INTERPRETING LITERACY AND NUMERACY TESTING REPORTS: WHAT DO TEACHERS NEED TO KNOW?

Robyn Pierce and Helen Chick
University of Melbourne, Australia
r.pierce@unimelb.edu.au

Australia's national testing of students' literacy and numeracy provides schools with a range of statistical reports analysing the school's results. The education authority intends that these data be used to inform planning to improve students' learning. Although lower secondary mathematics and English teachers see potential for using reports to identify weak students and curriculum topics needing attention, they consider the reports difficult to understand. This paper examines one such literacy test report together with typical "everyday" statistical graphics, and identifies the statistical literacy required to interpret each. The analysis suggests that the statistical literacy needed for interpreting the test report is greater than might be expected of teachers without any specific statistical education. Context-driven, targeted statistical learning will be necessary if testing reports are to inform teachers' professional practice as intended.

INTRODUCTION

National literacy and numeracy testing has become an established part of the education profile in many countries since the technological revolution has allowed the recording and analysis of the vast amounts of data this testing generates. The underlying assumption is that not only does such testing provide education authorities with a snapshot of the state of their system but that these tests provide evidence that school principals and teachers should use to inform their planning and teaching practice. Of interest to statistical educators are the issues of how this data is reported and whether teachers with no specialist statistical training are likely to be adequately equipped to interpret the data and reports to which they have access.

In the UK (Children, Schools and Families Committee, 2008), USA (Baker, 2007), and Australia (Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA), 2007) standardised testing is advocated to identify students' educational needs, promote data-driven decision-making for teaching, and increase schools' accountability to students and funding authorities. In Australia, for example, literacy and numeracy assessment results, plus socio-economic data, are recorded for all students in Years 3, 5, 7, and 9. These data are the basis for reports prepared for individual schools. The reports allow comparison of school results with national and state distributions, and include data allowing the consideration of possible impacts of the socio-economic background of the student cohort. In Australia reports are prepared for schools by the "data service" organization within each state's education authority.

Since the reports are generated by the data service organization, teachers do not need to conduct any sophisticated statistical analyses or tests. However, they *are* expected to interpret the reports supplied to them. Hence it is important to examine such reports to determine the statistical literacy that teachers need to make sense of this data. In particular, it is important to determine whether understanding such statistical reports requires more than "everyday life" statistical literacy.

The key features of statistical literacy have been well encapsulated by Gal (2002, pp. 2-3) as the ability to interpret and evaluate statistical information from diverse contexts, and discuss the meanings of, implications of, and concerns about such data and conclusions. This definition encompasses the expectation that teachers should be able to interpret national testing data (being data "encountered in diverse contexts"). The question remains, however, what "level" of statistical literacy do teachers actually require to do this? Is it more than might be expected from the general exposure to statistical thinking that individuals gain in their secondary school education? Do these reports, arising in the context of the workplace, require a level of statistical literacy more sophisticated than might be required for more everyday situations?

The issue about teachers' capacity to interpret and use statistical reports has been noted internationally, as illustrated by two examples. Matthews, Trimble, and Gay (2007), writing from their Georgia, United States experience, expressed concern that teachers need to know what the numbers in such reports indicate. An Organization for Economic Cooperation and Development (2004) report on the improvement of education in Chile discussed the introduction of national

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

testing. It found that constructive use of data seemed to be restricted by teachers' lack of capacity to interpret the reports they received. Locally, a pilot study with junior secondary mathematics and English teachers (Pierce & Chick, 2009) suggested that some teachers felt that Australian testing data were difficult to understand. The present study investigates whether statistical knowledge for the workplace really is more demanding than the statistical literacy requirements of everyday life.

IDENTIFYING AND CLASSIFYING STATISTICAL LITERACY

Chick and Pierce (2009) proposed a hierarchy for the abilities and understandings required to interpret data presented in tabular, graphic, and other condensed forms. This hierarchy is shown in Table 1 and will be used to analyse the statistical literacy needed for a typical everyday statistical interpretation task and that needed to interpret official reports from national tests.

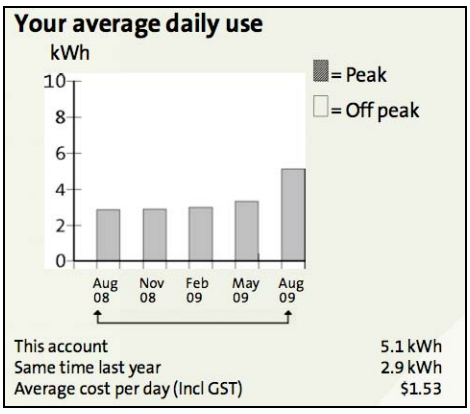
Table 1. Data interpretation hierarchy (Chick & Pierce, 2009)

Level	Characteristics	Examples
1. Attributes	Reads directly accessible single items explicitly evident in the data	Reads a single data point on the graph; can identify axis labels.
2. Comparisons	Attends to multiple aspects of the data to make direct comparisons	Compares data values; looks for trends; identifies skewed data from the shape of a boxplot.
3. Technical aspects	Applies relevant statistical tools to interpret the data as presented and can imagine how representations might change with changes in the data	Understands that differences between data sets may not be statistically significant.
4. Setting and context	Interprets the data in its broader or specific context, questions claims, and acknowledges alternative interpretations	Explains outcomes in light of local knowledge of the situation; questions the reason for unanticipated outcomes.

The national data reports typically include boxplots, which Pfannkuch (2006) carefully explains are not simple for the novice to interpret. Chick and Pierce (2009) conducted an analysis of such reports. The present paper looks first at a typical everyday graphical report, examining the statistical literacy required for understanding the data therein, and then conducts a similar examination of a non-boxplot national testing report (different from those in Chick & Pierce, 2009). The analysis was undertaken through careful consideration of the image, reading the explanatory notes, and discussion with an experienced teacher with a background in statistics.

Statistical literacy: Everyday life

In order to determine whether teachers need specialist statistical literacy skills as part of their professional learning we looked at the statistical demands of everyday life. Newspapers and incoming utility accounts were examined over a one-week period. The local major newspaper contained trend line graphs in the business section and an occasional bar graph in the news or sporting section. An electricity bill included a bar graph representing usage in previous billing periods. We selected two graphs, shown in Figures 1a and 1b, as representing the typical to most extreme complexity of the examples seen during the week. The skills needed to interpret these graphs are summarized in Table 2.



Figures 1a and 1b. Electricity bill graph and stock-market graph

Table 2. Data interpretation skills needed for electricity bill and stock-market graphs

Level	Skills for electricity bill graph	Skills for stock-market graph
1. Attributes	Read the title. Read the vertical scale and note the units. Read individual data points as indicated by ends of the bars. Read the legend. Understand horizontal axis is categorical not numerical.	Read the title. Read the vertical scale and note that it does not start from 0. Read individual data points on line graph. Read the legend. Understand the units on both axes (x – time, y – points). Read figures at top of the graph.
2. Comparisons	Compare the magnitude of numbers. Compare values (proportional/absolute comparisons)	Compare graph values against previous close (horizontal line). Make proportional/absolute comparisons (needs attention to 0 value).
3. Technical aspects	Understand what is meant by daily average. Understand about variation and its implications. Know which numbers to compare for interpretative purposes. Reconcile how big the difference is between values across billing periods.	Understand meaning of unlabelled figures. Understand about rising/falling trends, and variation and its implications. Know which numbers to compare for interpretative purposes. Reconcile how big the difference is between values over time.
4. Setting and context	Understand features of this graph in context (e.g., seasonal variations, holidays, extra house guests).	Understand features of this graph in context (e.g., other activity in market). Determine which trends might be a stimulus for action.

Statistical literacy: Interpreting literacy and numeracy reports

Typical of the graphical representations included in statistical reports made available to Victorian schools is Figure 2. This graphic presents a fictitious school's Year 3 reading results as a percentile rank within the State, for each year. Superimposed on this is a socio-economic indicator (Student Family Occupation or SFO). The graphic is meant to indicate the school's reading ranking against the relative level those students might be expected to achieve if socio-economic factors were the key determinant of their results. The skills needed to interpret this graphic are summarised in Table 3.

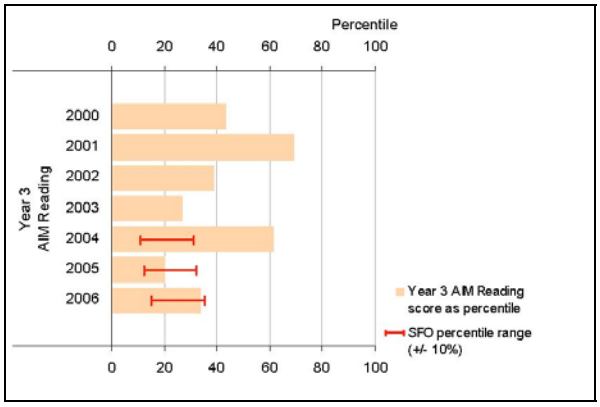


Figure 2. Sample literacy report including Student Family Occupation (SFO) (DEECD, 2009)

IMPLICATIONS AND CONCLUSIONS

One reason the reading report graph seems more challenging to interpret is because the data represented by it has been processed in a complex way. The electricity and stock-market graphs have straightforward variables: the measures displayed are directly associated with the situation. The electricity graph shows electricity consumption for each billing period over the previous year; the stock-market graph shows the value of a recognized stock-market index over time, and although this index is a composite value, it is generally understood as a simple absolute measure. The measures shown on the reading data graph, on the other hand, do not directly show the school's absolute reading performance on the testing. Instead, the school's position in relationship to other schools is displayed as a percentile ranking. The reader also has to understand that the SFO percentile range shows the school's percentile ranking by Student Family Occupation index in the middle of the line, with 10% whiskers on either side. This representation is based on the fact that SFO is, in general, a good predictor of performance, and so a school may be expected to have a performance percentile roughly matching their SFO percentile. In Figure 2, therefore, the 2004 reading performance for this school was better than expected for its SFO profile. Interpreting

this, however, is complex because of the fact that the data have been expressed as percentiles, and that there is a range of values presented to depict the school's approximate expected performance based on SFO.

Table 3. Data interpretation skills needed for the reading report graph

Level	Skills
1. Attributes	Read labels. Read percentile values. Relevant values are at the far right of each bar. Read SFO rank for school (middle of red bar). Read values at ends of SFO range. Read legend.
2. Comparisons	Understand the meaning of percentiles. Understand where this school ranks relative to all schools in the state. Compare results for different cohorts (year to year). Compare the value indicated by the end of the bar with the SFO estimate for that year.
3. Technical aspects	SFO red bar does NOT represent a confidence interval, nor error bar; the SFO bar extends +/- 10% of school's predicted percentile rank. Understand the graphic presents percentile rankings not scores—i.e., it does not tell you how well the students did on the test(s). Large change in percentile rank may or may not indicate large change in raw score. Percentile rank for school gives no indication of variability within the school cohort.
4. Setting and context	SFO estimate for the group is not related to teaching and learning. Need further information to determine when differences are educationally significant. Need to consider factors other than SFO that will impact on the student's learning and on their results relative to other schools.

As Tables 2 and 3 indicate, the skills and understanding required to interpret the test report graphic are considerably greater than those required for the everyday task of interpreting a utility bill or a stock-market graph. While various data services organizations provide some guidance on interpreting test reports there are many aspects to the statistical literacy required that are beyond that which non-specialists could normally be expected to possess. There is sufficient evidence to suggest that detailed analysis of the statistical literacy required for such reports should be undertaken and appropriate training should become a standard part of practicing teachers' professional development.

ACKNOWLEDGEMENT

The authors wish to thank Roger Wander, University of Melbourne, for finding examples and checking the data analysis.

REFERENCES

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

Chick, H. L., & Pierce, R. (2009). *The statistical literacy needed to interpret school assessment data*. Manuscript submitted for publication.

Children, Schools and Families Committee. (2008). *House of Commons: Select Committee on Children, Schools and Families. Third Report.* Retrieved 25 Apr 2009: http://www.publications.parliament.uk/pa/cm200708/cmselect/cmchilsch/169/16902.htm

Department of Education and Early Childhood Development (2009). *Percentile/SFO comparison charts*. Retrieved 9 Oct 2009: http://www.eduweb.vic.gov.au/edulibrary/public/account/operate/ptileSFO_v5.00.ppt.

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

Matthews, J., Trimble, S., & Gay, A. (2007). But what do you do with the data? *Principal Leadership, 7*(9), 31-33.

Ministerial Council on Education, Employment, Training and Youth Affairs. (2007). *Measurement framework for national key performance measures.* Retrieved 25 Apr 2009: http://www.mceetya.edu.au/verve/_resources/2007_National_Measurement_Framework.pdf.

Organization for Economic Cooperation and Development (2004). *Reviews of national policies for education: Chile.* Paris: Author.

Pfannkuch, M. (2006). Comparing box-plot distributions: A teacher's reasoning. *Statistics Education Research Journal, 5*(2), 27-45. Online: www.stat.auckland.ac.nz/serj.

Pierce, R., & Chick, H. (2009). *Teachers' intentions to use national literacy and numeracy assessment data: A pilot study*. Manuscript submitted for publication.