Math Ed Res J (2013) 25:189–205
DOI 10.1007/s13394-012-0046-3

ORIGINAL ARTICLE

Workplace statistical literacy for teachers: interpreting box plots

Robyn Pierce · Helen Chick

Received: 21 December 2011 / Revised: 2 May 2012 / Accepted: 2 August 2012 /
Published online: 30 August 2012

Abstract As a consequence of the increased use of data in workplace environments, there is a need to understand the demands that are placed on users to make sense of such data. In education, teachers are being increasingly expected to interpret and apply complex data about student and school performance, and, yet it is not clear that they always have the appropriate knowledge and experience to interpret the graphs, tables and other data that they receive. This study examined the statistical literacy demands placed on teachers, with a particular focus on box plot representations. Although box plots summarise the data in a way that makes visual comparisons possible across sets of data, this study showed that teachers do not always have the necessary fluency with the representation to describe correctly how the data are distributed in the representation. In particular, a significant number perceived the size of the regions of the box plot to be depicting frequencies rather than density, and there were misconceptions associated with outlying data that were not displayed on the plot. As well, teachers' perceptions of box plots were found to relate to three themes: attitudes, perceived value and misconceptions.

Keywords Statistical literacy · Box plots · NAPLAN · School reports · Teachers interpreting data

Introduction

The collection and dissemination of data is a major aspect of many contemporary workplaces; increasingly, these data are used as a driver for change and improvement.

R. Pierce (✉)
Melbourne Graduate School of Education, University of Melbourne, Level 3 234 Queensberry St, Melbourne, Victoria 3010, Australia
e-mail: r.pierce@unimelb.edu.au

H. Chick
Faculty of Education, University of Tasmania, Private Bag 66, Hobart, Tasmania 7001, Australia
e-mail: Helen.Chick@utas.edu.au

In this environment, with a prevalence of data use, it seems it is often assumed that those expected to interpret and act on data have the statistical skills to make sense of the reports, graphs, tables and other quantitative information that they receive. Analysing and interpreting such quantitative data, however, may not be a trivial task, requiring, as it does, both statistical skills and an awareness of the role played by the workplace context.

In the particular case of educational settings, such as schools, the principal and teachers are living within an increasingly data-driven policy environment. They are expected to use statistical information to inform school policy and to improve teaching practices and programmes and, as a result, must be able to think critically with data. This critical thinking requires *statistical literacy*, which is 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, coupled with a disposition to do so (Ben-Zvi and Garfield 2004; Wallman 1993; Watson 2006). To assist school personnel in their efforts, education systems provide data in what they believe are user-friendly formats. In the Australian state of Victoria, for example, schools are provided with graphical reports summarising outcomes from student literacy and numeracy achievement tests. Part of one of the Victorian NAPLAN Data Service reports, in this case Reading, is shown in Fig. 1. (NAPLAN is the National Assessment Program for Literacy and Numeracy.) It uses box plots to represent national, state and school data (presented left to right in the three boxes within each column) for the school as a whole and then subgroups: boys, girls, students with language background other than English (LBOTE), and Aboriginals and Torres Strait Islanders.

Statistics education research in other contexts (e.g. Makar and Confrey 2002), however, suggests that many people do not intuitively and correctly interpret and use such box plots or other data representations. In particular, Bakker et al. (2005) have highlighted the conceptual difficulties inherent in learning to interpret box plots (further details are presented later). In light of these issues, the research question addressed in this paper is:

> What aspects of statistical literacy affect teachers' capacity to interpret the box plots they encounter in student achievement reports?

After reviewing some of the background literature, a framework for professional statistical literacy, used to frame this research, is outlined. This is followed by specific details of the study reported in this paper that examine the research question, including the research methods, results and discussion. Finally, some implications and conclusions are outlined.

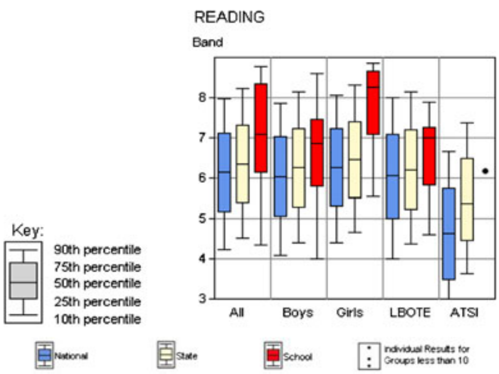


Fig. 1 Example of data provided to schools. (Graphs like these are generated online by the Victorian NAPLAN Data Service.)

Literature

As technology has advanced and statistical packages have been developed, the use of graphical displays to support exploratory data analysis has become commonplace. Among the new visual representations of distributions of data that have been invented is the box plot, first developed by Tukey (1977). Although there are variations in the way a box plot is presented, a classic box plot shows the minimum value, 25th percentile, median, 75th percentile and maximum value of a distribution. Over time, however, there has been discussion about what should constitute a "standard" box plot. Frigge et al. (1989) discussed a number of alternative box plot presentations that were facilitated by various software packages at that time and noted differences related to the length of the whiskers and the representation of outliers. Wickham and Stryjewski (2011) also describe and illustrate key variations that have been developed over the 40 years since Tukey promoted the simple box plot as a representation that would aid the comparison of distributions. These differences can still be observed in box plots from different sources, with some using the maximum and minimum of the data as the ends of the whiskers, others marking outliers beyond some defined cutoff used to determine the ends of the whiskers (e.g. 1.5 times the interquartile range), with yet others extending whiskers to a particular decile. (This is the case for the graphic reports provided to schools in Victoria, as in Fig. 1.)

Box plots are claimed to be a powerful method for displaying data because they emphasise the location of the centre of the data at the same time as illustrating the spread of the data, by showing both the interquartile range and overall range. They are designed to support the comparison of two or more distributions through comparisons not only of overall location and spread but also of quartiles. Bakker et al. (2005) posit several reasons for claiming the usefulness and simplicity of box plots as a way of representing data. They cite Mokros and Russell (1995) in support of the use of the median in the box plot, claiming that students understand the median as a measure of centre more readily than they do the mean. Bakker et al. go on to advocate the simplicity of the interquartile range as giving a measure of the degree of spread, as "an alternative to the computationally more challenging standard deviation" (p. 164). Nevertheless, Pfannkuch also notes that Bakker's own work suggests that students find the statistical ideas of medians and quartiles conceptually difficult (Bakker 2004, cited in Pfannkuch 2006), and observes that box plots, by their very nature, condense and obscure information, thus adding to their complexity.

Pfannkuch (2006) gives further insights into the complexity of box plot interpretation by looking in detail at a teacher's reasoning when interpreting box plots with her class of year 11 (senior high school) students. As a result of the study, Pfannkuch proposed a ten-element descriptive model of reasoning with box plots, highlighting the aspects that need to be attended to in order to make good sense of the data. Bakker et al. (2005), too, had earlier acknowledged four areas of conceptual difficulty for students learning box plots. First: "students are inclined to view data as individual cases, whereas box plots only provide aggregate information" (p 171). Second, box plots display densities rather than frequencies. This is the only commonly used graphical display of data to do this. Third, while it is easy to learn to identify the median, it is harder to recognise it as a measure of centre for the group. Finally, issues with defining quartiles and what to do with tied data values inhibit the use of the interquartile range as a measure of spread. Nevertheless, they suggested that the nature of the summary, with the cutoff points illustrated by the box and whiskers, encourages an intuitive comparison of key values. Like Pfannkuch, they particularly drew attention to the idea of attending to the “shift”, or displacement, between pairs of box plots, focusing on the ways in which the five-number summary values for one box plot are related to those of another. This encourages the notion of uniform (all values are displaced by about the same amount) and non-uniform shift.

To date, much of the focus amongst researchers has been on the issues of teaching and learning box plots. The focus of this paper is on teachers, not as teachers or learners but as users of data, and considers their ability and disposition to interpret box plots in the professional context of working with system reports on student achievement.

Framework

Pierce and Chick (2011a) proposed a framework to encapsulate professional statistical literacy; this was used as the basis for the study reported in this paper. The framework, summarised in Fig. 2, posits three levels for data interpretation with consideration of context. This framework was developed from earlier frameworks and discussion about interpreting data (e.g. Curcio 1987; Gal 2002; Shaughnessy 2007; Shaughnessy et al. 1996; Watson 2006), and the levels increase in complexity with higher levels depending on lower levels. The first level, *reading values*, requires a technical understanding of labels, scale, data type (e.g. numerical, categorical) and terminology such as percentage versus percentile. The next level, *comparing values*, requires awareness of relative and absolute differences, early informal inference and low-level statistical tools. Finally, *analysing the data set* requires consideration of the entire data set as a single object: observing and interpreting variation, trends and changes with time or other variables and attending to the significance of results. The surrounding box in Fig. 2 indicates the role of context for data interpretation. The *professional context* concerns knowledge of information recognised within the whole profession applicable to the data set (e.g. meaning of special terms such as “band” and “LBOTE”, which in this case come from the Australian educational context). *Local context* comprises contextual understanding that may be known by individuals about the specific data set but is not evident in the data set alone nor known by others from the wider profession (e.g. knowledge of the local school situation or timetabling issues affecting classes). This framework is used to examine the ways in which teachers interpret data supplied in box plot form.

Method

The participants

A mixed methods but primarily quantitative study was conducted. Ethical clearance for this study was given by the University of Melbourne (Ethics ID 1033125) and the

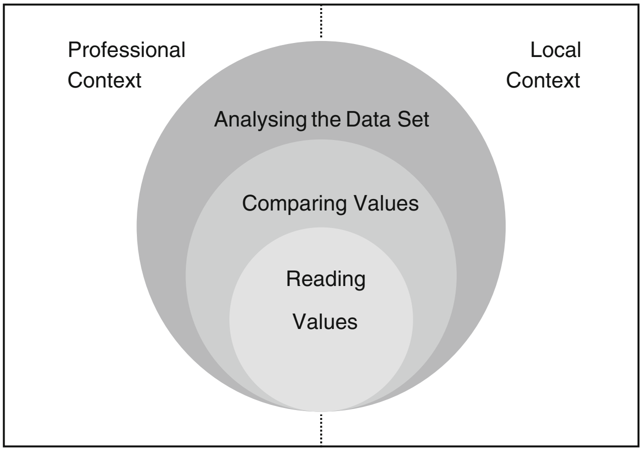


Fig. 2 A framework for considering professional statistical literacy

Department of Education and Early Childhood Development (2010_000564). Two phases of a larger study conducted with Victorian government schools contribute data to this research. The 150 participants for phase 1 came from a cluster sample of 20 schools—ten primary and ten secondary—selected randomly from within each of five organisational regions. In each school, seven teachers were randomly selected to participate, and the principal or data manager was also requested to take part. Phase 1, conducted in 2010, involved collecting data through a one hour paper-based survey with closed and open-response items about beliefs, attitudes and statistical understanding of school data (aspects of this are described in detail in the next section). The survey was administered face to face and followed by focus group discussion with each group of seven teachers from the participating schools, as well as a separate interview with the school principal or data manager. The teachers' responses to the phase 1 survey and focus group discussions then informed the development of an online survey designed to take approximately 30 minutes to complete, with predominantly closed responses. This survey, used in 2011 for phase 2 of the study, was made available to all teaching staff at 104 randomly selected primary and secondary schools across the Australian state of Victoria, with schools that participated in phase 1 deliberately excluded from the sampling frame for phase 2. The aim was to obtain data from approximately 1,000 teachers. Based on the research team's prior experiences, the sample of 104 schools allowed for a 60 % school response rate and then responses from 50 % of the teachers. In the end, 704 teachers supplied data for phase 2 of the study.

Instruments

The instruments were intended to examine (a) teachers' disposition towards data presented in box plots and to capture information about (b) teachers' skills in interpreting such data representation. Teachers' dispositions towards box plot representations were examined as part of phase 1 of the study. During the paper-based survey, participants were presented with Fig. 3 and were requested to respond to the following:

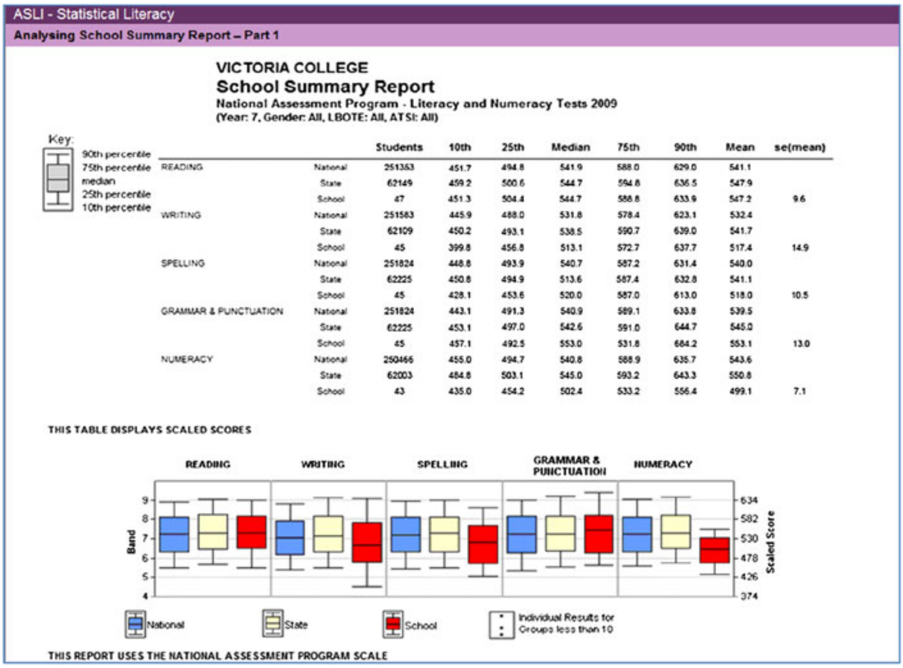
A. Please identify any aspects of this report which you think *might be of use to you as a teacher* by *circling it/them* and annotating the relevant aspects to indicate what is helpful and why.

B. Please identify any aspects of this report which you think *are hard to make sense of* by *drawing an arrow to it/them* and annotating the relevant aspects to indicate what may cause difficulties and why.

C. (Please write) any other comments about this particular report.

Additional attitudinal data regarding box plots were obtained from the semi-structured focus group discussions led by a researcher and held with each school group of seven teachers, together with the interviews conducted separately with the principal or data manager. The discussions and interviews were audio recorded and the transcripts analysed to identify themes among the comments made about data represented in the box plot format.

Results for teachers' skills in interpreting data presented in box plots come from the 2011 online survey (phase 2), in which there was a series of items focused on box plots. For the purpose of this study, the items were designed to examine participants' facility with aspects of the Professional Statistical Literacy framework given above. None of the items here assessed only reading values, but this skill was


ASLI - Statistical Literacy

Analysing School Summary Report – Part 1

VICTORIA COLLEGE

School Summary Report

National Assessment Program - Literacy and Numeracy Tests 2009

(Year: 7, Gender: All, LBOTE: All, ATSI: All)

		Students	10th	25th	Median	75th	90th	Mean	se(mean)
READING	National	251353	451.7	494.8	541.9	588.0	629.0	541.1	
	State	62149	459.2	500.6	544.7	594.8	636.5	547.9	
	School	47	451.3	504.4	544.7	588.8	633.9	547.2	9.6
WRITING	National	251583	445.9	488.0	531.8	578.4	623.1	532.4	
	State	62109	450.2	493.1	538.5	590.7	639.0	541.7	
	School	45	399.8	456.8	513.1	572.7	637.7	517.4	14.9
SPELLING	National	251824	448.8	493.9	540.7	587.2	631.4	540.0	
	State	62225	450.8	494.9	513.6	587.4	632.8	541.1	
	School	45	428.1	453.6	520.0	587.0	613.0	518.0	10.5
GRAMMAR & PUNCTUATION	National	251824	443.1	491.3	540.9	589.1	633.8	539.5	
	State	62225	453.1	497.0	542.6	591.0	644.7	545.0	
	School	45	457.1	492.5	553.0	531.8	684.2	553.1	13.0
NUMERACY	National	250466	455.0	494.7	540.8	588.9	635.7	543.6	
	State	62003	484.8	503.1	545.0	593.2	643.3	550.8	
	School	43	435.0	454.2	502.4	533.2	556.4	499.1	7.1

THIS TABLE DISPLAYS SCALED SCORES

THIS REPORT USES THE NATIONAL ASSESSMENT PROGRAM SCALE

Fig. 3 Report used as a focus for the 2010 paper-based and 2011 online survey questions. (Graphic based on hypothetical data provided by the Victorian NAPLAN data service for this research.)

covered by other items on the survey. The first set of items was associated with the sample school report shown in Fig. 3. This type of report is sent to schools to indicate their performance on national literacy and numeracy tests, although the school displayed in Fig. 3 is fictitious.

Participants were asked to examine the report and respond to the three statements below by choosing one of four options: “definitely true”, “definitely false”, “not enough information” or “I don't know”.

Question 1. The spread of the 50th–90th percentile results is wider in writing than in spelling.

Question 2. In the writing results, fewer students were between the 25th percentile and the median than were between the median and the 75th percentile.

Question 3. Victoria College's writing results have a greater range than the state results.

Question 4, a multiple-choice item, asked participants to identify the school's weakest area from the results displayed, with only one choice permitted from the five different literacy and numeracy components. This was followed by the open-ended question 5 that asked participants what features of the report had led to this conclusion.

The final item considered here was from later in the survey and concerned the histograms and box plots shown in Fig. 4. Participants were asked to identify which of the four box plots corresponded with the data displayed in the source B histogram and then to justify their choice (questions 6 and 7). (See endnote for a comment about question numbering.)

In terms of the Professional Statistical Literacy framework, questions 1 and 2 required respondents to compare, in the case of question 1 across box plots and for question 2 within a box plot. To answer these items correctly, the respondent needed to pay attention to the key and have an understanding of the words “percentile” and “median”, as well as have an implicit understanding about the number of values that would be included in both halves of the interquartile range. Question 1 also required underlying reading skills. A professional understanding of scores and bands was also helpful. Question 3 required participants to compare and analyse, and to pay attention to the key. In addition, knowledge of the professional context was necessary to understand that school data are a subset of state data. Question 4 involved analysing the whole data sets, with participants needing to read, compare and interpret the data within the professional and local contexts in order to identify the school's weakest area. The extent to which they could conduct such an analysis was revealed by their open-response justification for their choice of weakest area. Finally, question 6 also required participants to compare and analyse. A correct response required a good knowledge of box plots, including being able to identify and compare range and median values, and attend to the skewness of the graph, while also relating this information to a data set presented in histogram format. Evidence of the exact nature of the box plot features that participants regarded as significant for their decisions was revealed in question 7. Questions 6 and 7 did not, however, require any contextual interpretation.

The data for these seven questions were analysed quantitatively by frequency of the different multiple-choice options, with the exception of questions 5 and 7. In these latter cases, the responses were clustered according to the characteristics of the data

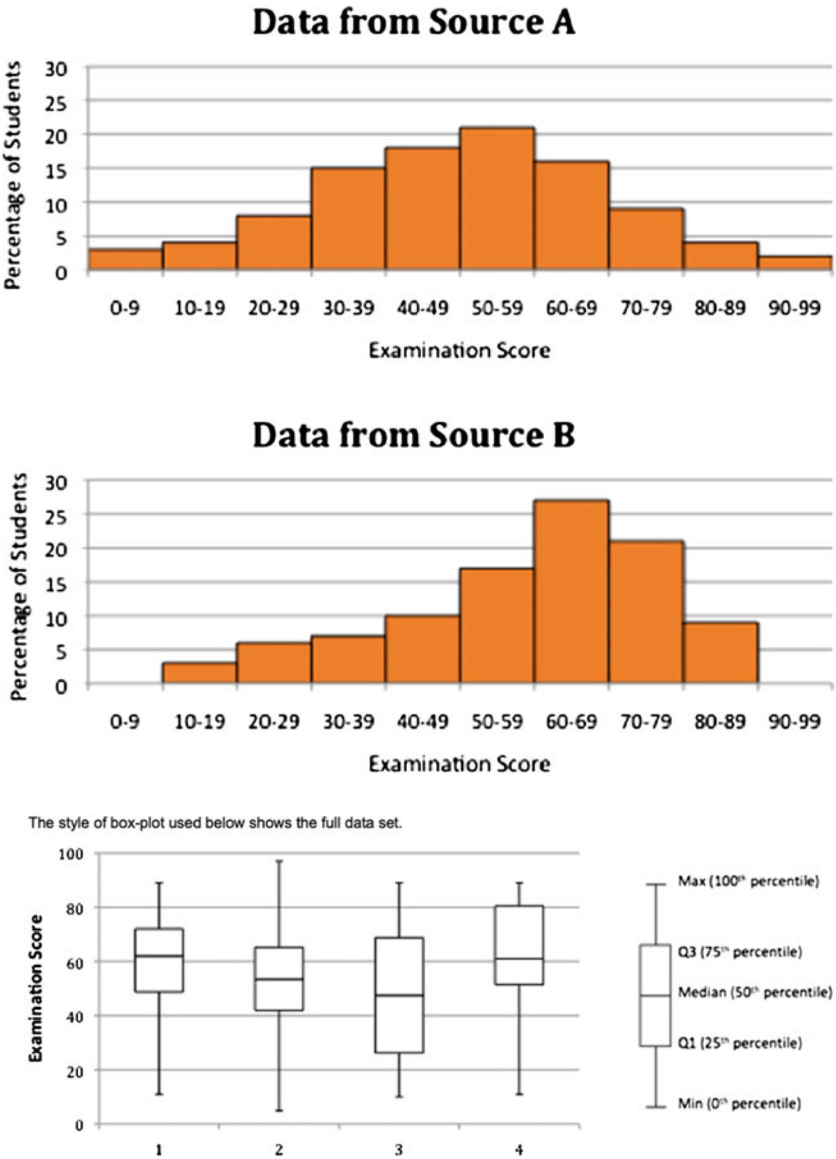


Fig. 4 Histogram and box plots showing school examination results

that were identified as contributing to the reasons and assigned hierarchical levels based on the strength of the evidence presented for the decision. More details about this analysis are presented in the "Results" section.

Results

Results of items probing phase 1 participants' perception of data are presented first. This will be followed by results from phase 2 of the study related to teachers' skills and understanding of box plots. Quotes from transcripts or open-response questions are provided to illustrate the data that were taken to be evidence of various themes that we will outline.

Teachers' disposition towards and perceptions of box plots

In response to questions A, B and C, which asked what the teachers found useful or difficult about reports showing box plots, three themes emerged from the responses given by the 96 teachers in phase 1 who provided such comments. The first theme concerned respondents' disposition towards box plots. Some expressed a positive attitude to the graphic representation of data (i.e. the box plots) in preference to the tabular presentation, but others admitted that they found these graphics difficult to interpret. The second theme related to the use made of the data presented in this graphic format. Some respondents commented that they valued box plots as particularly helpful for making comparisons and/or for their illustration of the spread or diversity of the data (in this case, students' ability). Thirdly, there were comments that indicated that some respondents held one of two common misconceptions about these particular box plots. In some cases, there was evidence that they had not read the key and did not realise that in these reports, the whiskers extend only to the 10th and 90th percentiles. The second misconception was the common frequency/density confusion, where, for example, the respondent thinks that a long box represents more students rather than a more diverse group of students. These misconceptions were targeted in items used in phase 2 of the study. The comments in Tables 1, 2 and 3 were typical of those taken as evidence of the three themes. These results have not been quantified, as the questions were open response and did not target particular themes, and nor did all teachers in Phase 1 choose to respond to all the parts. This means that percentage values would have the potential to be misleading.

Skills and understanding

Of the 704 teachers who submitted responses to the online survey (phase 2), 69 % were female and 31 % male, with 45 % working at the primary school level and 55 % at the secondary level. Asked to indicate "The last time I formally studied Statistics as

Table 1 Sample statements indicative of teachers' disposition towards box plots

Theme 1. Disposition
Positive attitude to the graphic representation of data [box plots]
Visually much easier and quicker to interpret [than a table].
Graphs are good and clear as an intro[duction], then look at figures [table] if in need of clarification.
A generally negative attitude: find box plots difficult to interpret
I find this report a bit difficult to interpret as I struggle to decipher box plots.
I am confused about the percentiles and what they mean.
I really hate box and whisker graphs—they make little sense to me and I feel frustrated even looking at them!

Table 2 Sample statements indicative of teachers' views of the use for which box plots are valued

Theme 2. Use for which box plots were valued
Box plots helpful for making comparisons
Good to see school compared to State/National.
Great to track progress over number of years. Good to see comparison between bigger picture and school…
Get an idea where our students sit in relation to like schools and schools around the state.
Good to compare; also good to show development or progression.
Box plots valued for illustration of the spread or diversity of the data
[Circled lower whisker on Writing plot] There is a large spread in this cohort and teachers will need to meet the needs of all.
[Circled whiskers of school Writing plot] Wide spread of results—varying abilities.
[Circled school Spelling plot] Spelling lower than state and national—need to work on in school.
It is useful to see the range of ability (P10–P90) and how the middle 50 % (P25–P75) is spread in terms of levels of achievement.

a subject or a topic in some other subject", 20 % chose "never", 10 % chose "year 10 or earlier", 21 % chose "year 11/12" and 47 % chose "post secondary". Asked about attendance at any of a range of professional learning programmes related to interpreting reports on student achievement data, 75 % indicated that they had attended at least one such programme.

The quantitative results for questions 1, 2, 3, 4 and 6 are presented in Tables 4 and 5, noting that about 10 % of the 704 participants did not respond to each of the items, and hence, the percentages are given as a proportion of the non-blank responses. Most respondents gave correct responses to questions 1 and 4, in contrast to questions 2, 3 and 6 where the majority of responses were either incorrect or "I don't know". Finally, in response to questions 5 and 7, only a minority of the respondents was able to articulate correct reasons clearly for their choices in questions 4 and 6. These results are discussed in more detail in the following subsections, and, where, relevant the results are cross-tabulated against the participants' demographic factors, with chi-squared tests of independence run (using SPSS) to check the statistical significance of any apparent associations.

Table 3 Sample statements indicative of teachers' misconceptions affecting their interpretation of box plots

Theme 3. Misconceptions impacting on interpretation of boxplots
Not reading the key and realising whiskers extend only to 10th and 90th percentiles
Highest 10 %, lowest 10 % —tells us they may need support eg assessment ! … and teaching strategies for these extreme students. Colour, position [of the box]—tells us where the group stands as a whole—whole class approaches.
Thinking in terms of frequency not density.
[circled low tail school Writing] The long tail means there is a big group of students who need extra support.
[circled low tail school Writing] Too many at this level below state/nat[ional].

Table 4 Percentage of respondents choosing different options for questions 1, 2 and 3

	Definitely true	Definitely false	Not enough information	I don't know
Question 1. The spread of the 50th–90th percentile results is wider in writing than in spelling	70 %[a]	12 %	4 %	14 %
Question 2. In the writing results, fewer students were between the 25th percentile and the median than were between the median and the 75th percentile	56 %	17 %[a]	6 %	21 %
Question 3. Victoria College's writing results have a greater range than the state results	71 %	9 %[a]	5 %	15 %

Percentages are the proportion of all actual responses, rounded to the nearest whole number

[a] Correct responses

Question 1: The spread of the 50th–90th percentile results is wider for writing than for spelling.

This was the simplest item and was answered correctly by 70 % of respondents. There was some statistically significant association between teachers' choice of response and their previous experience of statistics either through formal study or professional development programmes. Those who had never studied statistics or who had only met it at year 10 or lower were more likely (than would be expected if there were no association between the variables) to choose “I don't know” ($p<0.001$). This was also true for those who had not attended a relevant professional learning programme ($p=0.003$). There was no statistically significant association between response patterns and the type of school in which participants were teaching (primary or secondary).

Table 5 Percentage of respondents choosing the different options for questions 4 and 6

Question 4. From the results displayed [Fig. 3] what would you identify as the school's weakest area? (Choose one only)	
Numeracy[a]	86 %
Writing	9 %
Spelling	3 %
Reading	1 %
Grammar and punctuation	1 %
Question 6. [Looking at the histogram and box plot in Fig. 4] Which box plot represents the data from source B?	
Box plot 1[a]	26 %
Box plot 2	5 %
Box plot 3	10 %
Box plot 4	37 %
I don't know	22 %

Percentages are the proportion of all actual responses, rounded to the nearest whole number

[a] Correct responses

Question 2: In the writing results, fewer students were between the 25th percentile and the median than were between the median and the 75th percentile.

This item was intended to determine the extent to which participants appreciated the way in which the box plot depicts density rather than frequency. The longer length of the box showing the third quartile was misinterpreted as representing more students by 55 % of respondents (those picking “definitely true”). Only 17 % of the respondents chose the correct “definitely false” response to this item. The response pattern showed a statistically significant association with school type ($p=0.001$), last formal study of statistics ($p<0.001$) and attending a professional learning programme ($p=0.038$). More secondary (22 %) than primary (11 %) teachers chose the correct response, while fewer primary and more secondary teachers than expected chose the “I don't know” option. Of those teachers who had studied some statistics at a post-secondary level, more than expected were correct. By a smaller margin, more chose the incorrect response rather than indicating uncertainty. The pattern was the same for those who had attended a professional learning programme.

Question 3: Victoria College's writing results have a greater range than the state results.

This question addressed the fact that the particular style of box plots used for teachers in Victoria requires deeper interpretation beyond literally reading the displayed plot. The viewer must understand that a school's data form a subset of the state data and that the outlying data values are invisible in these particular box plots since, as the key indicates, the lowest value indicated by each box plot is the 10th percentile and the highest value is the 90th percentile. Only 9 % of teachers chose the correct response to this item (“definitely false”), whereas 71 % were, presumably, misled by the shapes of the relevant box plots and chose definitely true. There was no statistically significant association with school type, but again, those who had studied statistics in a post-secondary programme and those who had attended a professional development programme were more likely to choose correctly. Those with little or no background in statistics were more likely to choose “I don't know”.

Question 4: From the results displayed, what would you identify as the school's weakest area?

Question 5: What features of the report have led to this conclusion?

The best choice, numeracy, was chosen by 86 % of respondents, with 9 % choosing writing. Only 15 % provided a response in question 5 that included a definitive comparison with the state results, including noting that almost 90 % of the school results were in the lower half of the state results. The rationale given for their choices was sometimes unhelpful (e.g. “prior knowledge”) or vague, for example, statements like “table scores”, “the box and whisker graph” or “the school mean”. The more informative explanations varied from those that gave one clear reason to others that considered most or all critical features of the distributions. The following two statements were typical of those giving one good reason.

> [Has] the largest gap between school result and state result.

> The top level of numeracy for the school falls lower than the average level of the State and Nat[ional] levels.

The next two statements are representative of the most thorough responses.

> 90th percentile [for School Numeracy] just slightly higher than [50th percentile for] State and National [results]. [Comparison of] median[s]. [The School's] narrow spread of achievement [in Numeracy]. [School's] 10th percentile [for Numeracy is] lower than State and National.

> It [Numeracy] has the lowest average, median, top score, 75th and 25th percentile[s]. The low score [10th percentile] is higher than, Writing, but variables described before are higher [for Writing].

Question 6: [Looking at the histogram and box plot in Fig. 4] Which box plot represents the data from source B? Question 7: How do you know [which box plot matches the data from source B]?

In order to match correctly the negatively skewed data presented in the histogram with the correct box plot, participants had to attend not only to the range and the median, but also the skewness of the graph or the location and spread of the third quartile. As is seen from the comments in response to question 7, many of the participants (37 %) were distracted by the large box depicting the third quartile in box plot 4, associating this with the large “clump” of values in the upper range of the histogram. In contrast, only 26 % correctly chose box plot 1. A higher proportion of secondary teachers than primary teachers chose box plot 1. Teachers who had studied statistics at year 11/12 or the post-secondary level were more likely to choose either box plot 1 or 4, which could be identified using the range and the median. Surprisingly, of those who had studied some statistics at the post-secondary level, 42 % incorrectly chose box plot 4 and only 29 % chose box plot 1. Making the correct choice required paying attention to skewness as well as centre and spread.

The respondents clearly found question 7 a difficult question: the weakest responses made no, little or incorrect reference to the graphs; examples of such responses included “too difficult to read” or “percentile level”. Another group noted some distinguishing feature(s) but not enough to restrict the choice to the correct answer, for example: “The median [is] higher showing a positive skew” or “median is between 60–69”. Better responses noted features that would restrict the choice to box plots 1 or 4, for example:

> Because the highest and lowest score aligns with the first graph as does the data within the range 25–75th percentile. The median score can also be aligned.

> The examination scores range between 10 and 90. The median is sitting around 60.

The most complete responses gave enough detail to isolate one correct pairing, for example:

> Whiskers start and end around the correct numbers. Most of the results in the skewed graph are above 60, so are compressed in the M & Q1 area

> above 60 making it smaller as most of the results are in this area of exam scores.
>
> The box plot shows that the top 50% of students have less spread than the bottom 50%, with the median and extremes correctly positioned.

Not all of those respondents who attempted to identify the correct box plot (question 6) offered reasons for their choice (question 7). Of those who did answer question 7, 32 % had correctly identified the match with box plot 1. Within this subgroup, however, only 26 % gave reasons that could ambiguously identify box plot 1 or 4, and only 23 % gave complete responses that pinpointed only box plot 1. Within the subgroup who chose box plot 4, 24 % gave reasons that would apply to box plots 1 or 4, and 75 % gave weaker answers.

Discussion

In the following discussion, we first look at teachers' dispositions, acknowledging the role they play in statistical literacy. We then make some observations about what the data reveal about teachers' capacities using the professional statistical literacy framework.

Disposition

From the data, we can see that teachers' disposition to work with box plots is mixed. A few teachers find the graphic reports with box plots difficult to understand. They do not like them and express frustration at being expected to deal with them. In contrast, however, by far, the majority of teachers' comments from those expressing an opinion suggest that they find box plots to be a very helpful summary of the data that allows them to compare across test results within their school and to compare their school results with those of the state. Other comments focus positively on being able to see the location and spread in one set of test results and consider the implications for teaching. If all teachers are expected to deal with this kind of data, then work must be done to ameliorate the negative dispositions evident here and also from comments made within the focus groups.

Read, compare and analyse

In general, most of the teachers could correctly identify which of the data sets showed the school's poorest area of performance, revealing the capacity to make comparisons across box plots. They could also successfully read those aspects of the box plots that are directly represented by the graphic (e.g. in question 1, "spread" is indicated by the "length" of the box). Nevertheless, their attempts (question 5) to justify their choices made in question 4 revealed some difficulties with the appropriate technical language and lack of attention to all the relevant features of the distribution.

There was strong evidence of misconceptions about the box plot representation. It seems both from the number of open comments and from the skills results, particularly in questions 2 and 3, that there were teachers who were confident that they could interpret box plots but who held incorrect or incomplete conceptions regarding the box plot representation of data. In particular, we see evidence of the perception of frequency rather than density. This is the misconception that Bakker et al. (2005) note causes confusion for students and which requires careful teaching. This misconception is clearly present for many teachers, perhaps as a consequence of limited work with this kind of representation or perhaps because of a lack of experience in linking different representations of the data, for example dot plots and box plots.

Professional and local context

There is also evidence of a lack of attention to the details of the box plots provided in their professional context, which are truncated at the 10th and 90th percentiles. Friel et al. (2001) draw attention to the need to teach students to read the axes, scale labels, keys, etc.; this is a basic but important skill for professionals too. Frigge et al. (1989) discussed some issues and confusion that may be caused by a lack of standardisation for box plots. Tufte (2001), in his discussion of good graphics that communicate well, emphasises the importance of the use of standard graphics. The choice not to represent the full distribution in these reports causes confusion for some of the teachers, although those with a knowledge of graphics produced for their profession will be expecting this non-standard graphic.

Issues associated with local context were not especially evident from the research, since the data used were from a fictitious school, but there was evidence of the teachers being able to identify actions that could be taken to address students' results.

The skills results suggest that in general, those who have studied more statistics and/or attended professional learning programmes are more likely to be correct in their reading and analysis of the data. The results also hint, however, that there is an unexpectedly large proportion in this group who have erroneous conceptions, For questions 2, 3 and 5, we only see correct choices from a minority of these teachers who might be expected to have the strongest statistical backgrounds. Since, elsewhere in the online questionnaire, only 15 % of respondents reported that they did not feel able to adequately interpret school data, it seems that many teachers are oblivious to the problem.

Implications and conclusions

The results provide evidence that most of the teachers can read box plots at a superficial level, sufficient for making straightforward comparisons to determine which, of two groups, is “better”. What is not clear (and the study did not examine this specifically) is whether teachers are able to assess when and which differences might actually be significant. A more obvious difficulty, revealed by the study, is that many teachers seem to assume—either through a genuine misconception or just a lack of care in discussing the results—that the box plot depicts the complete data set, instead of omitting the extreme top and bottom 10 % of the group. Although there are administrative reasons for why the educational data service presents data in this way, the consequences for teachers' understanding of the data need to be acknowledged.

More subtle, but as important, is the issue of density versus frequency, with teachers being misled by the size of the components of the graph to interpret these as proportional to the numbers of students, rather than understanding that each component contains a set proportion of the data and that it gives an indication of the density of results. This issue needs to be addressed through explicit professional development and, perhaps, the provision of a reminder graphic to be discussed with staff at the time that reports come into schools.

Although the 70 % success rate on question 1 and the even higher success rate on question 4 are pleasing, a word of caution is in order. When we consider test results in a general education context, we would commonly report “most students are correct” as a positive result. However, in this context, we are concerned about professional statistical literacy. An incorrect interpretation of the data may waste considerable time and money and may fail to lead to improved learning outcomes for students. Even on items where most respondents were correct, there were still up to 30 % who were either incorrect or admitted that they did not know. It would seem important therefore to consider what roles in the school require a sound level of statistical literacy. Either the issue of statistical literacy should be addressed for all teachers, through targeted pre-service or in-service training, or, alternatively, individuals who operate in specified roles should receive additional education in statistical literacy before they are given the responsibility of interpreting school data reports.

Finally, although the study has revealed some insights into teachers' dispositions towards box plots as a particular statistical representation, there are interesting questions still remaining concerning teachers' attitudes towards such school assessment data in general (see Pierce and Chick 2011b for a report of some pilot work on this topic) and the more complex question of the extent to which dispositions are related to statistical literacy. If assessment data are to be put to effective use (where it is acknowledged that the role of “numbers” in the educational sphere might be contested; see, e.g. Lingard 2011), then clearly both capacity to interpret and appropriate dispositions with respect to statistical representations will need to be enhanced.

Acknowledgments The authors wish to thank all participating schools and teachers and to acknowledge the contribution of the other members of the research team: Roger Wander, Ian Gordon and Sue Helme (University of Melbourne); Jane Watson (University of Tasmania); Michael Dalton and Magdalena Les (VCAA) and Sue Buckley (DEECD).This research was funded by the Australian Research Council (LP100100388), the Victorian Department of Education and Early Childhood Development (DEECD) and the Victorian Curriculum and Assessment Authority (VCAA).

References

Bakker, R., Biehler, R., & Konold, C. (2005). Should young students learn about box plots? In G. Burrill & M. Camden (Eds.), *Curricular development in statistics education: International Association for Statistical Education 2004 Roundtable*. Voorburg: International Statistical Institute. http://www.stat.auckland.ac.nz/~iase/publications.php?show=rt04. Accessed 10 February 2011

Ben-Zvi, D., & Garfield, J. (Eds.). (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht: Kluwer Academic Publishers.

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*, 382–393.

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124–159.

Frigge, M., Hoaglin, D., & Igliwicz, B. (1989). Some implementations of the box plot. *The American Statistician, 43*(1), 50–54.

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

Lingard, B. (2011). Policy as numbers: ac/counting for educational research. *The Australian Educational Researcher, 38*, 355–382.

Makar, K., & Confrey, J. (2002). Comparing two distributions: investigating secondary teachers' statistical thinking. In B. Philips (Ed.), *Developing a statistically literate society.* (CD-ROM Proceedings of the 6th International Conference on Teaching Statistics). Cape Town, South Africa: International Statistics Institute.

Pfannkuch, M. (2006). Comparing box plot distributions: a teacher's reasoning. *Statistics Education Research Journal, 5*(2), 27–45.

Pierce, R., & Chick, H. L. (2011a). Reacting to quantitative data: teachers' perceptions of student achievement reports. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thornton (Eds.), *Mathematics: Traditions and (new) practices. Proceedings of the 23rd biennial conference of The Australian Association of Mathematics Teachers Inc. and the 34th annual conference of the Mathematics Education Research Group of Australasia Inc.* (pp. 631-639). Adelaide: MERGA/AAMT.

Pierce, R., & Chick, H. L. (2011b). Teachers' intentions to use national literacy and numeracy assessment data: a pilot study. *The Australian Educational Researcher, 38*, 433–447.

Shaughnessy, J. M. (2007). Research on statistical learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). Charlotte: Information Age Publishing.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics in education* (pp. 205–237). Dordrecht: Kluwer.

Tufte, E. R. (2001). *The visual display of quantitative information*. Cheshire: Graphic Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading: Addison-Wesley.

Wallman, K. K. (1993). Enhancing statistical literacy: enriching our society. *Journal of the American Statistical Association, 88*(421), 1–8.

Watson, J. M. (2006). *Statistical literacy at school*. Mahwah: Lawrence Erlbaum Associates.

Wickham, H., & Stryjewski, L. (2011). 40 years of boxplots. http://vita.had.co.nz/papers/boxplots.html. Accessed 20 April 2012

Endnote

In order for data from this report to be cross-referenced against other reports from the same study, it should be noted that questions labelled 1, 2, 3, 4, 5, 6 and 7 in this paper correspond to items 15(ii), 15(iii), 15(iv), 16, 17, 24 and 25, respectively, in the original online survey.