# The Successful School:
# A Genuine Trend or Statistical Artifice?

Rosemary Callingham
University of New England
rcalling@une.edu.au

Heather Mays
University of New England
hmays@une.edu.au

Schools are increasingly being expected to make improvements based on data about students' learning outcomes. Such an expectation implies that principals, teachers and key personnel within systems can read and act upon the data available. There is evidence, however, that many people have poor understanding of statistical information, and that many factors inside and outside the school have an effect on students' outcomes. This study considers one primary school's data from statewide testing programs. Trends across time are considered as a basis for making judgments about the school's performance in improving students' learning outcomes in literacy and numeracy.

## Introduction

In most Australian states, reform agendas have focussed on data-driven and outcomes based processes of internal school review (Hill, 1998). In Queensland, for example, a triennial review process within the School Improvement and Accountability Framework (SIAF) (Education Queensland, 2002) is based on an initial self review by a school community using data to provide evidence of the school's achievements and weaknesses. A similar situation occurs in New South Wales, where schools are expected to use data from surveys and other sources, including Basic Skills Test (BST) data about students' learning outcomes. Principals are explicitly charged with:

> Analysing school-based and system-wide student assessment data which impacts upon school priorities, targets and teaching and learning programs to improve student outcomes (Department of Education and Training, 2000, p. 3).

Such expectations raise several issues of concern. There is considerable evidence that students' outcomes are affected by a range of contextual influences in addition to school factors (Silins & Murray-Harvey, 1998; Hattie, 2002). The Victorian Quality Schools Project (Hill, Rowe, Holmes-Smith & Russell, 1996) indicated that although there is variation between schools, the greater source of variation is that between classes, and that it is individual teachers that make the difference. This finding has been supported by a range of studies (Alton-Lee, 2003;

Lingard, Martino, Mills & Bahr, 2002; Rowe & Rowe, 2002; Tomlin, Street & Baker, 2000).

Systems, however, insist that schools use data from monitoring programs, such as the BST, to determine targets and set goals for the improvement of students' learning outcomes. The question then arises as to how such data should be interpreted by school managers and the school community.

The problems that many adults have in reading statistical data are well documented. In relation to health risks, for example, there is evidence that adults cannot adequately assess data presented numerically (Schwartz, Woloshin, Black & Welch, 1997), and that causation and correlation are frequently confused (Paulos, 1995, pp. 133-139). Steen (1999) suggests that statistical information is one building block of modern-day numeracy. Misunderstandings of sampling are widespread (MacNeal, 1994, pp. 257) and teachers' misunderstanding of the meaning of "average" and lack of confidence with other statistical ideas has been reported (Callingham, 1994; Callingham, Watson, Collis & Moritz, 1995).

Gal (2002) has suggested that adults' statistical literacy requires critical thinking and a disposition to get involved with the issues. To achieve critical statistical thinking an understanding is needed not only of the social context within which the information is situated, but also of the associated mathematical ideas. Without this combination of social context and statistical skills, errors may occur in the interpretation of the data. When the data are being used for high stakes purposes, such as school accountability and improvement measures, such errors may lead to undesirable consequences.

Data from large-scale assessment programs, such as the various statewide monitoring tests, are, by their nature, complex. They include longitudinal monitoring information disaggregated by various subgroups, such as indigenous status, grade level data and, in the instance of the BST, value-added data. Data are provided in terms of mean scores, standard deviations, box and whisker plots, line plots, and at state, school and individual student level. The purpose of providing such a plethora of data is ultimately to enable teachers to plan appropriate programs to improve the learning of their students. Programs are funded by systems and schools making strategic decisions about their budget expenditure and priorities. When decisions that impact at the classroom level are made on the basis of data, there is an imperative to

ensure that interpretation takes account of many different factors, including statistical aspects that may be overlooked.

This paper reports on a study of one primary school's BST data. It was undertaken as part of a broader study, initiated by the school, into the reasons for some patterns that had been perceived in the BST data. The study considers trend data over ten years for achievement of male and female students in Years 3 and 5 literacy and numeracy tests. Different possible interpretations of the data are presented, and the implications of these interpretations, in the light of the statistical literacy needs of teachers and administrators, are discussed.

## Background to the school

Peppercorn Primary School (a pseudonym) is located in a regional centre in northern NSW. The staff is comprised of the principal, two assistant principals, 12 classroom teachers and two special literacy teachers. The school has been in operation since the 1880s. The buildings are old and run down and, in recent years, have been infested by mice and termites. Other problems have resulted from a long-term lack of maintenance including rotting walls, flooding of rooms when it rains, lack of adequate heating in winter and a lack of sufficient power outlets for computing and other facilities.

Many of the parents of the children at the school are themselves former students. The student population is largely drawn from low socio-economic backgrounds and there are a lot of single-parent families, many of which have no adult male presence in the home. Parents are encouraged to help in the running of the school and the principal supports these parents (mostly women) in taking up some form of education and/or paid employment.

In recent years, for a variety of reasons, total enrolments at the school have been steadily declining. There has been increased competition from other local schools, and this has been exacerbated by particular events in the local community. As a result, current enrolments are fewer than 300 students. Table 1 shows candidate numbers available for BST in Year 3 and Year 5 over the past ten years. Not only have the numbers dropped overall, but there also appears to be some loss from Year 3 to Year 5, as shown by the Year 3 numbers in 1996 and 2000, and the numbers for the same cohort when they reached Year 5 in 1998 and 2002. This may have an effect on the summary outcomes.

Table 1
*BST candidate numbers (where available)*

|       | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
|-------|------|------|------|------|------|------|------|------|------|------|
| Yr 3  |      |      | 57   | 35   | 44   | 35   | 35   | 34   | 34   | 26   |
| Yr 5  | 62   | 49   | 45   | 46   | 45   | 43   |      | 38   | 29   | 36   |

On the basis of the BST data, the school principal believed that the boys in the school were doing better than the girls, in an apparent reversal of current tends. The principal initiated a research project to investigate why this might be the case. A starting point was to reconsider the data available to establish whether boys were indeed doing better than girls.

## Data Analysis

The data were taken from the reports provided to the school from the BST. Although individual student data are provided to schools, these were not available for further analysis so that only summary data were used. Each year, a summary of the overall mean scores was provided for Year 3 and Year 5 Literacy and Numeracy, disaggregated by sex. Further detail was provided by other subgroups but there was a considerable amount of data missing from the school's records, and the numbers of students involved were very small. For these reasons, the initial exploration of the data was confined to the summary mean scores. Standard errors of measurement and standard deviations were not presented in the summary data.

Data from the BST are provided to schools in the form of tables of means, broken down by male and female achievement for both literacy and numeracy in Year 3 and Year 5. Table 2 shows a summary of the available data.

Table 2
*Summary of Mean Scores for Males and Females*

| Literacy | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Yr 3 M | 47.32 | 46.88 | 47.56 | 47.36 | 45.02 | 50.25 | 48.43 | 46.80 | 48.60 | 46.59 | 49.16 |
| Yr 3 F | 50.35 | 47.69 | 51.72 | 50.60 | 46.97 | 51.52 | 47.58 | 49.10 | 45.27 | 49.78 | 50.64 |
| Yr 3 | 48.73 | 47.20 | 49.91 | 49.24 | 45.97 | 50.88 | 48.09 | 47.80 | 47.33 | 48.37 | 49.67 |
| Yr 5 M |  | 53.98 | 54.74 | 52.60 | 54.55 | 54.74 | 50.82 | 55.33 | 55.25 | 55.80 | 57.18 |
| Yr 5 F |  | 54.35 | 54.54 | 54.85 | 57.14 | 56.49 | 51.93 | 57.39 | 53.40 | 53.00 | 53.20 |
| Yr 5 |  | 54.13 | 54.65 | 53.65 | 55.90 | 55.79 | 51.31 | 56.48 | 54.52 | 54.40 | 54.86 |

| Numeracy | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Yr 3 M | 49.62 | 52.65 | 50.20 | 52.74 | 48.12 | 52.57 | 48.30 | 49.60 | 48.69 | 53.64 | 52.81 |
| Yr 3 F | 50.21 | 53.73 | 54.36 | 52.81 | 48.83 | 53.54 | 47.26 | 48.60 | 45.33 | 50.55 | 52.23 |
| Yr 3 | 49.89 | 53.08 | 52.55 | 52.78 | 48.47 | 53.05 | 48.48 | 49.20 | 47.41 | 51.91 | 52.61 |
| Yr 5 M |  | 60.27 | 58.26 | 55.39 | 59.30 | 58.49 | 53.85 | 59.13 | 57.80 | 58.70 | 58.65 |
| Yr 5 F |  | 59.18 | 56.71 | 55.77 | 60.70 | 57.93 | 54.64 | 56.00 | 53.47 | 53.20 | 51.61 |
| Yr 5 |  | 59.81 | 57.60 | 55.56 | 60.03 | 58.16 | 54.20 | 57.38 | 56.09 | 56.10 | 54.54 |

On the basis of the recent BST data, the school principal believed that the boys in the school were performing better than the girls in literacy, in an apparent reversal of current trends. In Year 5 the boys' literacy scores were higher than the girls in 2002 and 2003. A similar situation was also apparent in numeracy in both grades. The principal initiated a research project to investigate why this might be the case. A starting point was to reconsider the data available to establish whether boys were indeed doing better than girls.

*Trends over time*

Because the data are Rasch (1960) scaled, and Year 3 and Year 5 results are placed onto the same scale, for each subject the data could be directly compared, both over time and for particular cohorts whose data could be tracked from Year 3 to Year 5.

Figure 1 presents ten year trend data for Year 3 students' scores in Literacy. These suggest that, although there is considerable variation from year to year, the overall trend is for Year 3 boys' literacy scores to increase slightly while Year 3 girls' literacy scores decrease. Despite the greater rate of increase, however, the Year 3 boys are still scoring less well than the girls.
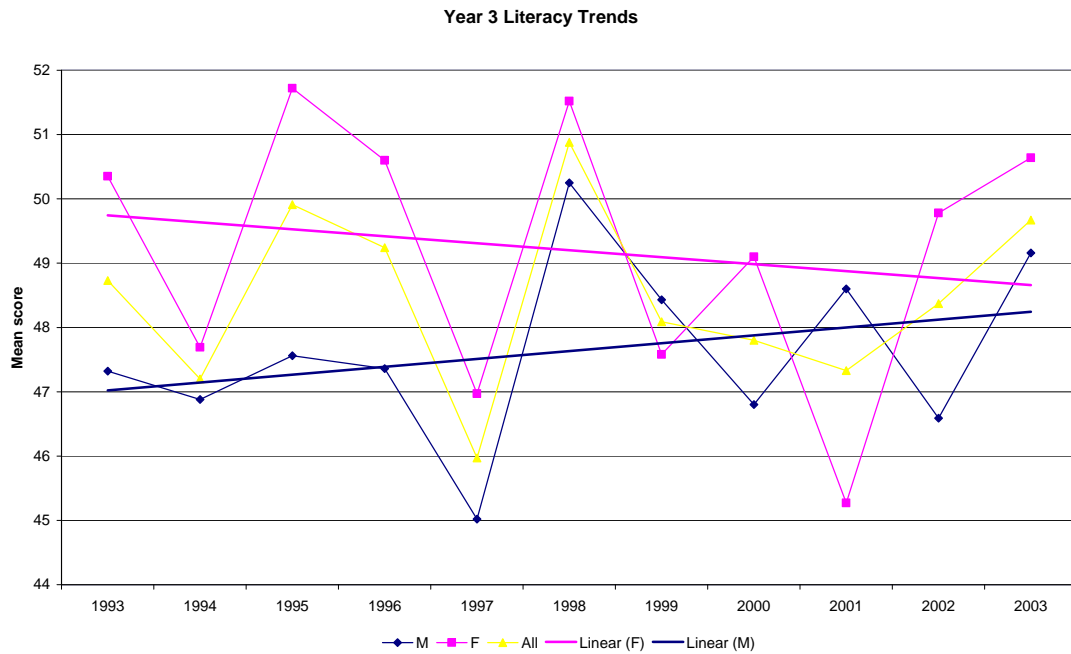
**Year 3 Literacy Trends**



*Figure 1.* Ten-year trends in Year 3 Literacy scores.

When similar data are considered for numeracy, the trend for boys to improve appears even more marked. Around 1997, boys' scores overtook those of girls and the apparent improvement has continued. These findings are shown in Figure 2.
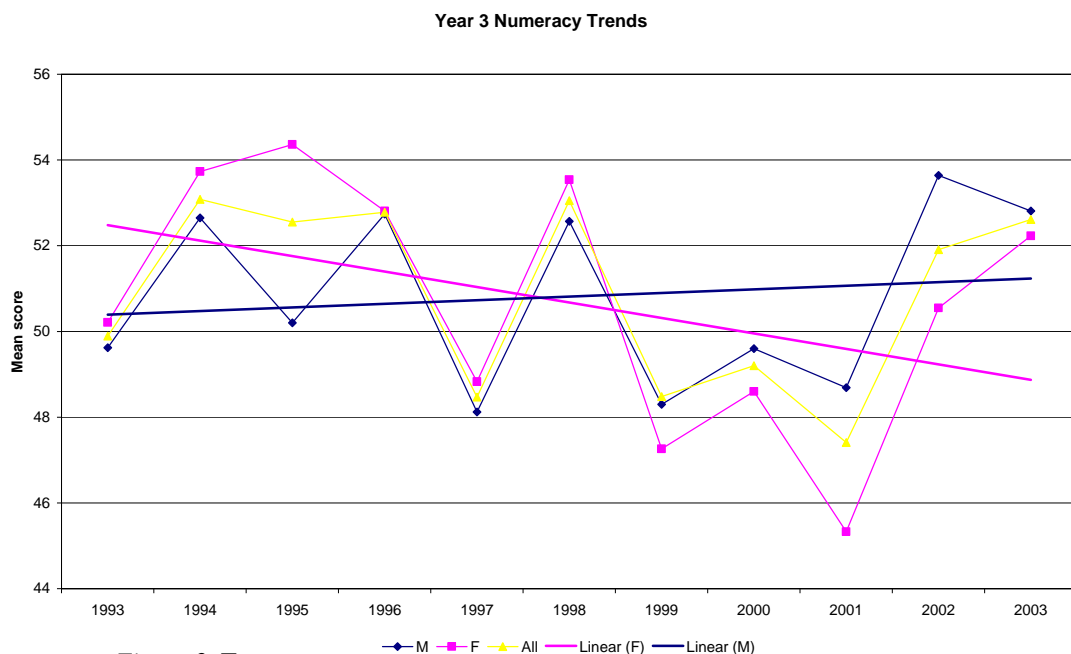
**Year 3 Numeracy Trends**



*Figure 2.* Ten-year

trends in Year 3 Numeracy scores.

The patterns in Literacy results for Year 5 show a similar trend. Figure 3 shows Year 5 Literacy trends. As with Year 3 data, the boys are scoring below the

girls at the start of the period. However, around 1999, the boys' trend score overtakes that of the girls, and continues to rise.
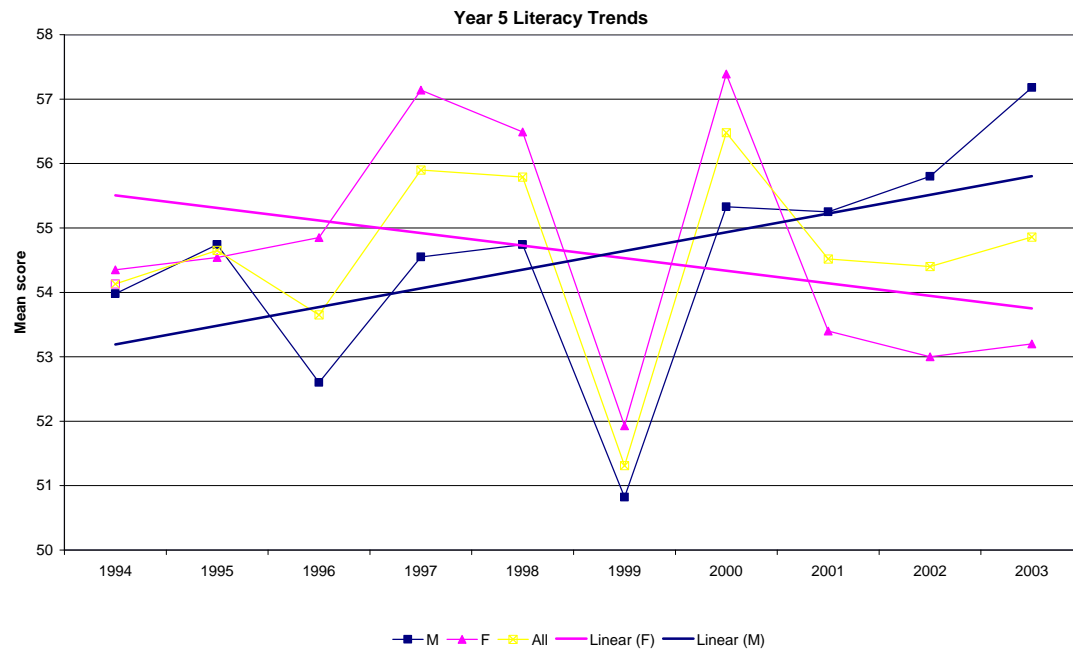
**Year 5 Literacy Trends**



*Figure 3.* Ten-year trends in Year 5 Literacy scores.

A similar, but more marked trend is shown in the Year 5 Numeracy results, presented in Figure 4. Girls' scores, in particular, show a downward trend, whereas boys' scores show some fluctuation from year to year but little change overall.
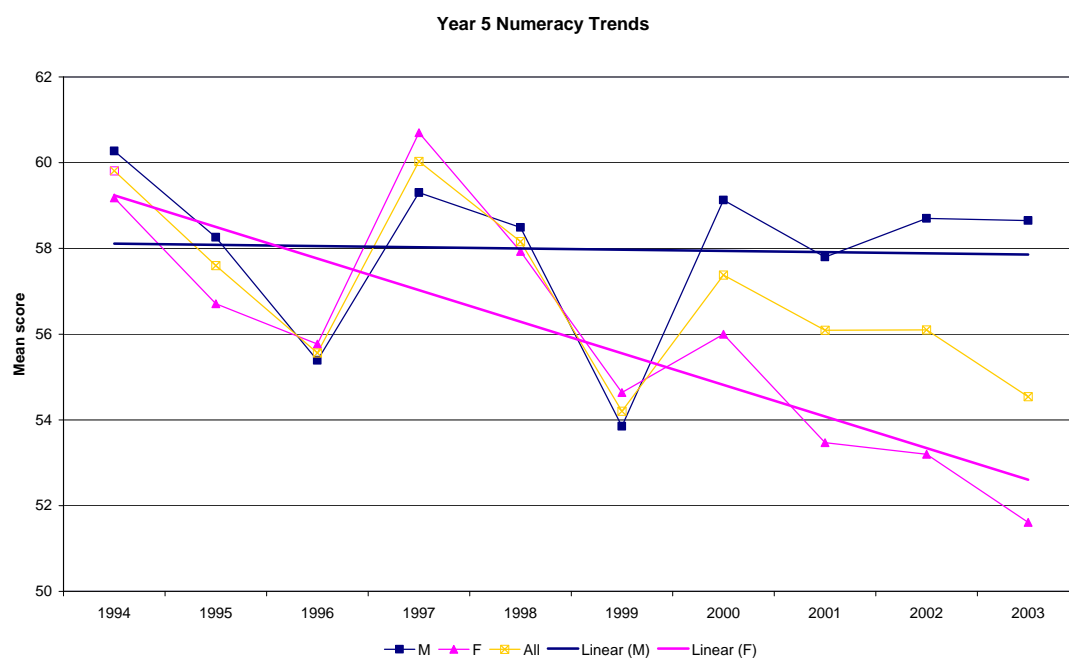
**Year 5 Numeracy Trends**



*Figure 4.* Ten-year trends in Year 5 Numeracy scores.

The visual impact of the graphical displays suggesting improving performances by boys is compelling. However, the regression equations are less convincing. Table 3 shows the trend equations for literacy and numeracy achievement over ten years for Years 3 and 5. These suggest that over that period there has been little overall change in achievement, except for some gains in Year 5 boys' literacy and a drop in both Year 3 and Year 5 girls' numeracy. In general, when the relationships are considered, the data suggest a small improvement in boys' achievement in literacy in Year 5, and a drop in girls' numeracy achievement, rather than an improvement in boys' performances overall.

Table 3.
*Trend Equations for Male and Female Achievement over Ten Years*

| Year 3 | Equation | $R^2$ |
|---|---|---|
| Male Literacy | $y = 0.12x + 46.90$ | 0.08 |
| Female Literacy | $y = -0.11x + 49.85$ | 0.03 |
| Male Numeracy | $y = 0.08x + 50.31$ | 0.02 |
| Female Numeracy | $y = -0.36x + 52.84$ | 0.17 |
| Year 5 | | |
| Male Literacy | $y = 0.29x + 52.61$ | 0.25 |
| Female Literacy | $y = -0.20x + 55.90$ | 0.10 |
| Male Numeracy | $y = -0.03x + 58.17$ | 0.00 |
| Female Numeracy | $y = -0.74x + 60.71$ | 0.62 |

*Moving Averages*

One problem with the use of simple trends is that they have momentum, which means that it takes some time to change direction. Another approach to considering the summary data available is to take moving averages. These have the advantage for a school that they can be calculated over any appropriate period of time, such as the three years of a school review process. They also provide a means of damping the effects of irregularities in time series data to reveal a smoother picture of underlying trends. Figures 5 to 8 present the three-year moving averages, with the associated trendlines, for Years 3 and 5 in literacy and numeracy.

Results for Year 3 literacy, presented in Figure 5, show that the gap between girls' and boys' performances is narrowing. The gap was initially about 2.5 score points in favour of the girls, as measured by the mean scores. By the end of the data

collection period, the gap closed to about 0.5 score points, again in favour of the girls. However, the trend lines show that while the boys have made marginal improvement ($\beta = 0.12$, $R^2 = 0.35$), the girls have shown a significant decline in performance ($\beta = -0.27$, $R^2 = 0.64$). Both groups have shown an improvement in the last two or three time periods as evidenced by the fact that the actual values are greater than the forecasted values. Although the moving average trend suggests that boys have now overtaken girls, more data are required before it can be assumed that this trend will continue.
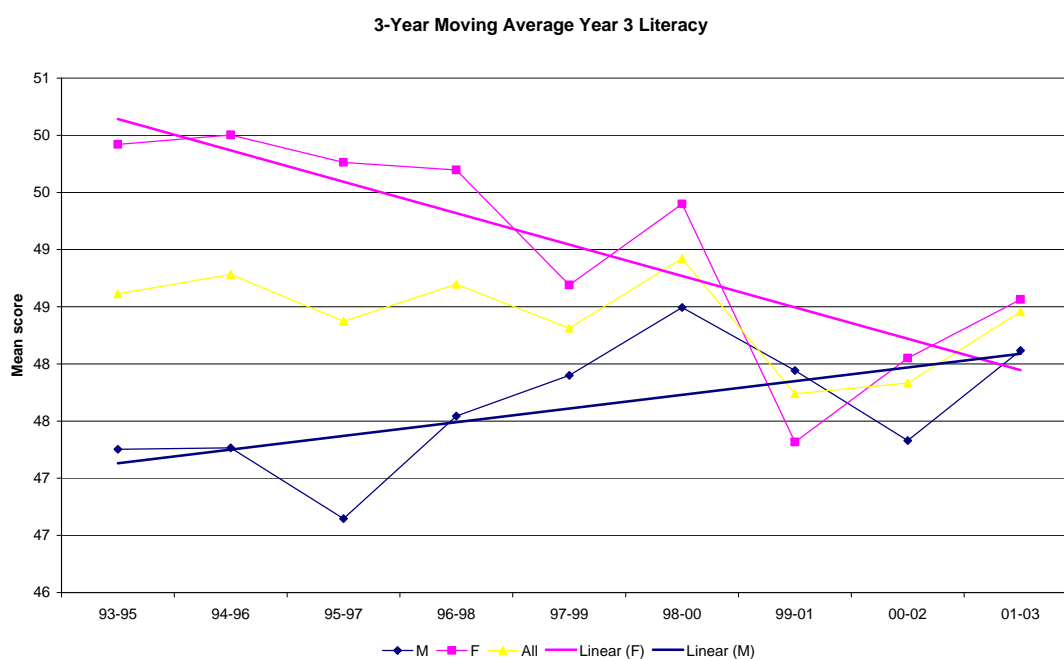
**3-Year Moving Average Year 3 Literacy**



*Figure 5.* Three-year moving averages of Year 3 Literacy data.

Similar trends can be observed for the Year 3 numeracy data, although this data set shows less variability than does the literacy data. The numeracy data for Year 3 students show that although girls outperformed boys at the start of the data collection period, this trend was reversed in the second half of the data collection period (see Figure 6). The performance of the boys remained fairly static over the period ($\beta = -0.07$, $R^2 = 0.04$), whereas the performance of the girls showed a significant downward trend ($\beta = -0.70$, $R^2 = 0.76$). Again, both groups showed an improvement in the last two time periods as evidenced by the fact that the actual values are greater than the forecasted values.
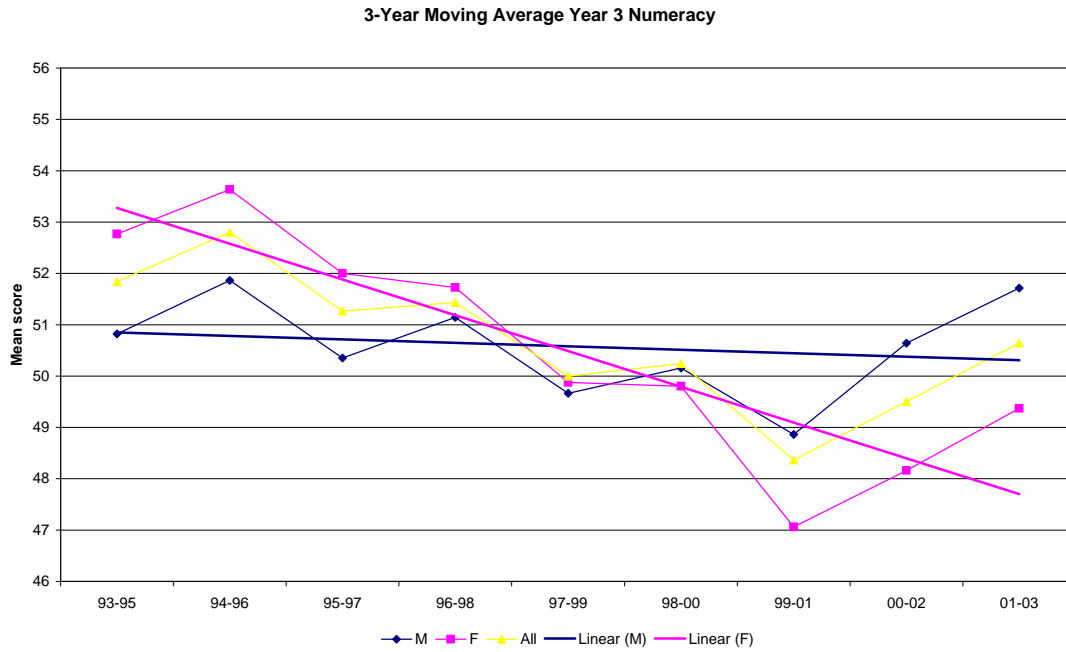
**3-Year Moving Average Year 3 Numeracy**



*Figure 6.* Three-year moving averages of Year 3 Numeracy data.

The story regarding Year 5 is somewhat different. The literacy data, presented in Figure 7, show divergence in the behaviour of boys' and girl' scores. At the start of the data collection period, the average of the girls' scores was one point higher than that of the boys' scores. However, by the end of the ten-year period, the average of the boys' scores was three points higher than the average of the girls' scores. Further, the boys showed a significant improvement over the ten-year period ($\beta = 0.28$, $R^2 = 0.50$) and their last two actual values were above the forecast values, whereas the girls showed a significant decline in performance ($\beta = -0.24$, $R^2 = 0.42$), and their last actual value fell considerably below the forecast value.
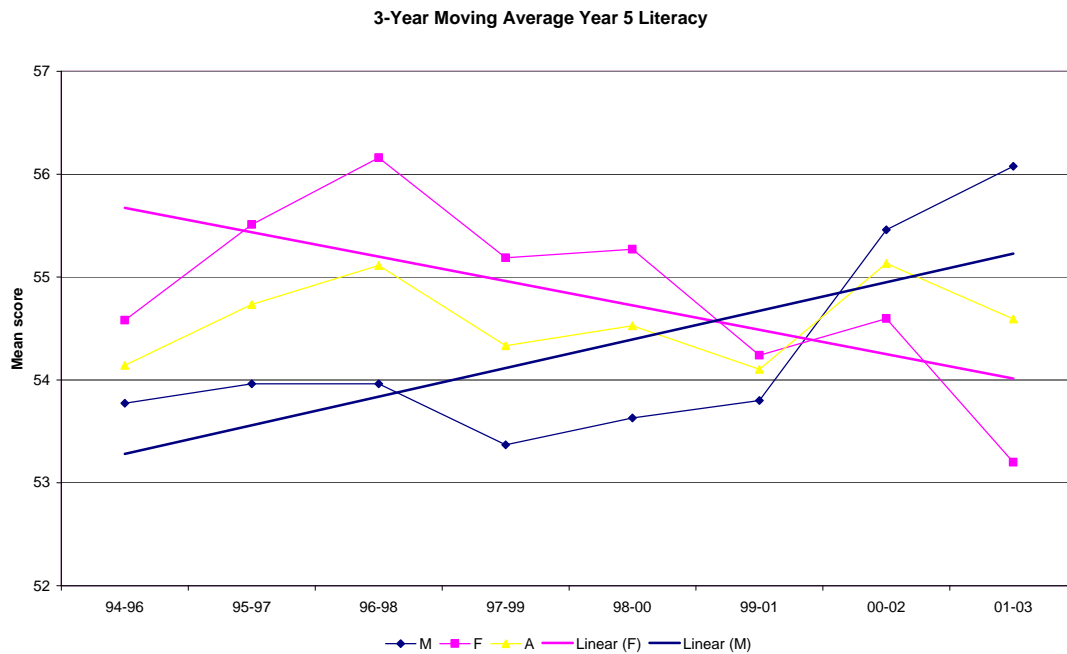
3-Year Moving Average Year 5 Literacy

*Figure 7.* Three-year moving averages of Year 5 literacy data.

The numeracy data for Year 5, shown in Figure 8, indicate that the performance of the boys has been fairly static over the entire data collection period ($\beta = 0.06$, $R^2 = 0.06$), whereas the performance of the girls has shown a steady decline after the first two time periods ($\beta = -0.72$, $R^2 = 0.80$). This resulted in a significant decline in overall performance. The last two periods show that boys are beginning to show improvement, with actual scores above the forecasts. The position is reversed for the girls where, for the last three time periods, the actual values have dipped below the forecast values. These observations about trends could be expected from the Year 3 data, in that the particular groups involved also showed similar results when they were in Year 3. However, the apparently widening gap between the numeracy performance of boys and girls is of concern.
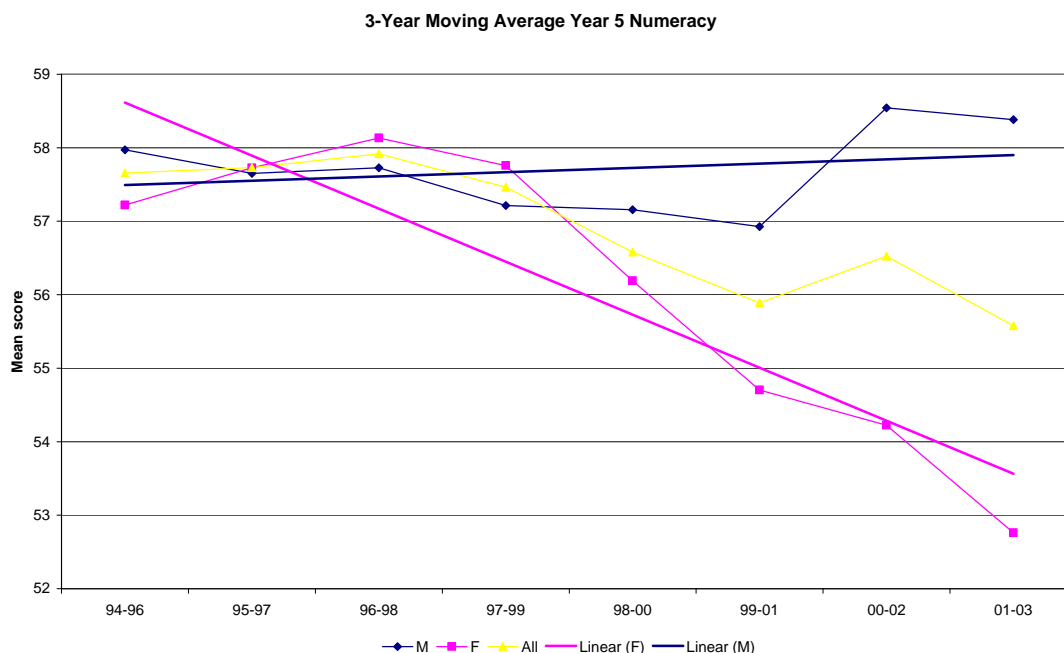
**3-Year Moving Average Year 5 Numeracy**



*Figure 8.* Three-year moving averages for Year 5 Numeracy data.

Although these data suggest something about the performance of the school over ten years, the nature of schools is that over ten years they can change considerably. For this reason, trend data over longer periods may be less useful than a consideration of performance from year to year, and the performance of individual cohorts of students as they move from Year 3 to Year 5. Because of the nature of the underlying Rasch model, the data from year to year are presented on the same scale and may thus be directly compared. The change in performance of a particular cohort from Year 3 to Year 5 may provide a more accurate picture of the schools' success at particular points in time.

*Growth Measures*

Figure 9 shows the growth measures in literacy performance for those cohorts for which data were available. The lowest overall growth occurred during the period from 1997 to 2000, for the Year 3 cohorts of 1997 and 1998. In general, the comparison between the mean growth of boys and girls showed little difference, usually around one or two score points. There are two exceptions, however. The Year 3 cohort of 1993 and the Year 3 cohort of 2000 showed much greater gains for boys. Apart from these two groups, there is little to suggest that the school is making a greater difference for boys in literacy.
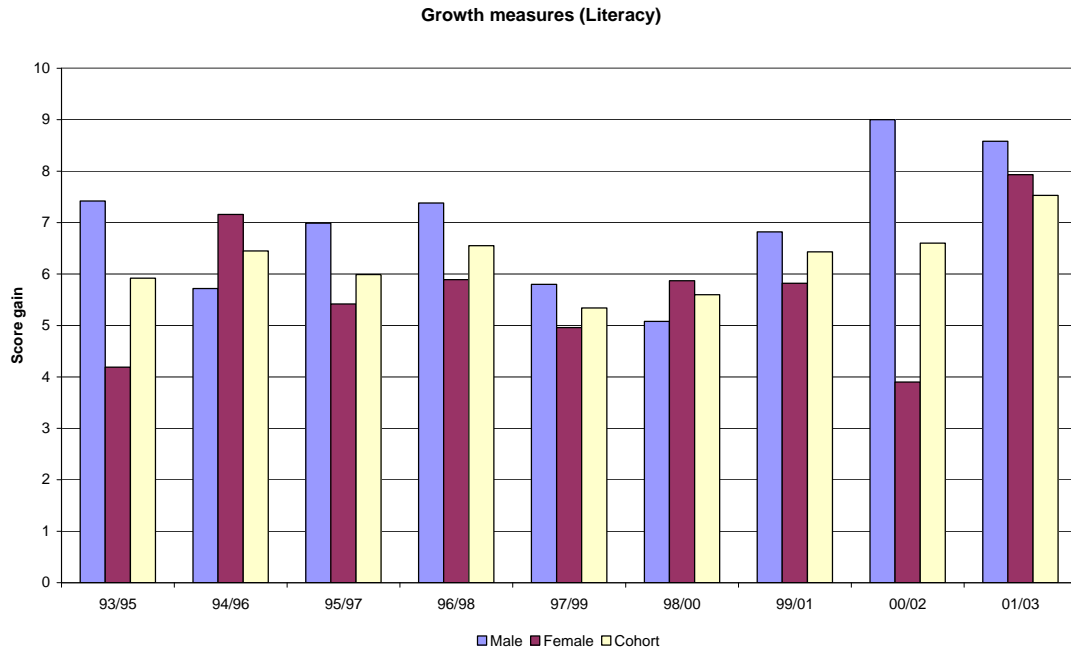
12

**Growth measures (Literacy)**



*Figure 9.* Growth in Literacy performance from Year 3 to Year 5.

The patterns for numeracy performance are rather different, as shown in Figure 10. In numeracy, the lowest growth period was from 1994 to 1996. For the Year 3 cohorts of 1996 and 1997 there was little difference in growth between boys and girls. For all other years, however, boys' performances appeared to improve more than those of girls, and this was particularly marked in the period from 1998 onwards.
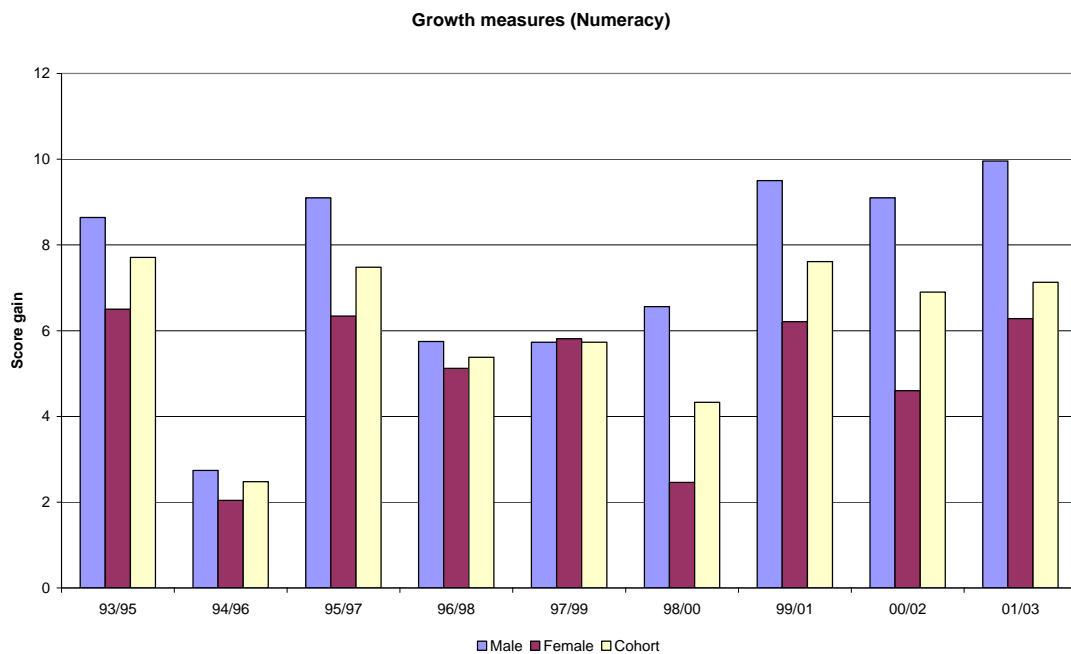
**Growth measures (Numeracy)**



*Figure 10.* Growth in Numeracy performance from Year 3 to Year 5.

13

*Year to Year Performance*

Although long term trends may have statistical importance, from an educational perspective events that affect students' performances in the short term may be more significant. The year-to-year patterns shown by the data are different for Year 3 and Year 5, although within a particular grade level, the patterns for literacy and numeracy are similar overall. In Year 3 there is a marked drop in performance in both literacy and numeracy in 1997. From 1999 to 2001, Year 3 performance in both literacy and numeracy appeared to fall below the levels of 1993 to 1995.

In Year 5, there is a drop in both literacy and numeracy performance in 1999. This is the same cohort that showed a drop in performance in Year 3, and is also a cohort that shows relatively small growth over time. In Year 5, girls' performance in numeracy dropped from 2000 to 2003, The 2003 Year 5 girls were the same group that showed a low performance in 2001, when in Year 3.

The patterns from Year 3 to Year 5 suggest that the year-to-year variation is mainly due to the particular cohorts of students. Correlations between Year 3 and Year 5 scores were calculated for each group to determine the strength of the cohort effect. The results are shown in Table 4.

Table 4
*Correlations of Year 3 and Year 5 scores*

|  | $r$ | $t$ | $p$ | $R^2$ |
|---|---|---|---|---|
| Boys' Literacy | 0.71 | 2.68 | 0.03 | 0.51 |
| Girls' Literacy | 0.82 | 3.77 | 0.01 | 0.67 |
| Overall Literacy | 0.91 | 5.76 | 0.00 | 0.83 |
|  |  |  |  |  |
| Boys' Numeracy | 0.18 | 0.49 | 0.64 | 0.03 |
| Girls' Numeracy | 0.85 | 4.27 | 0.00 | 0.72 |
| Overall numeracy | 0.65 | 2.27 | 0.06 | 0.42 |

The findings show that there is a strong cohort effect for females in both numeracy and literacy ($R^2$ = 0.72 and 0.67 respectively, both significant at the 0.01 level). Similarly, there is a moderately strong cohort effect for boys' literacy scores ($R^2$ = 0.51, significant at the 0.05 level). However, the numeracy scores for boys do not conform to this trend. In this instance, the correlation is extremely low ($R^2$ = 0.03), indicating that factors other than the cohort effect are having a major impact. These may include other factors within the school. A new principal was appointed in

1996, and this appointment was followed by several changes in staffing and organisation within the school. These factors may also have had an impact on the fluctuations in performance shown in the period 1997 to 2001.

## Discussion

The different approaches to the summary data presented above appear to indicate diverse interpretations. The particular interpretation chosen could have an impact on how the school uses its resources.

The ten-year trend data show the boys improving over that period, and could lead to the view that this school is successful with boys more than girls. The numerical data, however, are rather less convincing than the graphically presented data, showing small rates of change and small $R^2$ values, except for the variation in Year 5 girls' numeracy achievement and, to a lesser extent, Year 5 boys' literacy achievement. The Year 3 literacy data over ten years, in particular, appear to show regression to the mean, the phenomenon where, over a period of time, lower scores apparently rise, while higher scores decrease. The small variations in the trend data may be more due to statistical artifice than school effects.

The graphs generated using moving averages tell a somewhat different story. These indicate that in Year 3, girls are showing a significant decline in performance for both literacy and numeracy, whereas the boys have remained fairly static in their performance in numeracy and have made minor improvements in literacy. Similarly, the girls in Year 5 have shown a significant decline in performance in both areas, while the boys have made significant improvement in literacy and have remained fairly stable in their numeracy performance. The Year 5 data also indicate that the gap between the performances of girls and boys is growing in each of the two areas, although whether this trend will continue cannot be determined without further data.

The growth data, however, suggest that there is a difference between the achievement of boys and girls in numeracy, although there is little difference in literacy. There are fluctuations from cohort to cohort, and this may reflect aspects of the school context at the time. Such variations in the context may impact on students' performances quite strongly in the short term. This can be seen in the depressed performance of the 1997 Year 3 cohort, for example, which appeared to carry through to the same cohort's Year 5 results in 1999. This period coincides with a period of

15

transformation in the school context, with changes to organisation, different emphases in the school's programs, staff transfers, and pressures from outside the school.

The school principal indicated that there was a heavy emphasis on literacy programs within the school, and that she believed that these were making a difference to the boys. When the mean values from year to year are considered, this may appear to be a reasonable conclusion to draw. The trend data, and moving average data, however, tend to show a different picture. Over time, there has been a slight improvement in boys' literacy performance but girls' numeracy performance does appear to drop. This may create a perception that the boys are improving relative to girls, rather than that the girls' performances are dropping. The growth data also suggest that the greatest area of growth is generally boys' numeracy, rather than boys' literacy.

Two issues arise as a result of these analyses:

1. Differing interpretations, such as those presented here, may have a considerable impact on a school's disposition of its resources; and

2. Most teachers and school principals, however, have neither time nor skills to go below the surface of the data and determine the underlying patterns.

From a consideration of the data presented here, the school's focus would appear to have been literacy programs in the early years of schooling, with an emphasis on boys' achievement. The initial summary mean data appear to confirm that this focus has been successful. Further data analysis, however, suggests that the school may need to change its perspective. Girls' performances are seemingly falling, particularly in numeracy. In addition, Year 5 data, and the growth data, indicate that programs in the middle and upper primary years may need more attention.

Such findings raise questions about the statistical literacy of teachers. If teachers and administrators are to make use of the sophisticated data that are provided to them, then teachers need professional development in developing the analysis skills required. The use of a mean, for example, may be greatly affected in a small cohort by one or two exceptional children, and the interpretation of the data needs to take account of such outliers. In the instance presented in this paper, it is possible that the better performing girls, for example, are transferring to other schools between Year 3 and Year 5, with the effect that the mean BST scores are reduced for girls overall. Without further data this conjecture cannot be tested by outsiders, but teachers within

the school who were statistically literate might include considerations such as this in their interpretations.

## Conclusion

Several different statistical analyses of a small data set comprising BST results over a ten-year period for Years 3 and 5 from one small primary school have been presented in this paper. The data discussed here are typical of the data that school administrators and teachers are expected to use when planning their programs. The different analyses demonstrate the different interpretations that can be made from a set of data, depending upon the choice of statistical techniques. If teachers and administrators are to be able to make considered decisions about resource allocation, they will need to be able to engage with the statistics in meaningful ways, in order to reconcile the data with their own school's context.

## Acknowledgement

## References

Alton-Lee, A. (2003). Quality *Teaching for Diverse Students in Schooling: Best Evidence Synthesis*. Retrieved 24 April 2004 from http://www.minedu.govt.nz/web/downloadable/dl8646_v1/quality-teaching-for-diverse-students-in-schooling.doc

Callingham, R.A. (1994). Teachers' cognitive functioning in the context of questions involving the arithmetic mean. In G. Bell, B. Wright, N. Leeson, & J. Geake (Eds.) *Challenges in mathematics education: Constraints on construction.* (Proceedings of the 17th annual conference of the Mathematics Education Research Group of Australasia*,* pp. 129-136). MERGA: Lismore, NSW.

Callingham, R.A., Watson, J.M., Collis, K.F. & Moritz, J.B. (1995). Teacher attitudes towards chance and data. In B. Atweh & S. Flavel (Eds.) *Galtha.* (Proceedings of the 18th annual conference of the Mathematics Education Research Group of Australasia, pp 143-150). MERGA: Darwin.

Department of Education and Training (DET), (2000). *Leading and managing the school*. Sydney: Author. Retrieved 19 April 2004 from http://www.schools.nsw.edu.au/media/downloads/schoolsweb/edu_leadership/manage_lead.pdf

Education Queensland, (2002). *School improvement and accountability framework*. Brisbane: Author. Retrieved 19 April 2004 from http://education.qld.gov.au/strategic/accountability/docs/triennial.pdf

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*, 1-51.

Hattie, J. (2002) *What are the attributes of excellent teachers?* Paper presented at the annual conference of the New Zealand Council for Educational Research. Retrieved on 28 April 2004 from http://www.nzcer.org.nz/default.php?cPath=21_148_194&products_id=485

Hill, P. (February, 1998). *Student learning outcomes and the role of the system*. Paper presented at the

Tasmanian Education Department Teacher and School Development Branch conference, Hobart.

Hill, P.W., Rowe, K.J., Holmes-Smith, P., & Russell, V.J. (1996). *The Victorian Quality Schools Project: A study of school and teacher effectiveness. Report (Volume 1)*. Melbourne: Centre for Applied Educational Research, University of Melbourne.

Lingard, B., Martino, W., Mills, M. & Bahr, M. (2002). *Addressing the Educational Needs of Boys*. Retrieved on 8 September 2003 from http://www.dest.gov.au/schools/publications/2002/boyseducation/Boys_Report_Final1.pdf

MacNeal, E. (1994). *Mathsemantics. Making numbers talk sense*. New York: Penguin.

Paulos, J.A. (1995). Ranking health risks: Experts and laymen differ, The dyscalculia syndrome. In J. A. Paulos. *A mathematician reads the newspaper*, pp. 133-139. New York: Basic Books.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research (Expanded edition, 1980. Chicago: University of Chicago Press).

Rowe, K. and Rowe, C. (2002). What matters most: Evidence-based findings of key factors affecting the educational experiences and outcomes for girls and boys throughout their primary and secondary schooling. Retrieved on 8 September 2003 from http://www.acer.edu.au/research/documents/RoweBoysInquiry2.pdf

Schwartz, L.M., Woloshin, S., Black, W.C. & Welch, H.G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med, 127*, pp. 966-972.

Silins H. & Murray-Harvey, R. (December, 1998). *Students as a central concern: The relationship between school, students and outcome measures*. Paper presented at the annual conference of the Australian Association for Research in Education, Adelaide.

Steen, L. A. (1999). Numeracy: The new literacy for a data-drenched society. *Educational Leadership, 57*(2), pp10-13.

Tomlin, A., Street, B. & Baker, D. (September, 2000) *Leverhulme Numeracy Research Programme: Home/school Relations and their Significance for Effective Learning and Teaching of Numeracy*. Unpublished paper presented at the British Educational Research Association Conference. Cardiff: