continued and that new cohorts were added at appropriate intervals (including the use of tests of achievement similar to those from 1975). Lyn Robinson has played an important and invaluable role in many aspects of the *Youth in Transition* and *Longitudinal Surveys of Australian Youth* projects, and helped in the preparation of this report. Geoff Parkinson contributed through his pivotal role in the *Australian Youth Survey*, as an important source of knowledge on the practice of longitudinal surveys, and in his administrative role with respect to the present program of longitudinal surveys. Ray Adams and Margaret Wu contributed their expertise in Rasch modelling in the construction of common scales that allowed comparisons over time.

Each one of these people had a vital role to play in the development of the data on which this report is based and there are many more but space precludes mentioning each by name. However, those who contributed are in no way responsible for the analyses and interpretations contained in this report.

EXECUTIVE SUMMARY

This report examines the performance, and levels of mastery, on tests of reading comprehension and numeracy of Australian students in junior secondary school between 1975 and 1995. Data for the analyses reported were drawn from two national monitoring studies and other studies of representative samples of junior secondary school students designed to monitor the progress of young people through school into further education, training and work. In each of the studies students completed tests that are similar to those conducted in the Australian states and territories and overseas to monitor student performance. The tests focussed on reading skills (which correlate highly with other aspects of literacy) and numeracy.

Although both the reading and numeracy tests were not identical over time they included a substantial number of common items. The existence of common items allowed the construction of single measures of achievement in reading and numeracy. In addition to reporting average levels of achievement on these two scales, the concept of mastery developed for the 1975 study was applied at later points in time. Mastery was defined as competence in the basic skills necessary for active participation in society. Non-mastery on the reading tests does not correspond to 'illiteracy' in the conventional meaning of the term.

The analyses reported include: comparisons of the percentage of students correctly answering common items; the mean levels of achievement and percentage of students attaining mastery at different times; and comparisons over time of achievement and in some instances mastery, by gender, ethnicity, and social background. In addition, the report discusses the results obtained from multilevel multivariate analyses of both school-level and individual-level factors.

Over the period from 1975 to 1995 there was little systematic change in the performance on the common items in the reading tests. There were two counteracting tendencies. Performance on items referring to newspapers articles improved slightly, but declined on items dealing with more difficult textual passages. With the numeracy tests there were also counteracting tendencies; performance declined marginally on computational items but improved on conceptual items. Analysis of the common achievement scales showed no change in the mean levels of either reading comprehension or numeracy. There was no substantial change in the proportion of 14 year-old students achieving mastery in reading since 1975 but the percentage attaining mastery in numeracy appears to have increased.

Associations between achievement and social and educational factors are detailed in the report. Some of the findings include: a decline in the proportion of boys attaining mastery in reading; a decline in differences in achievement relating to parental occupation for both reading and numeracy; little net change in the performance of students whose home language is not English; and an improvement in the performance of Aboriginal and Torres Strait Islander students in numeracy.

Reading Comprehension and Numeracy among Junior Secondary School Students in Australia

INTRODUCTION

In 1977 the authors of a report on the first national study of literacy and numeracy noted that there was community concern about declining educational standards (Bourke & Keeves, 1977: 5; Keeves et al., 1978). Similar concerns were noted in the report from the second national study conducted five years later and again in 1988 with the release of a report on literacy and numeracy among Victorian Students (Bourke et al., 1981: 3; McGaw et al., 1989: 2). In 1996 that concern has not disappeared and it is not uncommon to read or hear comments by parents, employers and teachers, about the poor word and number skills of young people. If word and number skills have declined, this would represent a serious problem, given the context of decreasing demand for unskilled labour and the emergence of a post-industrial information-based society.

Although there is a number of state-based programs in Australia for monitoring achievement (especially in the primary school years), there is nothing that can provide an integrated national perspective (Lokan & Ford, 1994). In addition, there is very little information on patterns in achievement for junior secondary school students even at state level. Their performance at this stage in their school careers has an important bearing on their pathways through senior school, post-secondary education, and the world of work. As a nation we do not know the proportion of students in junior high school who have not mastered the 'basic skills' which enable citizens to effectively function in modern society. This paper attempts to fill these gaps by using existing data to report on the levels of achievement in reading and numeracy in the 20 years since the first national study of school performance.

In addition to examining achievement and mastery over time this paper also reports on the social and educational correlates of achievement in reading and number work. There has been a longstanding interest in the extent to which achievement at school is associated with social background, gender and ethnicity. More recently, there is a renewed interest in the roles of school systems and schools in shaping patterns of achievement. If it is known which students in what type of schools are not achieving mastery, then it helps in seeking explanations as to why. Detailed information at both the individual and organisational levels is needed to determine why some students have lower achievement levels. A strong empirical foundation is essential to develop appropriate policies for improving standards in reading and numeracy.

DATA AND ANALYSES

As the investigation was based on data from several studies, procedures were designed to establish the best possible basis for comparisons. This section briefly describes the data that were used, the achievement tests, the common scales constructed to link the different tests, and the analyses that were conducted. Appendix 1 provides more information about these issues and the general research design.

Samples and Data

The analyses made use of eight data sets spanning a 20 year period. The four main data sets were based on large representative samples of students who completed tests in 1975 (*Australian Studies in School Performance*/ASSP), 1980 (*Australian Studies in Student Performance*/ASSP), 1989 (*Youth in Transition*/YIT) and 1995 (*Longitudinal Surveys of Australian Youth*/LSAY). Four smaller data sets from 1989, 1990, 1991 and 1992 (*Australian Youth Survey*/AYS) were used in some parts of the analyses. An outline of the size and structure of the data sets is contained in Appendix 1. Generally, these studies were based on samples of 14 year-olds, but the 1995 LSAY sample was of year 9 students and the AYS samples included some 13 and 15 year-old students. For this reason three lines of data are presented in many of the tables in the report: all students tested, all 14 year-olds tested and 14 year-olds in year 9. Although the patterns for the means are similar whichever group is compared, the prime reference is 14 year-old students.

Any estimate based on a sample will have a sampling error associated with it. As a guide to the accuracy of these samples, Keeves and Bourke (1976) estimated that the 95 per cent confidence interval for a percentage of 70 was plus or minus 2.2 percentage points on the reading and numeracy tests in the 1975 sample. This means that a difference of up to 4.4 percentage points could be attributed to sampling variation. For smaller groups (such as the samples from smaller states and territories) the sampling errors and confidence intervals would be larger than this.

The Achievement Tests and the Common Scales

The reading comprehension tests and numeracy comprised 20 to 35 items, more commonly multiple choice questions but also short answer questions. The tests include a substantial number of items common to two or more tests. The reading and numeracy tests in the 1975 ASSP study were intended to assess minimum competency but later tests were designed to provide information about a wider range of student performance and thus incorporated more difficult items. In this report student performance on individual common items is compared but it is especially advantageous to use information on student performance from all items. To that end common scales for reading and numeracy based on all items were constructed and used in reporting results. The procedure for establishing these scales is outlined in Appendix 1. Even though the scales were calibrated to range from 0 to 100, the scale is arbitrary and does not correspond to the percentage of items correct.

Achievement and Mastery

Achievement is represented by performance on the reading and numeracy tests and is measured on the common scales. In this report overall levels of achievement in reading and numeracy at each time are summarised by the mean (average) score. This statistic provides the most stable indication of achievement for a particular cohort or group. Mastery refers to satisfying the criteria defined in 1975 as necessary to function in adult society. For later studies it was measured by applying the cutting score on the common scale that differentiated mastery from non-mastery in 1975 to the particular test completed by the student. This paper reports the percentages of

students who attained mastery. These percentages tend to be less stable than mean levels of achievement since they are based on cut-offs and therefore are sensitive to small changes in distributions. This is especially the case for small samples such as the *Australian Youth Survey* samples and the social groups within the larger studies.

Analyses

In the report, analyses of reading comprehension and numeracy are presented separately but within a similar structure. First, item facilities (the percentage correct) for items common to two or more tests are presented and discussed. These data can provide an indication of changes in achievement in particular areas; detail that is lost when overall scales are compared. Secondly, changes over time in achievement and mastery are reported. In this, and in following analyses, the common achievement scales are employed. Thirdly, the report examines the association of achievement with gender, social background, and state or territory. These investigations use simple measures of association as well as multivariate and multilevel analyses.

Multivariate procedures are required because there are multiple influences on achievement and these influences are inter-related. Multilevel (or hierarchical) analyses are necessary since students are clustered within schools and there may be contextual influences from schools. The multivariate and multilevel statistics and tests of significance are derived from analyses of achievement in reading and numeracy that incorporate all predictor variables in a model of achievement. These analyses uncover the net effects of predictor variables, that is the effect of a predictor variable controlling for other influences (specified in the model). These relationships are reported for different times (when the data are comparable) thus enabling the identification of trends. The multivariate multilevel procedures determine whether a change in the performance of particular social group is because of changes in the social characteristics of that group (for example changes in its socioeconomic composition) or represents a real change.

ACHIEVEMENT IN READING COMPREHENSION

This section presents the results of some of the analyses of achievement in reading comprehension. The first part focuses on individual items, the second considers general trends in achievement and mastery over 20 years and the final part explores patterns of association with demographic factors.

Changes in the Performance on Individual Items

Table 1 presents the percentage of students correctly answering common items on the reading tests. Items common to only the 1975 and 1980 ASSP are not included in this table and have been discussed elsewhere (Bourke et al., 1981).

Table 1 Percentage of 14 year-old Students Correctly Answering Common Reading Comprehension Items

Item	All	1975 ASSP	1980 ASSP	1989 AYS	1989 YIT	1990 AYS	1991 AYS	1992 AYS	1995 LSAY

Newspaper Articles									
How did Oil get there	88	89	88	87	86	-	-	-	-
Who Shot the Birds	85	82	81	89	91	-	-	-	-
Tiger Article	77	73	73	75	79	-	-	-	-
Pay Grab Means	84	79	-	84	83	86	84	87	87
On Strike	75	-	78	80	79	80	76	78	69
Heave-Ho Battle Means	82	-	83	83	84	82	78	79	81
Valour Award Tells of	93	-	92	90	92	93	92	91	94
Acne Occurs when	51	-	-	-	-	50	-	53	-
Acne Becomes Severe if	69	-	-	-	-	69	-	70	-
Don't Squeeze Pimples	68	-	-	-	-	69	-	68	-
Floating Best Described as	68	-	-	-	-	-	55	57	71
Floating Depends on	49	-	-	-	-	-	49	50	-
Bee-Line is	85	-	-	-	-	-	72	-	86
Comprehension of Text									
Paracutin-Name of What	65	75	-	70	70	-	-	-	55
Paracutin-Destroyed in Eruption	89	93	-	91	90	-	-	-	85
Paracutin-Asleep Means	86	87	-	87	88	-	-	-	83
Paracutin- is Now	63	57	-	65	67	-	-	-	-
Paracutin-Location of Volcano	61	65	-	55	57	-	-	-	-
Paracutin-Learn about Volcanoes	77	86	-	70	69	-	-	-	-
Train-Affect of Bad Weather	62	-	-	-	-	69	-	66	60
Train-Man didn't Go into Niche	85	-	-	-	-	93	-	90	83
Train-Sign of Second Train	61	-	-	-	-	60	-	61	-
Train-Two Trains in Tunnel	57	-	-	-	-	62	-	55	57
Train-Mistaken about Distance	64	-	-	-	-	68	-	66	63
Train-Good Lord with Me	52	-	-	-	-	54	-	59	51
Nuclear War-Role of Scientists	38	-	-	39	41	37	32	32	-
Nuclear War-Attitude to Bomb	32	-	-	-	-	36	31	31	-
Nuclear War-Working On Bomb	34	-	-	-	-	37	33	33	-
Nuclear War-Duty	54	-	-	54	54	-	-	-	-
Nuclear War-Evidence	32	-	-	28	33	-	-	-	-
Nuclear War-End Humans	56	-	-	56	55	-	-	-	-
Other Items									
Crossword Clue-Anger	33	-	-	32	33	30	33	36	-
Letters Page	63	-	65	62	63	64	61	61	64
Perth's Maximum Temperature	83	-	88	82	80	83	81	-	-

Before considering trends, it is worth noting that the results for the two studies conducted in 1989 are consistent on most items with the percentage of students answering correctly differing by only a few percentage points. This result suggests that the tests are reliable. They are not subject to wide and unaccountable changes.

Overall, two counteracting changes have occurred. Performance on the items relating to short newspaper stories have improved or stayed the same while that on items relating to longer textual passages has generally declined.

Of the items relating to newspaper articles, performance has improved on the 'pay grab', 'valour award', 'floating described as' and 'bee-line' items. In contrast performance on the 'on strike' item declined. On the other items, the percentages who answered correctly are remarkably constant.

On the more difficult passages, the item facilities have declined (ie. a smaller percentage of students were correct). Several questions pertaining to the Mexican volcano, Paracutin, show a decline, both between 1975 and 1989 and between 1989 and 1995. Similarly, most questions about the express train passage, show a decline between 1990 and 1995. The item facilities for the first question on the Manhattan project also suggest a decline. Since these declines cannot be attributed to changes in the format of the questions or their position in the test, they are most probably real.

Trends in Reading Comprehension

Table 2 records the mean scores for reading achievement on the common scale for each of the studies. Since 1975 the level of reading achievement has either stayed about the same or declined slightly. With 14 year-old students, or with all students tested, the decline is around one or two scale points, well within sampling error. When the analysis is restricted to 14 year-olds in year 9, a greater decline is discernible but again the variations are within (the larger) sampling variation. Overall, the conclusion from these data is that there is no substantial change in performance in reading comprehension over the last two decades.

Table 2 Trends in Reading Comprehension from 1975 to 1995

	All	1975 ASSP	1980 ASSP	1989 AYS	1989 YIT	1990 AYS	1991 AYS	1992 AYS	1995 LSAY
Mean Scores on Common Scale									
All Students	65	66	66	66	66	65	62	63	64
14 year-old Students	65	66	66	66	66	66	62	64	65
14 year-olds in Yr 9	67	69	68	67	67	68	63	66	65
Percentage of Students Achieving Mastery									
All Students	70	71	72	73	72	76	64	67	68
14 year-old Students	71	72	72	73	72	77	63	68	70
14 year-olds in Yr 9	75	79	78	75	76	84	66	74	70

Another perspective on looking at changes in performance over time is to examine the percentage of students who attained mastery in reading. Table 2 also presents the percentage of students in each study who achieved mastery for Australia as a whole. (The figures for the 1975 data correspond to the 72 per cent who achieved mastery as reported by Bourke and Keeves, 1977: 54). Overall, there has been no change in the proportions of students who have attained mastery of these literacy skills of reading comprehension. The changes observed are within sampling variation. There was a larger difference when comparing just 14 year-olds in year 9. This is fluctuation rather than decline and due to the instability of cut-offs in severely restricted samples. The data for means in the full 14 year-old sample show much greater stability.

Patterns of Achievement in Reading Comprehension

In this section relationships between achievement in reading comprehension and gender, ethnicity, parental occupational background (a measure of socioeconomic status) and state or territory are investigated using both bivariate and multivariate methods of analysis.

Differences Between Male and Female Students

Table 3 shows the average levels of achievement in reading comprehension for boys and girls. Consistent with results reported in the research literature, achievement in reading is slightly higher among girls than boys. Moreover these data suggest a tendency towards a widening of the gap between males and females during the 1990s. Analyses of the smaller data sets from the *Australian Youth Survey* support this interpretation. This widening gap between males and females over the 1990s is consistent with trends in other outcomes such as performance on final secondary school assessments (McCann, 1995; McGaw et al., 1989).

Table 3 Performance on Reading Comprehension for Males and Females

	1975		1980		1989		1995	
	Male	Female	Male	Female	Male	Female	Male	Female
Mean Scores on Common Scale								
All Students	65	67	65	66	65	66	62	66
14 year-old Students	65	67	65	66	65	66	63	66
14 year-olds in Yr 9	69	69	67	68	67	68	63	66
Percentage of Students Achieving Mastery								
All Students	70	73	71	73	70	75	63	72
14 year-old Students	70	73	72	73	70	75	66	74
14 year-olds in Yr 9	78	80	76	79	74	79	66	74

Table 3 also shows the same trend from a different perspective. It records the percentages of males and females who demonstrate mastery on these reading tests. In 1975, 70 per cent of 14 year-old male students, compared to 73 per cent of females attained mastery, whereas in 1995 the percentages were 66 per cent and 74 per cent respectively. The divergence in mastery levels between boys and girls is substantial and suggests that further investigation is warranted.

Multivariate analyses confirm that the gap between boys and girls in reading comprehension increased over time. In 1995 boys scored two scale points lower than girls. In the 1975 and 1989 studies, the differences were smaller and not statistically significant.

Differences Associated with Language Background

In three of the data sets it was possible to examine the difference in achievement between students whose home language was English and those whose language at home was another language. The results are recorded in Table 4.

Table 4 Mean Scores and Mastery of Reading Comprehension by Language Spoken at Home

	1975 Language at Home		1980 Language at Home		1995 Language at Home	
	English	Other	English	Other	English	Other
Mean Scores on Common Scale						
All Students	66	56	66	57	65	57
14 year-old Students	66	56	66	57	65	59
14 year-olds in Year 9	69	58	68	60	65	59
Percentage of Students Achieving Mastery						
14 year-old Students	73	47	74	53	72	53

Not surprisingly, students who spoke a language other than English at home performed substantially less well than their peers. There did appear to be a closing of the gap over the period from 1975 to 1995 and this was investigated further with multivariate analyses.

Multivariate analyses showed that the effect of language at home was partly due to socioeconomic differences between the two language groups. After an allowance was made for socioeconomic background the difference between the groups became smaller and was approximately six scale points in both 1975 and 1995.

These findings suggest that the apparent reduction of the gap between the two language groups over the period is due to higher levels of socioeconomic status of students whose language at home is not English in 1995 compared to 1975. In other words there was no improvement in reading performance for students of comparable socioeconomic background.

There is an increase in mastery of reading among students whose main language at home was not English. In 1975 47 per cent of these students attained mastery compared to 53 per cent in 1995, still trailing the English at home group by nearly 20 percentage points. This is much larger than differences in mean scores. As discussed above, the narrowing of the gap in mastery is mainly due to changes in the socioeconomic profile of non-English language groups.

Differences Associated with Parents' Occupational Group

Table 5 presents the mean scores on the common reading scale according to parental occupational group. Parents' occupations were categorised into four broad groups that allow consistent categorisation over time. These data show that the difference between the professional-managerial group and the unskilled-semiskilled group has

narrowed over the past 20 years. In 1975 the gap was approximately eight scale points, whereas in 1995 it was five or six scale points. This apparent equalisation may be due to greater heterogeneity in the professional-managerial group or it may reflect a real decline in this aspect of social inequality.

Table 5 Mean Scores and Mastery of Reading Comprehension by Occupational Background

	1975 Occupational Group				1989 Occupational Group				1995 Occupational Group			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
Mean Scores on Common Scale												
All Students	71	68	67	63	71	67	66	63	68	65	63	62
14 year-old Students	71	68	67	63	71	67	66	63	68	65	63	63
14 year-olds in Yr 9	73	71	69	66	72	67	67	65	68	65	63	63
Percentage of Students Achieving Mastery												
14 year-old Students	85	76	75	64	85	76	75	64	79	72	67	67

Note. Group I comprised Professionals, Para-Professionals, Managers and Administrators
Group II comprised Clerical, Sales and Service Workers
Group III comprised Trades Persons
Group IV comprised Process Workers and Labourers (unskilled-semiskilled)

Multivariate analyses confirm that the impact of occupational background has declined over 20 years. For example, in 1995 the net difference between students from professional-managerial and unskilled backgrounds was around three points compared to five points in 1975 and 1989.

The multivariate analyses also indicate that part of the difference observed in Table 5 could be associated with school effects such as differences between school sectors. The net difference in reading achievement between students from professional-managerial and unskilled backgrounds was around five points in 1975 and 1989, compared to the eight points shown in Table 5. Therefore approximately three points could be attributed to associated factors including school sector.

There are more substantial socioeconomic differences with mastery than with overall performance (averages). In 1975 and 1989 20 per cent more students from professional and managerial backgrounds attained mastery compared to students from unskilled and semi-skilled backgrounds. This gap had narrowed to about 12 per cent in 1995. There was a decline in the mastery levels of reading comprehension for students from a trades or skilled worker background between 1989 and 1995.

Differences Associated with Aboriginal and Torres Strait Islander Background

Two of the large studies (1975 and 1995) distinguished Aboriginal and Torres Strait Islander students. In 1975 Aboriginal and Torres Strait Islander students were identified by the teacher, whereas in 1995 this information was provided by the student. The 1975 study also included a sizeable proportion of students in Aboriginal

schools. Table 6 shows that in 1975 the reading achievement of Aboriginal and Torres Strait Islander students was some 20 scale points lower than that for other students. Multivariate analyses suggest that much of this difference was due to socioeconomic, language and school differences since the net difference was only eight scale points.

Table 6 Mean Scores and Mastery of Reading Comprehension for Aboriginal and Torres Strait Islander Students

	1975		1995	
	Aboriginal	Non-Aboriginal	Aboriginal	Non-Aboriginal
Mean Scores on Common Scale				
All Students	44	66	55	64
14 year-old Students	44	66	55	65
14 year-olds in Year 9	49	69	55	65
Percentage of Students Achieving Mastery				
14 year-old Students	28	72	44	71

By 1995 the gross difference in achievement between the two groups had declined to approximately 10 scale points. Multivariate analyses suggest this reduction is probably due to socioeconomic and other differences between the Aboriginal and Torres Strait Islander samples, rather than a real improvement in reading comprehension among indigenous Australians. After adjusting for the effects of social background and schools, the gap was eight scale points in 1975 and seven in 1995. Therefore much of the improvement suggested by the data in Table 6 is apparent rather than real.

Differences between the studies in the schooling, home language and socioeconomic characteristics of Aboriginal students are also reflected by the substantial increase in the level of mastery, rising from 28 per cent in 1975 to 44 per cent in 1995. This difference in mastery levels appears larger than the corresponding differences in mean scores. Although this represents a substantial improvement, mastery levels among Aboriginal and Torres Strait Islander students remain the lowest of the social groups discussed in this report.

Summary

From analyses of performance in reading comprehension of junior secondary school students the following conclusions can be drawn.

- There has been little overall change in average levels of achievement or in the percentage of 14 year-old students attaining mastery.

- Differences between social groups in mastery of reading comprehension are most often greater than mean differences in performance.

- There has been a decline in the percentage of 14 year-old boys attaining mastery in reading and a widening of the gap between boys and girls.

- Test scores on reading comprehension are lower for students whose language at home is not English and the gap between these and other students of comparable socioeconomic background has not changed. The level of mastery for the home language other than English group has improved over time but still is considerably lower than for the English at home group.

- Higher achievement in reading is associated with higher socioeconomic status but the size of the effect appears to have diminished over time. There are larger socioeconomic differences in regard to mastery compared to mean level of achievement.

- Test scores and mastery for reading are lower for Aboriginal and Torres Strait Islander students and the net difference in performance between these and other students has not altered.

ACHIEVEMENT IN NUMERACY

This section discusses the results of the analyses of numeracy in a similar manner to the previous section on reading. It includes an examination of the individual item facilities, an examination of trends over time and an exploration of associations between numeracy and social factors. Where relevant the results for numeracy are compared with corresponding results for reading comprehension.

Changes in the Performance of Individual Numeracy Items

Table 7 presents the percentages of students correctly answering common items in the numeracy test. It excludes the very simple items common to only the 1975 and 1980 ASSP studies. For convenience of presentation, items in the table are classified somewhat arbitrarily into three groups: computational, practical and conceptual. The data indicate that performance on the 1980 ASSP is often (but not always) better than performance on the 1975 ASSP; a point noted both in the report on the 1980 study and the three-wave comparison of the performance of Victorian students (Bourke et al., 1981:144-145; McGaw et al., 1989: 102). The 1980 sample probably have undersampled low achievers in numeracy in Victoria. As was observed in the case of reading, most of the item facilities for the two tests performed in 1989 were very similar, suggesting that the sampling and testing procedures provide accurate estimates.

Table 7 Percentage of 14 year-old Students Correctly Answering Common Items in Numeracy

Item	All	1975 ASSP	1980 ASSP	1989 AYS	1989 YIT	1990 AYS	1991 AYS	1992 AYS	1995 LSAY
Computational									
Multiply 38 by 9	79	79	85	75	74	-	-	-	-
Division 125/5	91	90	93	-	-	-	-	-	-
Estimate Long Division 1240/29	67	69	74	64	62	-	-	-	66
Fraction Closest to 3/16	61	63	58	65	62	-	-	-	60
Which Fraction Correct	76	-	-	75	76	-	-	-	-
Which Equation is Wrong	58	-	-	60	57	-	-	-	-
Value of 2^3 x 3^2 is	53	-	-	56	52	-	-	-	54
Square Root of 12 x 75 is	59	-	-	61	58	-	-	-	-
Practical									
Chemist: Hours Open each Day	87	83	84	89	89	89	87	90	89
Interest on $900 at 10%	72	70	78	72	70	-	-	-	-
Michaelangelo's Pizza-Fractions	60	-	-	-	-	61	58	61	-
How Far can Snail Travel	93	-	-	94	94	93	92	94	94
Price of Tee-Shirt Before Sale	32	-	-	-	-	-	31	31	32
Average Rainfall During Period	61	-	-	61	60	56	58	59	63
Distance Brite To Dalton	83	79	84	-	-	-	-	-	85
Graph: Apples Mr Green Picked	87	88	90	85	84	-	-	-	-
Town with Highest % River Fever	39	-	-	40	40	36	35	40	-
Conceptual									
Perimeter of Rectangle	74	-	-	-	-	70	-	72	75
End Point of Cross-Country Run	80	-	-	-	-	80	76	77	82
Which Statement must be True	49	-	-	51	49	-	-	-	-
Which Figure Completes Pattern	73	-	-	71	72	74	71	74	73
Axis: Shows Complete Figure	54	-	-	47	49	50	50	51	62
Distance Car A Ahead in 3 Hours	79	-	-	-	-	76	-	73	81
When Trains will Pass each Other	44	-	-	-	-	41	41	44	45
Who is Taller than Spiro	85	-	-	-	-	83	-	84	85
Total Membership of Sports Club	68	-	-	71	67	70	62	70	69
Length of Videotape Needed	65	-	-	-	-	64	65	65	-
Symbols on Zyplon	61	-	-	-	-	-	59	63	-
Donuts Left Over	57	-	-	-	-	-	55	58	-
No. of Blocks to Make Stack	34	-	-	-	-	30	31	34	35
Weight of Sponge	27	-	-	-	-	-	27	26	28
Jane Needs to Turn x Degrees	20	-	-	20	21	18	20	18	21
Sum of Dots on Two Faces of Dice	38	-	-	-	-	35	-	-	39
How Many Blocks Fit into Box	60	58	62	63	61	-	-	-	-
Cheapest Marmalade	41	-	-	42	41	-	-	-	-

On most common items little change has occurred suggesting that there has been no overall change in numeracy between 1975 and 1995. There was a small decline on the computational items, no consistent change on the practical items and a slight improvement on the conceptual items.

Among the computational items on which performance declined was the item asking students to calculate the product of 38 and 9. On this item the percentage of students

answering correctly declined from 79 per cent in 1975 to around 75 per cent in 1989. Similarly performance on the item asking students to estimate long division declined from 69 per cent in 1975 to about 63 per cent in 1989. An item requiring understanding of fractions also showed a decline from a facility of 63 per cent in 1975 to about 60 per cent in 1995, although the results in the intervening studies are not consistent.

These declines could be attributed to a decreasing emphasis on these more traditional aspects of primary and junior secondary school mathematics. It is of interest to note that on the Third International Mathematics and Science Study, Australian 13 year-old students performed relatively poorly on computational items (multiplication, division, combining fractions) and relatively well on items concerned with conceptualisation and interpretation (Lokan et al., 1996:47-48).

Generally, performance on the practical items did not change. However, there are a few exceptions on which performance did change. These include the simple item on asking students how many hours a chemist is open, if it opens at 8 am and closes at 10 pm. On that item there was an improvement from 83 per cent correct in 1975 to 89 per cent in 1995. The item testing graph reading skills showed a decline from a facility of 88 per cent in 1975 to 84 per cent in 1989.

Performance on the conceptual items appears to have improved. Between 1989 and 1995 on at least seven of the 12 items, the facilities in 1995 were higher than at earlier times. This is not due to differences between sampling procedures since performances on the computational and practical items do not show this improvement.

Trends in Numeracy

In terms of mean scores on numeracy scale, Table 8 shows that among 14 year-olds there has been no change in performance on numeracy between 1975 and 1995. The slight decline among 14 year-olds in year 9 may simply reflect sampling variations or possibly changes in the year levels in which aspects of numeracy are taught.

Table 8 Trends in Numeracy from 1975 to 1995

Year /Study	All	1975 ASSP	1980 ASSP	1989 AYS	1989 YIT	1990 AYS	1991 AYS	1992 AYS	1995 LSAY
Mean Scores on Common Scale									
All Students	64	64	66	63	63	62	62	63	64
14 year-old Students	64	64	66	64	63	62	62	63	64
14 year-olds in Yr 9	65	67	68	64	64	64	62	65	64
Percentage of Students Achieving Mastery									
All Students	79	74	80	77	74	78	78	76	83
14 year-old Students	78	74	80	77	74	79	78	77	85
14 year-olds in Yr 9	82	81	85	79	78	84	81	82	85

Table 8 also indicates that there has been a small improvement in the percentage of students attaining mastery in numeracy. As suggested earlier, the 1980 result is probably due to under-sampling of low achievers rather than a temporary rise in the proportion of students attaining mastery. Against the 1975 reference point, the results for 1995 suggest a substantial improvement but allowance needs to be made for the fact that the later sample only included 14 year-olds in year 9. When comparisons are made with a similarly restricted sample, the most reasonable conclusion is that there has been an improvement in the percentage of students attaining mastery for numeracy. The increase in the percentage with mastery in numeracy in the context of stable mean levels suggests that remedial education in mathematics is succeeding.

Patterns of Achievement in Numeracy

In this section the differences in numeracy between boys and girls, several social groups and the states and territories are examined through both bivariate and multivariate analyses.

Differences Between Male and Female Students

Table 9 shows the average levels of achievement on the numeracy scale for males and females at four points in time. Consistent with the research literature, achievement in numeracy is marginally higher for males than females; the difference being approximately two scale points. This differences is similar in magnitude but in the opposite direction of that for reading. Over time there does not appear to have been any change in the magnitude of the difference.

Table 9 also records the percentages of males and females achieving mastery in numeracy. The difference between males and females was one or two percentage points in three of the four studies. The large difference in 1980 was not replicated in the other studies.

Table 9 Performance on Numeracy for Males and Females

	1975 ASSP		1980 ASSP		1989 YIT		1995 LSAY	
	Male	Female	Male	Female	Male	Female	Male	Female
Mean Scores on Common Scale								
All Students	65	63	67	65	63	62	65	63
14 yr-old Students	65	63	67	65	63	62	65	63
14 year-olds in Yr 9	68	66	69	67	65	64	65	63
Percentage of Students Achieving Mastery								
All Students	75	73	82	78	75	74	84	82
14 yr-old Students	75	73	82	78	75	74	85	84
14 year-olds in Yr 9	82	80	87	83	78	78	85	84

Multivariate analyses showed that the gender gap in numeracy was statistically significant in the three studies examined (social background data were not available for 1980). It should be kept in mind that the gender difference in means for both

reading and numeracy is of the order of two scale points, smaller than most other significant differences noted in these analyses.

Differences Associated with Language Background

Table 10 presents the mean numeracy levels of students by home language: English or another language. It indicates a diminishing gap between the two groups over the period from 1975 and 1995. As might be expected, the effect of home language is much less for numeracy than for reading.

Table 10 Mean Scores and Mastery of Numeracy by Language Spoken at Home

	1975		1980		1995	
	Language at Home		Language at Home		Language at Home	
	English	Other	English	Other	English	Other
Mean Scores on Common Scale						
All Students	64	56	66	57	64	61
14 year-old Students	65	56	66	57	65	61
14 year-olds in Year 9	68	59	69	58	65	61
Percentage of Students Achieving Mastery						
14 year-old Students	75	58	81	61	86	77

Multivariate analysis showed that for the 1975 study, the differences in numeracy between home language groups was largely due to socioeconomic differences. The nine point difference observed in Table 10 is reduced to only three points when allowance is made for socioeconomic background. Socioeconomic differences were less important in 1995. Half of the four point difference between language groups observed in Table 10 for 1995 was attributable directly to language at home. Therefore these investigations show that differences in performance directly attributable to language background declined only a little, from three to two scale points, over the 20 year period.

As was the case for reading, there are greater differences between home language groups for mastery than for overall (mean) performance. There was a reduction in the mastery gap from 17 percentage points in 1975 to 9 percentage points in 1995. However, as suggested above, this reduction is due mainly to the changing socioeconomic profiles of non-English language groups.

Differences Associated with Parents' Occupational Group

Mean scores on the common numeracy scale by parental occupational group are presented in Table 11. Students from advantaged backgrounds score higher than other students even though the differences are not large. Students from professional-managerial backgrounds scored higher on the numeracy tests than students from

lower white collar backgrounds who in turn scored higher than those from trades and unskilled or semiskilled backgrounds.

In multivariate analyses the differences observed are smaller. For the 1975 study the average scores of students from professional-managerial and unskilled-semiskilled backgrounds differed by six points, compared to nine points in the unadjusted analyses. The difference could be partly due to the associated influence of school sector which also has an effect on achievement.

Table 11 Mean Scores and Mastery of Numeracy by Occupational Background

	1975				1989				1995			
	Occupational Group				Occupational Group				Occupational Group			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
Mean Scores on Common Scale												
All Students	70	66	65	61	68	64	63	59	68	64	63	61
14 year-old Students	70	66	65	61	68	64	63	59	68	64	63	62
14 year-olds in Yr 9	72	68	67	64	69	64	64	61	68	64	63	62
Percentage of Students Achieving Mastery												
14 year-old Students	86	77	77	68	86	81	77	66	90	87	84	80

Note: Group I comprised Professionals, Para-Professionals, Managers & Administrators
Group II comprised Clerical, Sales and Service Workers
Group III comprised Trades Persons
Group IV comprised Process Workers and Labourers (unskilled-semiskilled)

Table 11 shows a decline in differences in achievement according to occupational background. In 1975 there was a difference between the highest and lowest status groups of eight to nine scale points but in 1995 it had declined to about six scale points. This change is similar in magnitude to that found for reading.

The multivariate analyses confirm the decline in the relationship between achievement and occupational background. In 1975 and 1989 the net difference in performance of students from professional-managerial and unskilled-semiskilled backgrounds was six points, but had declined to four points in 1995.

Socioeconomic background has a stronger influence on mastery than on mean achievement. This result was also found for reading. In 1975 and 1989 respectively, 18 and 20 per cent more students from professional and managerial backgrounds attained mastery in numeracy than students from unskilled backgrounds. In 1995 this difference had declined to 10 per cent.

Differences Associated with Aboriginal and Torres Strait Islander Background

Table 12 shows that differences between Aboriginal and non-Aboriginal students in numeracy have diminished considerably. Among 14 year-olds the gap was 27 scale points in 1975 but it had contracted to nine points in 1995. The decline in the performance gap was more dramatic for mastery. The percentage of Aboriginal and

Torres Strait Islander students attaining mastery increased from 22 per cent in 1975 to 65 per cent in 1995.

Further analyses examined the extent to which this change in mean levels was a real improvement or reflected differences in the Aboriginal and Torres Strait Islander samples. Multivariate analyses confirmed that there had been a real improvement in numeracy among Aboriginal and Torres Strait Islander students (although of smaller magnitude than that suggested by the data in Table 12). The difference in numeracy between indigenous and non-indigenous Australians declined from ten to six points.

Table 12 Mean Scores and Mastery of Numeracy for Aboriginal and Torres Strait Islander Students

	1975		1995	
	Aboriginal	Non-Aboriginal	Aboriginal	Non-Aboriginal
Mean Scores on Common Scale				
All Students	37	64	56	64
14 year-old Students	37	64	56	65
14 year-olds in Year 9	44	67	56	65
Percentage of Students Achieving Mastery				
14 year-old Students	22	75	65	86

Summary

Analyses of performance by 14 year-old school students on tests of numeracy over the last 20 years has led to the following conclusions.

- There has been little overall change in average levels of achievement in numeracy but there has been a small improvement in the percentage of 14 year-old students attaining mastery.

- As for reading, social differences in mastery are greater than for all overall performance.

- There has been little change in the relative performance of boys and girls on these tests with boys continuing to do slightly better than girls.

- Performance in numeracy is lower for students whose language at home is not English but the gap is less than for reading. The apparent improvement is due to changes in socioeconomic composition of non-English speaking students. The level of mastery in numeracy of students from non-English speaking backgrounds has increased considerably.

- Achievement in numeracy is associated with socioeconomic status but, as for reading, its influence has declined. There are stronger socioeconomic differences for mastery.

- Achievement in numeracy is lower for Aboriginal and Torres Strait Islander students than for other students but the gap has diminished somewhat over time. The levels of mastery for this group have increased dramatically.

DISCUSSION

The major conclusion of this investigation is that there has been little change in average standards of reading or numeracy over the past 20 years. This conclusion is based on trends in the mean scores on the common scale established to link the various studies. It is also indicated by facilities on the common items. In the case of numeracy there was an increase in the percentage of students who had reached the level defined as representing mastery. In reading there was no detectable change in this percentage. Whether these findings are interpreted as 'no decline' or 'no improvement' is a question of perspective.

Influences Outside of School

It is important that trends in achievement be interpreted in terms of a range of possible home and social influences as well as school factors. For example, an important factor in the development of reading skills is the extent to which parents read to, and engage in other language interactions with, children from an early age (Rowe, 1991). Therefore, changes in the extent to which parents read with their children would be expected to impact on the standards of reading comprehension. Similarly, changes in the composition of the school age population in terms of the extent to which a language other than English was spoken at home would be expected to influence scores on tests of reading comprehension. This is not to deny the need to investigate the effects of changes in school practices and approaches to teaching reading.

Reading

The results presented in this paper indicate that there remains a number of students (approximately 30 per cent) who have not attained mastery in the important area of reading. Although this does not mean that these students are 'illiterate', in the most common usage of that term, it does indicate that they cannot comprehend the written word at an appropriate level. The level set for mastery is not beyond dispute but it is arguable in terms of what was defined in 1975. Developing high levels of reading comprehension is a primary purpose of schooling, and traditionally a purpose that has been pre-eminent for primary schools.

In the literature there is support for the view that improvements in this aspect of literacy are more likely to be successful if they are focused on what happens in primary school rather than being attempted as remedial programs at a later stage. The work of Rowe & Hill (1996) shows a progressive widening of performance among students with the lowest 10 per cent (and to some extent the lowest quarter) making rather less progress beyond Year 4 than their peers. This result is consistent with a study in the United States which suggested that students from poorer backgrounds achieved as well as the general population up to Year 3 but fell behind from Year 4 onwards (Chall, Jacobs & Baldwin, 1990).

The data in this paper also indicate that lack of mastery in reading is greater among boys than girls and that the gap between girls and boys has widened in recent years; as it has for other aspects of performance in school. This result points to an area that deserves closer scrutiny and thoughtful action. In other studies it is generally found that girls have more favourable attitudes to school in the late primary and in the secondary years than boys (Ainley, 1995) but it is not clear why this should be so and whether girls are more actively engaged with at least some of the learning tasks at school. Rowe & Hill (1996) suggest that the gap in performance in reading between boys and girls widens after about Year 4 in primary school.

Numeracy

In terms of numeracy, it appears that, even though average levels of achievement have not changed over 20 years, there has been a reduction in the percentage of 14-year olds who did not achieve mastery. On the basis of the analyses in this paper, it is estimated that between 15 and 20 per cent of 14 year-olds failed to demonstrate this level of mastery. In numeracy, boys performed a little better than girls, a difference that had not changed over the period investigated.

Monitoring Achievement in Other Countries

In a high percentage of member countries of the Organisation for Economic Cooperation and Development, national programs for monitoring achievement have been implemented since the mid 1970s (Nuttall, 1993). The countries with identified national systems for monitoring achievement include Britain, France, Finland, the Netherlands, Spain, Sweden and the United States. These countries have diverse educational, political and cultural traditions but their national monitoring systems are quite similar. Typically these national monitoring systems test a limited range of skills and content (reading and language, maths and science) and proceed through them in a cycle. Most commonly they are based on data gathered from samples of schools and students. The National Assessment of Educational Progress surveys in the United States are notable for the range of student background data which are collected, and for the reporting of the performance of different groups of students.

This investigation was conducted using data from a number of different studies, only two of which were designed to monitor achievement on a national basis. It indicates something of what might be possible through regular monitoring programs based on light sampling procedures. Australia began a national monitoring program in the mid 1970s but did not continue with it. Since the mid 1980s most Australian states and territories have implemented differing monitoring programs at primary school level (Masters et al., 1990; Titmanis et al., 1993). Now it would appear to be an opportune time to consider a national monitoring program that also included secondary schools.

The study is in no sense an evaluation of the effectiveness of Australian schools. The development of skills in reading and numeracy is a central concern of schools but by no means the only one. With the general picture showing little overall change in performance, the analyses provide mixed messages about changes on particular groups of items and variations in performance among groups of students. The detail in those results signal areas where further inquiry might be warranted. These more

detailed investigations could then provide the basis for thoughtful action. It is unlikely that there will be simple universal solutions. However, if reading and numeracy are considered essential skills in our society, there is a moral imperative to reduce the proportions of young people who do not achieve mastery of those skills.

REFERENCES

Ainley, J. 1995. Students' perceptions of their schools. *Unicorn*, 21 (3), 5-16.

Bourke, S. F, Mills, J. M., Stanyon, J. & Holzer, F. 1981. *Performance in Literacy and Numeracy: 1980.* A Report to the Australian Educational Council on the Australian Studies in Student Performance Project. Canberra: AGPS.

Bourke, S. F. & Keeves, J. P. 1977. *The Mastery of Literacy and Numeracy: Final Report.* Australian Studies in School Performance. Volume III. ERDC Report no. 13. Canberra: AGPS.

Chall, J. S., Jacobs, V. & Baldwin, L. 1990. *The Reading Crisis: Why Poor Children Fall Behind*. Cambridge, Mass.: Harvard University Press.

Keeves J. P. & Bourke, S. F. 1976. *Literacy and Numeracy in Australian Schools: A First Report.* Australian Studies in School Performance Volume I. ERDC Report no. 8. Canberra: AGPS.

Keeves, J. P., Mathews, J. K. & Bourke, S. F. 1978. *Educating for Literacy and Numeracy in Australian Schools*. (Australian Education Review No. 11). Melbourne: ACER.

Lokan, J. & Ford, P. 1994. *Mapping State Testing Programs*. Melbourne: National Industry Education Forum.

Lokan, J., Ford, P. & Greenwood, L. 1996. *Maths and Science on the Line: Australian Junior Secondary Students' Performance in the Third International Mathematics and Science Study*. Melbourne: ACER.

Masters, G., Lokan, J., Doig, B., Khoo, S. K., Lindsey, J., Robinson, L. & Zammit, S. 1990. *Profiles of Learning: The Basic Skills Testing Program in New South Wales*. Hawthorn: Australian Council for Educational Research.

McCann, R. 1995. Sex differences in participation and performance at the NSW Higher School Certificate: A method which adjusts for the effect of differential selection. *Australian Journal of Education*, 39 (2), 163-188.

McGaw, B., Long, M. G., Morgan, G. & Rosier, M. J. 1989. *Literacy and Numeracy in Victorian Schools: 1988*. ACER Research Monograph No. 34. Hawthorn: Australian Council for Educational Research.

Nuttall, D. 1993. Monitoring national standards: United Kingdom. In T. Husen & N. Postlethwaite (Eds) *International Encyclopedia of Educational Research.* London: Pergamon.

Rosier, M. 1980. *Changes in Secondary School Mathematics in Australia 1964-1978*. Melbourne: ACER.

Rowe, K. J. 1991. The influence of reading activity at home on students' attitudes towards reading, classroom attentiveness and reading achievement: an application of structural equation modelling. *British Journal of Educational Psychology,* 61, 19-35.

Rowe, K. J. & Hill, P. W. 1996. Assessing, recording and reporting students' educational progress: the case for 'subject profiles' *Assessment in Education*, 3 (3), 309-352.

Titmanis, P., Murphy, F., Cook, J., Brady, K. & Brown, M. 1993. *Profiles of Student Achievement: English and Mathematics in Western Australian Government Schools 1992*. Perth: Ministry of Education.

APPENDIX 1: DATA AND METHODS OF ANALYSIS

This Appendix outlines a number of technical details of the research. An understanding of these features is essential to an appreciation of the data presented in the report. The report makes use of data that were derived from several sources incorporating similar but non-identical tests. It was therefore important to establish the best possible basis for analysis and comparison. The Appendix describes the data that were used and the nature of the achievement tests, defines the terms achievement and mastery, and outlines the procedure for the development of the common scales of literacy and numeracy and the measure of mastery.

Sources of Data

A total of eight data sets spanning a 20 year period from 1975 were used in these analyses. The four main data sets were as follows:

- The sample of 6,647 14 year-old students from the 1975 *Australian Studies in School Performance* (ASSP) that became the first cohort of the Youth in Transition project in 1978;
- The sample of 5,103 14 year-old students from the 1980 *Australian Studies in Student Performance* (ASSP);
- The sample of 5,597 14 year-olds from the 1989 *Youth in Transition* (YIT) study; and
- The sample of 13,615 year 9 (the modal year level for 14 year-olds) students from the 1995 *Longitudinal Survey of Australian Youth* (LSAY).

In addition, data from smaller samples of four of the cohorts for the *Australian Youth Survey* (AYS) were used in the calibration of the common scales and in some instances to confirm national trends between 1989 to 1995. These students completed achievement tests in the years from 1989 to 1992. The AYS samples comprised students who were mostly aged 14 years (but with varying ages and across years 8, 9 and 10). Later, when they were 16 years old, most were added to the ongoing *Australian Youth Survey*.

In general, these samples were two stage cluster samples, with random selection of schools in the first stage and random selection of students in the second stage. In the ASSP, YIT and LSAY studies, states were disproportionately sampled so that sufficient numbers of students were obtained to provide accurate population estimates. For some studies, the samples were also stratified by school sector. Weights appropriate to the sample design were calculated to provide population estimates of means and proportions.

Comparisons could be made between all students tested, 14 year-old students, or just 14 year-old students in year 9. Even though the data for all students provide the greatest amount of information, differences in the age structure of the samples need to be taken into account. The sub-sample of 14 year-olds is the most important because it allows simple direct comparisons between the ASSP, YIT, and AYS samples and

the majority of the students in the LSAY sample. This is the prime reference in the test and in the analyses and is shown as the middle line in the tables. The sub-sample of 14 year-olds in year 9 allows direct comparisons of the data from these studies with the LSAY sample but the estimates are less accurate and subject to the effects of changing policies affecting the age-grade distributions of students.

Achievement and Mastery

Achievement is typically defined as the degree to which a person has acquired skills or knowledge in a given field. It is therefore a continuous quantity that can be assessed by testing as well as by other methods. Even though the 1975 national study of literacy and numeracy focused on minimum competence, it did not preclude the construction of a graded scale. The second national survey in 1980, and the later tests used in this report aimed to assess a wider range of performance by including more difficult items.

Mastery refers to reaching a defined standard with respect to skills or knowledge in a given field and is usually a dichotomous concept. It is often used in association with the skills required to function effectively in society, or to effectively perform a given task. The concept of mastery employed in this paper is the same as that developed in the 1975 *Australian Studies of School Performance* study. A major aim of that study was to estimate the number of students failing to attain basic skills literacy and numeracy using a specified criterion. In the report the measures reported are the percentages of students attaining mastery in reading or numeracy.

The Achievement Tests

The reading and numeracy tests used in the initial 1975 ASSP study aimed at assessing 'basic skills' and incorporated items common to tests for the accompanying sample of 10 year-olds. To an outsider these tests are easy. The literacy test included questions as simple as asking students to complete the alphabet, identify the picture that corresponded to sentences such as 'He is taller than his sister', insert the correct form of verbs, articles, prepositions, and other word forms in everyday sentences, and to locate information in newspapers. Included were items asking students to find in newspapers, the letters page, a telephone number, a recipe, and a top 25 album. Most of the more difficult items were concerned with comprehension of either newspaper articles or textual passages.

The later reading tests included progressively fewer items on locating and retrieving information. The 1980 ASSP included exercises in using an encyclopedia, a table of contents, and an index. In the later studies the locating information tasks were restricted to finding in a newspaper the letters page, and Perth's maximum temperature.

As for the 1975 ASSP, the comprehension items in the later reading tests were based on newspaper articles or textual passages covering a diverse range of topics. The material from newspapers includes stories on a man rescued from a tiger attack, an armed robbery, birds trapped by dumped oil, a tug of war with a camel, a hang gliding flight, scientific explanations of floating and acne, a strike in which workers continue to work, and the flight of bees. The textual passages were on such diverse topics as,

the invention of printing, the birth of a volcano, a dispute between two motorists, a railway worker's near fatal experience with an express train, an attack by locusts, and the Manhattan project. Some of these articles and passages were also included in the original 1975 ASSP study.

As was the case for its reading counterpart, the 1975 numeracy test was easy for 14 year-olds. Questions included mathematical operations with one and two digit numbers; the reading of, scales, maps, graphs, and an analogue watch; and performing relatively simple calculations involving money. The more difficult items included fractions, estimating long division, and calculations involving volumes. Several of these items, including the simple computational items, were included in the 1980 numeracy test. The simplest computational items were not included in the later tests.

In general the numeracy items can be considered as three groups. Items in the first group deal with mathematical operations (mainly computations) with little or no practical component. These include simple operations such as addition and subtraction, and more complex operations such as long division, fractions, squares, cubes, and square roots. Those in the second group are items on practical applications of numerical skills. Examples are questions about buying things, reading scales, tables, and graphs, and calculating interest. The third group consists of items that require the application of abstract mathematical concepts. These are mainly logical and spatial problems. This categorisation is by no means absolute but serves as a useful guide to the types of questions asked.

Most items on each test were developed at the Australian Council of Educational Research and, before an item was incorporated into a test, it had undergone a field trial or had been used in another study. Once included in a test, the performance of the items was monitored in terms of difficulty and correlations with general performance on the overall test.

Establishing a Common Scale

Although student performance on common items can be compared, it is especially advantageous to use the information on performance from all the items. Common scales involving all items were constructed using a powerful statistical procedure based on the common items. It has several names including Rasch modelling, item response theory, and latent trait analysis. The procedure assumes that a student's performance on an item has two components, the student's ability and the relative difficulty of the item. It is based on the principle that items can be ranked in terms of difficulty and this relative ranking is constant between tests. Information on how students performed on both the common and unique items, and the relationships between the two, allow the construction of a common measure. The reliabilities of the common scales for reading comprehension and for numeracy were high and the total scores (a simple additive total) on each of the tests correlated highly with the relevant common scale (in excess of 0.95). It is also of interest that the correlation coefficient between reading comprehension and numeracy was 0.60 which indicates that the two measures are conceptually distinct.

A common scale established through this method has no absolute starting and end points. To produce a scale that could be readily interpreted, a linear transformation was applied so that the lowest score was assigned a value of zero and the highest score was assigned a value of one hundred.

Distinguishing Mastery from Non-Mastery

The 1975 study produced a measure distinguishing masters from non-masters. Its development involved establishing agreed-upon criteria, the construction of specific test items which satisfied these criteria, and finally the calculation of cut-off scores distinguishing masters from non-masters. Following the work of others the investigators decided to set the required level of performance as 'correct answers to 80 per cent of the population of all possible items associated with a particular task or objective'.

This report employs this concept of mastery as developed in the 1975 study as a baseline to assess mastery levels at later time points. Since the subsequent tests were more difficult, the percentage of correct items that represented mastery was less than 80 per cent. The highest score on the common scale for non-masters in the 1975 ASSP study was used as the cut-off score to distinguish 'non-masters' from 'masters' in the later studies.

The scores for the individual items obtained by Rasch modelling are on the same scale as the scores for individual students. When a student's score is the same as the difficulty for an item, it means that the student has a 50 per cent chance of correctly answering that item. For more difficult items, the student has a less than 50 per cent probability of correctly answering the item, and for less difficult items, a greater than 50 per cent probability.

Mastery in reading corresponds to a 50 per cent probability of correctly identifying the origin of the name of the volcano 'Paracutin', locating the letters page in a newspaper, and identifying the effect of bad weather in the story about an express train.

The cut-off for mastery in numeracy corresponds with about a 50 per cent probability of correctly estimating the result of 1240 divided by 29, and converting the time difference between 23:30 hours and 1:45 hours into minutes.