The statistical literacy needed to interpret school assessment data

Robyn Pierce, Helen Chick
University of Melbourne

Abstract

The growth in state-wide and national school testing in areas such a literacy and numeracy has led to the generation and dissemination of reports containing graphs and tables illustrating school and individual performance. These are intended to inform teachers, principals, and education organizations about student and school outcomes, so that the data can then be used to guide change and improvement. This aim relies on the assumption that the data are readily able to be interpreted by users. Given the complexity of the information being conveyed, it is of interest to determine the critical statistical skills needed to be able to make sense of such data. This paper, therefore, examines the statistical literacy required interpret the graphical presentations of school assessment data for the NAPLAN testing process. The work of Curcio (1987) and Pfannkuch (2006) is used to identify different levels of interpretation, and the implications for helping users to make better use of such data are discussed.

Introduction

In education—as in other workplace sectors—quality control, accountability, and forward planning may be informed by examination of statistical data. The technological revolution has facilitated the collection, analysis, and sharing of vast quantities of data. National literacy and numeracy testing has become an established part of the education profile in the UK (Children, Schools and Family Committee, 2008), USA (Baker, 2007), and Australia (Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA), 2007). Such tests are advocated to identify students' educational needs, promote data-driven decision-making when planning teaching practice, and to promote schools' accountability to their students and their funding authority. For example, Australia has developed a *Measurement Framework for National Key Performance Measures* (MCEETYA, 2007) for monitoring and advancing quality outcomes from school education. Integral to this framework are the set cycles for assessment and data collection. Extensive databases, incorporating assessment results and socio-economic data from every student across Australia, allow complex linking and analysis of this information. The government's claim is that:

> Literacy and numeracy assessments provide rich data about individual student performance and assist teachers to plan learning activities for students. They also enable schools to develop a more objective view about the performance of their students compared to those in other schools and in relation to state-wide standards. (MCEETYA, n.d.)

Governments expend significant resources on collecting such data from the education sector with the intention that this should inform planning and practice. The National Assessment Program - Literacy and Numeracy (NAPLAN), for example, involves students in Years 3, 5, 7 and 9 from all States and Territories. In Victoria, approximately 260,000 students from all Government, Catholic, and Independent schools participate in the yearly NAPLAN testing, coordinated by the Victorian Curriculum and Assessment Authority (VCAA, n.d.).

Planning and conducting tests, collecting and analysing data, and producing data reports is a time-consuming and expensive process. It is therefore important that the outcomes of the process are used in ways that will benefit students. There are several scales at which these data have application, with consequential demands on the statistical knowledge of the relevant users. At the systemic level, when governments or individual schools might look at across-system or school performance, it may actually be necessary to employ a statistical expert in order to conduct a detailed analysis of the data and highlight issues for which schools may be expected to be accountable. Principals of schools need to engage with the data to identify whole school issues; they too might employ an expert for in-depth analysis, but ideally they need to be able make sense of the data for their school and determine the applications. For teachers, the data can be examined at the levels of the individual student and also the class. Identification of students within the data who are performing below national benchmarks is normally a simple task; what is usually more complex is the analysis of class data in order to inform planning for teaching. Teachers are a well-educated work force but most will not have any formal training in statistics and many will have only limited experience with analysis or interpretation of quantitative data.

Indeed, less than 20 years ago, most adults would have experienced schooling with little emphasis on statistics. It was only in 1991, with the publication of *The National Statement on Mathematics for Australian Schools* (Australian Education Council, 1991), that "Chance and Data" became significantly and formally recognised as part of the entire P-12 curriculum in Australia. In the ensuing few years the states produced their own curricula (e.g., Board of Studies, 1995) based, in varying degrees, upon this document. This increased emphasis on statistics has remained in versions of curricula since that time (e.g., Victorian Curriculum and Assessment Authority, 2008). The relative recency of this expansion of statistics into the school curriculum means that teachers over the age of 35 at the time of writing (2009) may not have had much experience with statistical thinking activities. The push for statistical literacy is more recent still, implying that a significant proportion of the general teaching workforce may not be well equipped to undertake the sort of data interpretation necessary to make best use of school assessment data like NAPLAN.

The purpose of this paper is to consider the statistical demands made of principals and teachers who need to examine such data. It presents a preliminary analysis of the statistical literacy required to read and interpret reports provided by the Victorian NAPLAN data service. First we provide some background from literature, and then conduct an analysis of three of the graphical report formats provided to schools. We conclude with a discussion of the implications of these findings.

Background

Reading and interpreting statistical reports requires more than conventional literacy, it requires *statistical literacy*. The scope of this term encompasses sufficient knowledge and understanding of numeracy, statistics, and data presentation to make valuable use of quantitative data and summary reports in a personal or professional setting (Ben-Zvi & Garfield, 2004; Watson, 2006). It need not imply the capacity to actually conduct sophisticated statistical tests, but it does include the ability to question the sampling techniques used to collect the data, to interpret possible explanations and consequences of the data, and to identify limitations in the data and the conclusions.

Statistical graphs are commonly used in data reports to convey an image of the data. There is a common understanding that graphs provide a picture of the data that may convey a key message. Good graphical communication requires good graph design. Tufte (2001) summarises six principles of such design. These include using physical representations that are proportional to the actual numerical quantities represented; showing data variations rather than design variation, and including clear labels and explanation of the text. Conventions for simple designs, which keep the visual coding transparent, have been established over time. Details of the design of "standard" graphs are included in most introductory statistics text books and even school mathematics texts. Such standard graphs include bar graphs, boxplots, and stacked dot plots.

The producers of school reporting data, perhaps reasonably, assume that it is sensible to use standard graphs in data reports. If a graph is clearly constructed, without distortion, then it might be expected to convey a message; however, reading such a message requires knowledge of graph interpretation. Some of the interpretation may be straightforward, but other aspects—particularly more serious "informal inference"—require some specialised skills and knowledge. This informal inference, which goes beyond straightforward reading of data values, is critical if the data are to be used as a basis for decision-making.

Curcio's 1987 study of graph comprehension in Year 4 and Year 8 students highlighted the ideas of "reading the data", "reading between the data", and "reading beyond the data". The items that she used in her research allowed her to examine each of these three aspects of data interpretation. "Reading the data" is concerned with the capacity to read literally the direct factual information on the graph: graph title, axes, and single data values, for example. Much of this information is what a practised graph reader gathers quickly when first examining a graph: reading the title, checking the axes for variable names and ranges of values, and perhaps a quick check of one or more specific data points. This kind of activity might be termed "getting oriented to the graph". "Reading between the data", on the other hand, requires attention to two or more data points on the graph, often for comparison purposes. This aspect is necessary for starting to "tell stories" about the data and make sense of what the data might be able to say. Finally, and significantly, is the capacity to "read beyond the data". Curcio talks about extending, predicting, and, most importantly for our purposes, inferring from the data. This is where interpretative skills are required, and deeper statistical reasoning.

One of the standard statistical graphs used by the NAPLAN data service is the boxplot (see Figure 2). Pfannkuch (2006), drawing on the work of others, highlighted that for boxplots, attending to several critical components is essential for reading between the data, prior to fully interpreting the data. She points out that the middle part of the data usefully characterises the group and that comparing boxplots should incorporate comparison of these middle parts of the data. Components of the spread, including the interquartile range and the whiskers, allow a determination of the variation within and between boxplots, and give a sense of the 'shape' of the data (e.g., if it is symmetrical or skewed). Furthermore, it may be appropriate to compare equivalent and non-equivalent key values; for example, a difference between data sets may be evident if the 25^{th} percentile for group 1 is above the 75^{th} percentile of group 2.

The work of Pfannkuch (2006) provides a comprehensive analysis of the kind of reasoning that might generally be involved at the "reading beyond the data" stage. Her report, part of a larger study, found that informal inference is a complex matter. As part of that study five statisticians used their knowledge of formal inference in interpreting boxplots to identify what is required for successful informal inference (i.e., to be able to draw considered conclusions from the data, and understand the variation within it). Based on their experience and expertise they identified four key elements for informal inference: comparing centres (e.g., mean, median), comparing differences between centres while simultaneously taking into account variation, checking the distribution of the data (e.g., attending to its shape, presence of outliers, clusters of data values), and considering sample size effects. They also noted that there had been no research on how teachers reason when comparing boxplot distributions and no definitive accounts reporting on the process of going about drawing informal inferences.

Principals and teachers working with reports, such as those provided by the NAPLAN data service need to be able read between and beyond the data in order to make informal inference. It may be seen from Pfannkuch's summary that this is not a trivial exercise. In this study we examined three demonstration reports to identify more precisely the statistical literacy required at each level: reading the data, reading between the data, and reading beyond the data.

Methodology

The Victorian NAPLAN data service supplies assessment data for schools. They also provide a data set for a fictitious school that can be used for demonstration purposes; the reports on these data were analysed for this paper. The data, summarising individual, class, school, state, and national NAPLAN results, are presented in graphs and tables, according to the type of report selected. The range of reports includes analysis of individual items on the NAPLAN tests, data on content areas (e.g., reading; spelling; number; measurement, chance and data), and school summary data with state and national comparisons. For the purposes of this report, only the graphical representations were analysed. The most common representations shown in the NAPLAN data reports are boxplots and bar graphs (see Figures 1 to 3, which use data for the demonstration school).

The graphs were analysed to determine the statistical literacy skills required to interpret them. A first pass on the analysis was completed by the first two authors, who have good to high statistical literacy skills, in the company of an experienced senior mathematics teacher also having good statistical literacy skills. We looked at each of the graphs in Figures 1 to 3, with which we were not initially familiar, and made explicit amongst ourselves the aspects that we were attending to and the skills we were using as we read and interpreted the graphs. We made notes about the skills that were required. At the time of this initial analysis we did not have Curcio's (1987) framework specifically in mind as an analytical tool, but instead merely recorded each component that we noticed. We were, however, mindful that we were seeing several levels of statistical literacy in this process, and can, post hoc, note that Curcio's framework was evident. At the basic level, there were the skills needed to simply read the directly accessible information on the graph (Curcio's "read the graph"), and then to be able to make comparisons across data values (Curcio's "read between the data"). Finally we noted that in order to understand, for example, a school's performance it was necessary to "read beyond the data". Having identified all the aspects of statistical literacy that we felt we had used in interpreting the graph the first author, in consultation with the second author, then classified each skill under one or other of Curcio's categories.

Since the graphs here are more sophisticated than those used in Curcio's (1987) research, we expected that the "read beyond the data" stage would be more complex for this situation. We were interested to see which of the aspects identified by Pfannkuch (2006) as necessary for informal inference were required. Having categorised the statistical skills as described above, we then examined those in the "read beyond the data" category to determine if they corresponded to Pfannkuch's informal inference requirements.

Results

In the results that follow, we consider the statistical literacy required for interpreting the data as presented in the three different graphical reports produced by the NAPLAN data service. In each case, the results are classified according to Curcio's categories.

Graph 1: Assessment Area Report

The first graph, shown in Figure 1, is the Assessment Area Report and presents data about different components of literacy and numeracy. The skills required to read this graph are outlined in Table 1. Unlike boxplots, where the box and whiskers capture several points of information, the only part of the "bar" of a bar graph that provides information is the endpoint. Although Pfannkuch's analysis of the skills necessary to interpret boxplots is not applicable here, there are parallel skills needed, such as comparison of endpoints (instead of centres), comparison of differences between endpoints (with a need to take into account variation), and a need to consider sample size effects.

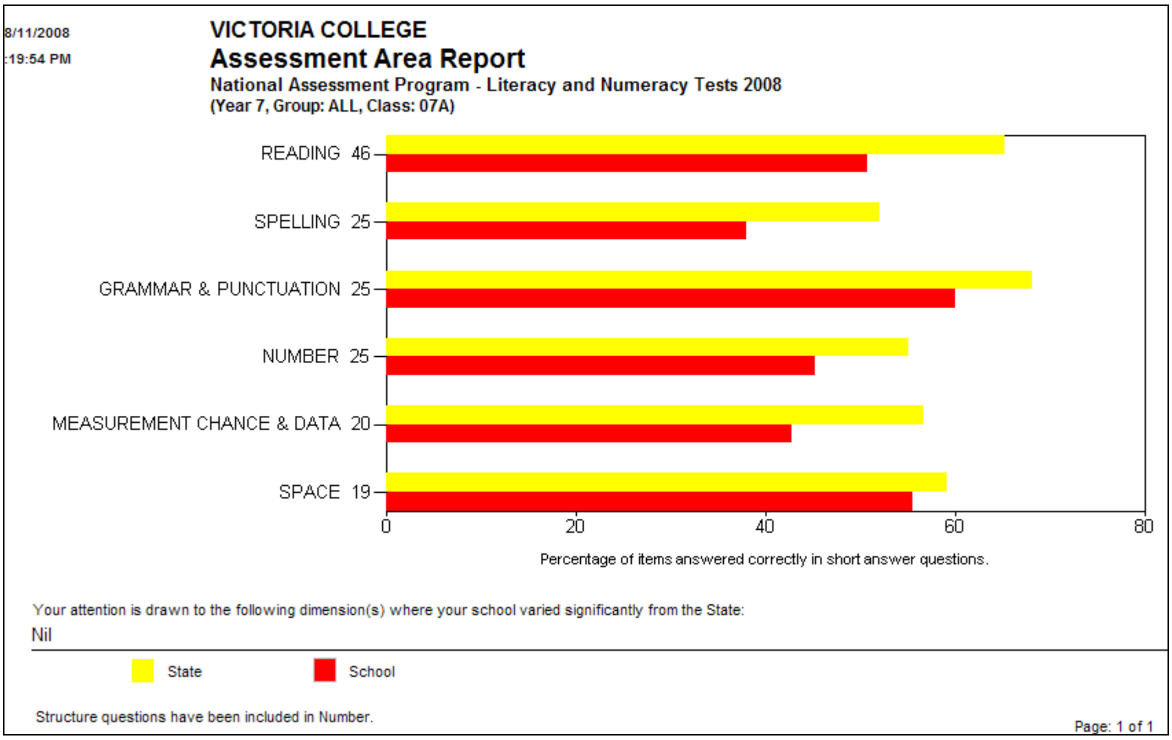


Figure 1 Demonstration school's Assessment Area Report from NAPLAN data service.

Table 1
Data reading skills required to interpret Figure 1.

Reading the data	*Reading between the data*	*Reading beyond the data*
Read the scale (e.g., note that it does not go to 100) Read individual data points as indicated by bar values Read labels Read the legend Understand that the left axis is categorical not numerical	Compare the magnitude of numbers Know how to compare values (proportional/ absolute comparisons)	Consider the number of questions when comparing values Understand that it is not appropriate to draw conclusions about differences between small groups and state or national values. Know which numbers to compare for interpretative purposes Reconcile how big the difference is between the state and the class in terms of real numbers Hence, determine when differences are numerically important (perhaps not “statistically significant” in a formal sense) Determine when are differences educationally significant. Implicit understanding about variation and its implications Understand features of this graph in context

The final skill mentioned in “reading beyond the data” involves being able to understand features of this graph in context. In the case of this graph this would require knowing how to interpret the statement that “your school varied significantly from the State” (what does “significant variation” mean?). It is also necessary to be able to recognise that although the horizontal axis is claimed to show the “percentage of items answered correctly in short answer questions” it seems to mean the *average* percentage of items answered correctly by the students in the school. This use of percentage is different from the way it is used in some of the other NAPLAN tables. Finally, contextual knowledge is required to understand the meaning of the numbers on the end of the labels for the subjects.

Graph 2: School Summary Report

The second graph, shown in Figure 2, is the School Summary Report. This presents data about the distribution of the performances of students from a school on different components of literacy and numeracy and compares these with state and national outcomes (the national outcomes were not available when Figure 2 was generated). The skills required to read this graph are outlined in Table 2. The use of a boxplot (or box-and-whisker plot) provides more information to the user but adds to the complexity of interpretation. The graph reader now has to attend to multiple facets of the boxes in order to make meaningful comparisons and understand the data. This includes the skills mentioned by Pfannkuch (2006): comparing centres (medians), attending to distribution or shape (i.e., the components of the box and the whiskers), and considering sample size effects (e.g., are differences between school and state data significant).

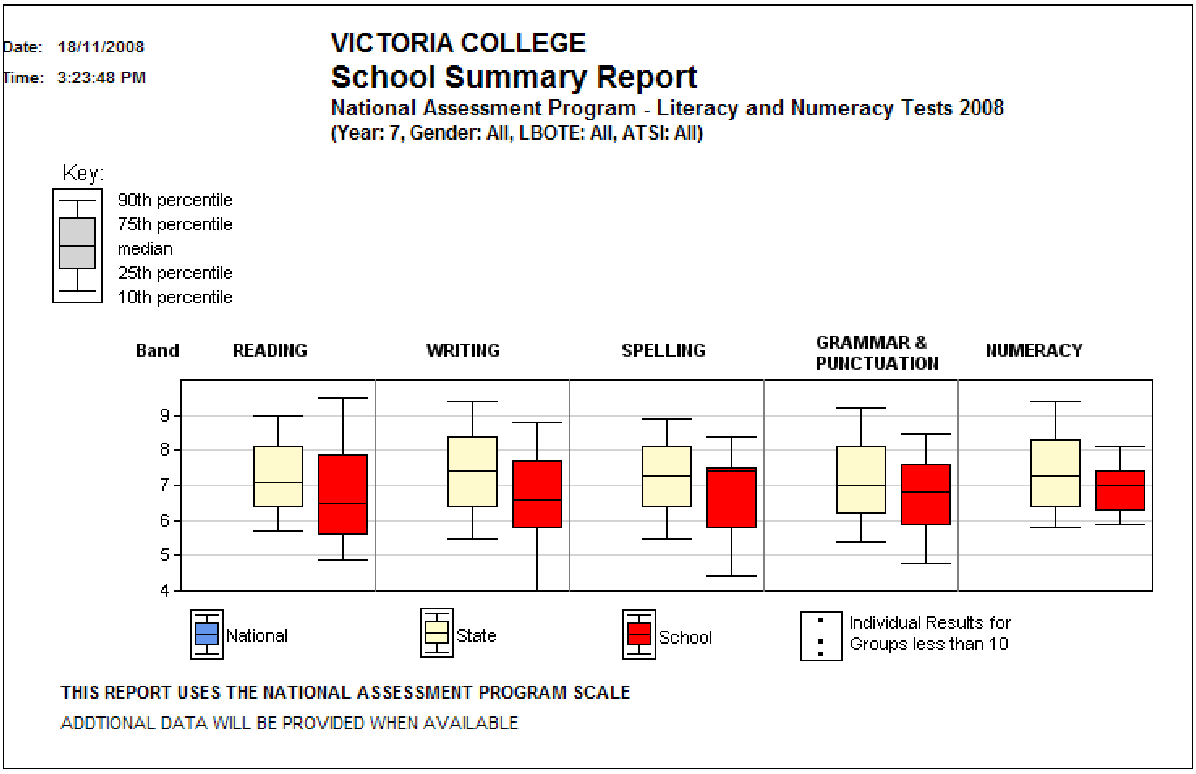


Figure 2 Demonstration school's School Summary Report from NAPLAN data service.

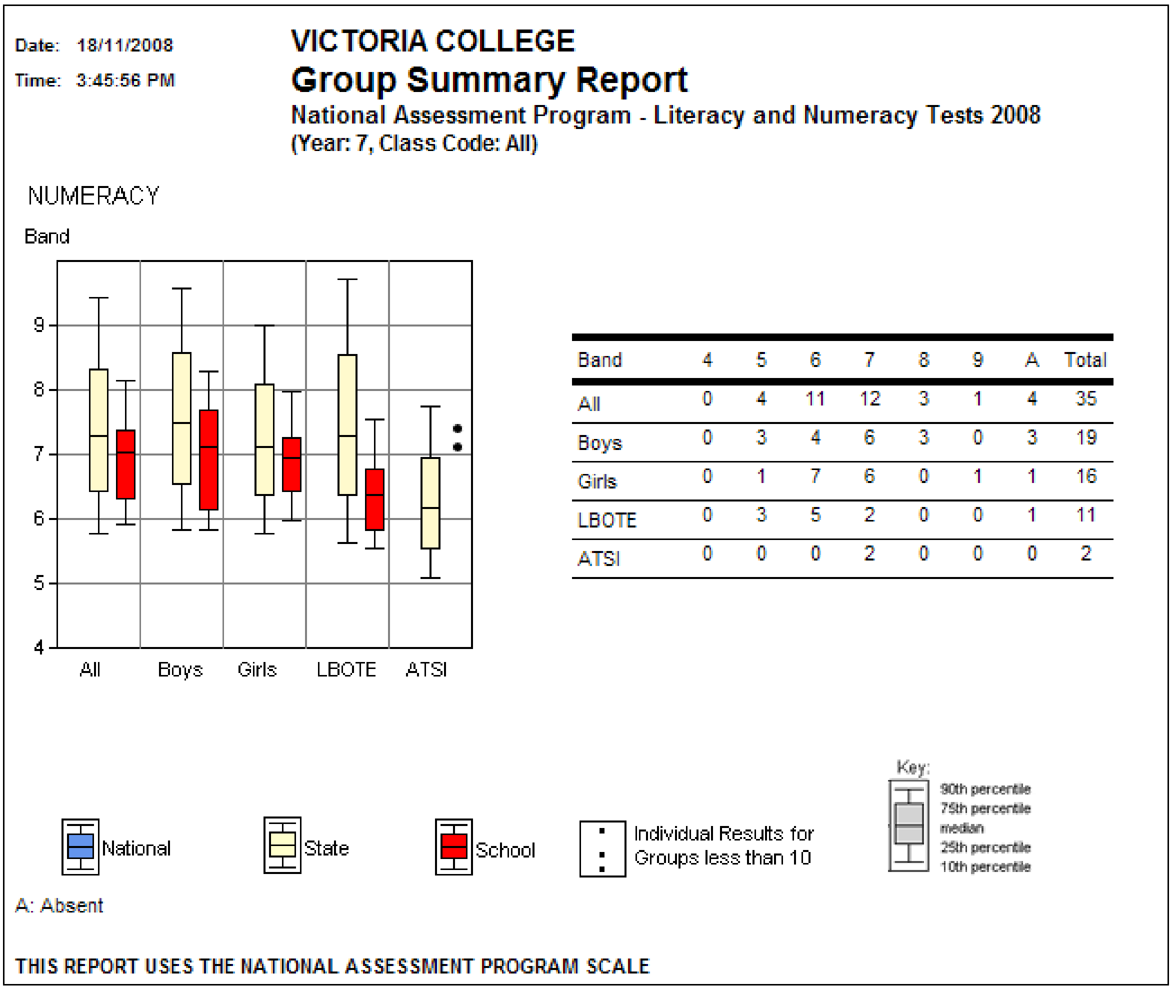
Table 2
Data reading skills required to interpret Figure 2.

Reading the data	*Reading between the data*	*Reading beyond the data*
Read the key (understand that this boxplot is different from textbook plots) Read values against the band-axis Read labels Read median value, including when the median and one or more of the quartile values coincide (see Spelling) Read percentile values Read the value for a 'whisker' including when one when falls on the boundary of the graph (see Writing) Understand that the width of the boxes is irrelevant Understand that results for groups less than 10 are displayed as points	Understand the meaning of median in context Understand the meaning of percentiles in context Understand the implication for the shape of the distribution of the data when the median is very close to or at one end of the box Understand about skewed data Know how to compare both within and between boxplots (which involves multiple comparisons (inter-quartile range, medians, etc) Comparing magnitude of numbers	Understand that it is not appropriate to draw conclusions about difference between small groups and state or national values. Understand and interpret the absence of shown outliers Determine when are differences numerically significant (perhaps not "statistically significant" in a formal sense) Determine when are differences educationally significant Implicit understanding about variation and its implications Understand features of this graph in context

In this case, the demands of "reading between the data" are heavier, because of the complex information presented in a box. The school's spelling data in Figure 2, for example, is skewed, with most of the top half of the students falling in a narrow range between BAND 7.3 and 8.3 or so; whereas the lower half of the students are spread over a much wider tail. As in Table 1, the requirement to understand features of this graph in context demands knowledge of contextual information beyond the details presented on the graph itself. In this case it is necessary to understand the meaning of "BAND" on the left-hand scale, and its connection to the "National Assessment Program Scale" mentioned at the bottom of the chart. It is also necessary to understand whether BAND was originally categorical but treated numerically to get values, or if it was numerical (i.e., can a single student get a BAND score of 7.3 or is he/she just categorised in BAND 7). Given that Figure 1 has Number, Measurement Chance & Data, and Space as separate components, it is important to understand how the "numeracy" construct arises from these.

Graph 3: Group Summary Report

The final graph, Figure 3, is the Group Summary Report. This allows comparisons among subgroups of students from a school with state and national results (e.g., by gender or language background). The skills required to read this graph are outlined in Table 3. The comments made for Graph 2 about interpreting boxplots are relevant here.



Band	4	5	6	7	8	9	A	Total
All	0	4	11	12	3	1	4	35
Boys	0	3	4	6	3	0	3	19
Girls	0	1	7	6	0	1	1	16
LBOTE	0	3	5	2	0	0	1	11
ATSI	0	0	0	2	0	0	0	2

Figure 3 Demonstration school's Group Summary Report from NAPLAN data service.

Interpreting this graphic requires the same skills and understanding as Figure 2, together with the abilities described in Table 3.

Table 3
Data reading skills required to interpret Figure 3.

Reading the data	*Reading between the data*	*Reading beyond the data*
Understand the acronyms Understand what the subsets mean Understand symbols used (e.g., A is used for absent) Understand that with small numbers individual students may actually be revealed (see ATSI)	Know how to connect the table and the boxplot	Need to understand that you can't "add" the boys and girls boxplots Need to understand how the size of the subsets will affect the overall results (e.g., could the good performance of the class be due to a small number of high-performing girls (this case is not shown)) Read and interpret the table Understand that it is not appropriate to draw conclusions about differences between small groups and state or national values. Understand features of this graph in context (e.g., derivation of the "numeracy" construct)

Discussion

The graphs provided by the NAPLAN data service largely comply with the guidelines for good graphics suggested above by Tufte (2001). They are presented simply, without unnecessary distracting ornamentation. It should be noted, however, that the boxplots with whiskers extended to the 10^{th} and 90^{th} percentiles are not standard, and the reader must be aware of the range of the scale used in each report. Nevertheless, despite this variation—which is clearly indicated by the key supplied with the data—the graphs are good examples of appropriate summary data representation.

The statistical literacy required to "read the data" for these graphics has three main components. First, the reader needs to be aware of and understand the key components and labels on the graph: title, scale labels, and scale. In Figure 1, for example, the left hand side scale is categorical while the horizontal scale is numeric. Second, the reader needs to note the legend or key in order to know what values they are reading. For example, in Figure 1 they must note which bar refers to the State and which to their school. In Figure 2 the key to the boxplots indicates that these are not drawn in the way that is typically presented in school and statistics text books where it is conventional for whiskers to be extended to one-and-one-half times the length of the box, with points beyond this indicated as outliers. In the NAPLAN data service boxplots the whiskers extend to the 10^{th} and 90^{th} percentiles respectively, and outliers are not shown. This leads to the third element of statistical literacy required by the reader: knowledge and understanding of the values that may be read from each type of graph. For example, in Figure 1 the only data values that can be read from this graph are the values associated with the endpoint of each bar. In Figure 2 there are five key values that may be read from each boxplot, associated with the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles.

"Reading between the data" requires readers to have some further understanding of the graphs' components. For the bar graphs they need to consider the meaning of absolute and proportional differences, and for boxplots they need to be able to recognise and understand the implications of the shape of the distributions. Importantly, reading between the data involves paying attention to both the location of the centre of the distribution, indicated on the boxplot by the median (50^{th} percentile) but also the spread, indicating the variability in results within the school group. In Figure 2, for instance, reading between the data for boys and girls shows that the centre of the distribution of girls' results was below that of the boys' but the variation in the boys' results was greater.

"Reading beyond the data" requires not only an understanding of the key components and structure of the various graphs but also an understanding of both the types of claims that it might be appropriate to make based on informal inference and also of the context within which to interpret these claims, such as the school, the students, and the tests used to collect the data. This is where the critical thinking that is at the heart of statistical literacy must be applied. The reader, at this stage, is making inferences, albeit informal, based on the data. If such data reports are to be of value for decision making and planning then after the two previous stages of identifying and comparing values the school principal or teacher must consider the implications of these numbers within the context of their school. For example, in the Assessment Area Report (Figure 1) the reader might think "We scored below the state average in Measurement Chance & Data; is this something we need to address in our teaching or is there some other explanation for this result?" Other explanations may relate to group size or local events that impacted on the testing. There is a need to reconcile how big the difference is between the state and the class results in terms of real numbers. For example, in Figure 1 perhaps the State result was 65% against a school result of 61%. It is important to recognise that in a school group of 50 this would represent only a difference of 2 students and could be accounted for by other factors such as test anxiety or other factors on the day of the test. "Reading beyond the data" for the boxplots of Figure 3, for example, means paying attention to the differences between groups relative to the spread or variability within each group, leading to questions about whether the difference in performance of the LBOTE students (with a Language Background Other Than English) in comparison to the whole school cohort is significant. It also requires recognition that few, if any, conclusions can be drawn about the ATSI (Aboriginal and Torres Strait Islander) students as a group because only two were involved in the testing.

It should also be noted that the four key elements identified by Pfannkuch (2006) and her statistician colleagues were, indeed, critical for a full understanding of the data. In order to determine how the school's results compared with state and national results—or to conduct any of the other important comparisons among the results represented as boxplots in Figures 2 and 3—it was necessary to attend not only to the centres (medians) of the boxplots, but to the distribution indicated by the components of the boxes and the whiskers, with consideration of variation necessary in order to determine the magnitude and significance of any differences. The influence of sample size on the significance of differences is particularly critical for informal inference, and hence for making valid conclusions about when the results indicate causes for concern or satisfaction. Without these skills it is likely that teachers will incorrectly focus on differences that are not statistically significant or dismiss those that are.

Conclusions

The analysis of the graphs has highlighted that relatively complex and critical thinking is required to make sense of the data, particularly in the light of their context. It should not be assumed that everything that the data might reveal is evident just through the graphics alone, nor that someone with no statistical literacy could read and make sense of these reports. The picture may, indeed, be worth a thousand words, but its subtle nuances require a careful reading to get the full story. Having said this, it is not the case that the skills required to read the graphs are particularly difficult either. The user needs to be "fluent" in reading the technical aspects, which is essentially the scope of "reading" and "reading between" the data: the capacity to attend to scale and labels, for example, together with reading data values as represented on the graph and making comparisons among them. These skills, if not already understood, are relatively straightforward and could be taught at a simple hands-on professional development session.

More complex are those skills needed to "read beyond the data". This requires understanding of informal inference, and the statistical literacy to know what level of importance to place on the various results. Principals and teachers need to be able to determine which differences should be given attention, and then use knowledge of their own context to identify appropriate responses and strategies, or provide alternative explanations for the outcomes. This demands a capacity to question the data, to be aware of sampling issues, and the kind of relational thinking that makes it possible to keep track of how one variable may affect another in a variety of circumstances. These skills are likely to be more difficult to develop, but teachers' own familiarity with their school context might mean that it is possible to develop some scenarios and associated graphs that allow such a level of thinking to be fostered.

Finally, it is important to consider the graphs themselves and to ask whether or not they present data in the best possible format. The graphical reports are constrained to two formats—compared bar graphs, and compared boxplots—and are consistently presented (e.g., apart from width there is no change in the way the boxplots are presented between Figures 2 and 3). Although the boxplots are not standard, a key is provided and, as mentioned, the box and whisker values do not change between graphs. What is missing, which may be of interest to teachers, are the outlying values, although teachers also have access to individual student data.

In conclusion, the reports are consistently presented but, as they convey complex information, there is a need for statistical literacy and critical thinking in order to understand and make best use of them. This also requires attention to information about the local context. To what extent teachers currently have sufficient statistical literacy to make sense of these data is another question; moreover, it would be interesting to know what affects their motivation and capacity to make use of it.

References

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Carlton, VIC: Curriculum Corporation.

Baker, E. (2007). Presidential address: The end(s) of testing. *Educational Researcher*, *36* (6), 309-317.

Ben-Zvi D., & Garfield J. (Eds.). (2004). *The challenge of developing statistical literacy, reasoning and thinking.* Dordrecht, The Netherlands: Kluwer Academic Publishers.
Board of Studies. (1995). *Curriculum and standards framework: Mathematics.* Melbourne: Author.
Children, Schools and Family Committee. (2008). *House of Commons: Select Committee on Children, Schools and Families. Third Report.* Retrieved: 25-04-2009 from http://www.publications.parliament.uk/pa/cm200708/cmselect/cmchilsch/169/16902.htm
Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*, 382-393.
Ministerial Council on Education, Employment, Training and Youth Affairs. (2007). *Measurement framework for national key performance measures.* Retrieved 25-04-2009 from http://www.mceetya.edu.au/verve/_resources/2007_National_Measurement_Framework.pdf
Ministerial Council on Education, Employment, Training and Youth Affairs. (n.d.). *Benefits of participating in national assessments.* Retrieved 08-05-2008 from http://www.mceetya.edu.au/verve/_resources/Benefits_of_participation_in_national_assessments1.pdf
Tufte E.R. (2001). *The Visual Display of Quantitative Information (2nd edition)*, Cheshire, Conn : Graphics Press.
Victorian Curriculum and Assessment Authority. (2008). *Victorian essential learning standards. Discipline-based learning strand: Mathematics.* Melbourne: Author.
Victorian Curriculum and Assessment Authority. (n.d.). *National Assessment Program Literacy and Numeracy Testing* http://www.vcaa.vic.edu.au/prep10/naplan/index.html
Watson, J.M. (2006). *Statistical literacy at school.* Mahwah, NJ: Lawrence Erlbaum Associates.