GRAPH CREATION AND INTERPRETATION: PUTTING SKILLS AND CONTEXT TOGETHER

NOLEINE FITZALLEN
University of Tasmania
Noleine.Fitzallen@utas.edu.au

JANE WATSON
University of Tasmania
Jane.Watson@utas.edu.au

Although the creation of graphs to display data has been part of the school curriculum for some time, the call in the new Australian Curriculum for "numeracy across the curriculum" provides both the opportunity and challenge to link the skills of graph creating with the understanding of context in order to produce meaningful interpretation of the messages held in graphs. This paper reports on classroom experiences of and follow-up interviews with 12 grade 5/6 students who were introduced to the software *TinkerPlots* to assist in graph creation. The focus is on their success at graph creation and interpretation in contexts that provide potential links to other subjects in the school curriculum. Implications for the curriculum and teaching are drawn from the students' experiences.

Introduction

As statistics continues to occupy a place in the Australian Curriculum—Mathematics (ACARA, 2010), the justification for this place is often built upon its application "in context" (Mooney, 2010). In fact, many years ago Rao (1975) claimed, "Statistics ceases to have a meaning if it is not related to any practical problem" (p. 152). These statements suggest that the cross-curriculum opportunities of statistics should make it a major component of numeracy across the curriculum, as set out for the Australian Curriculum by the National Curriculum Board (NCB, 2009). The question for those developing the pedagogy for statistics is, "How do teachers link the development of skills with the provision of motivating context?" In turn, the question for the assessment of student learning is "What mix of skills and understanding of context is expected of students when they are drawing informal statistical inferences?" These questions may be further complicated by the stipulation by the NCB (2009) that technology be employed across the curriculum as appropriate. The Australian Curriculum–Mathematics (ACARA, 2010) carefully avoids mention of specific technology, and in particular of statistical software packages. Acknowledging the importance of software to enhance the teaching and learning of statistics, this study employed the software for middle school, *TinkerPlots* (Konold & Miller, 2005), to explore the issues raised in these questions. Twelve students' experiences with *TinkerPlots* and data sets are explored in two settings: four teacher-led classroom sessions and a one-to-one assessment interview with researchers.

Underlying model

The model used for the analysis of the data in the study was developed by Watson and Fitzallen (2010) in a research report commissioned by the NSW Department of Education and Training. The model posits three cycles of development of graph understanding for the purpose of drawing informal inferences while employing graphical representations of data sets. The model is based broadly on the work of Biggs and Collis (1982, 1991) in cognitive psychology. Although acknowledging other factors that influence learning, this model is based on the combining of basic cognitive content elements to construct the understanding required to be successful in each of the three cycles. The elements involved for each cycle are derived from the research in the field and are combined in various ways to produce the understanding necessary to create meaningful graphs.

The first cycle consists of building up the Concept of Graph from the elements of Attribute, Data, Variation, and Scale. Because of the several types of complexity associated with the data that are represented in graphs, there are two parallel second cycles in the overall model. One cycle considers multiple attributes and the other, large data sets. For the purpose of this study, only the cycle related to multiple attributes is employed in the data analysis. The Concept of Graph for Multiple Attributes is based on the following elements: Concept of Graph (from the first cycle), Types of Attributes (e.g., categorical and numerical), Two-dimensional Scaling, and the Relationship of Two Attributes to a Single Case. The aim of this study is to explore the creation of graphs representing two or three attributes. The third cycle, Graph Interpretation is based on the elements: Concept of Graph (from cycle 1 or cycle 2), Concept of Variation, Concept of Average, Context, and Critical Questioning Attitude.

In this study variation and average were not explicitly taught in the classroom but the introduction of hat plots, where the crown covers the middle 50% of the data and the brims extend to the extremities, was intended to develop intuitions about these concepts. Context played an important role in motivating the students in both the classroom and interview settings and students were encouraged to make hypotheses and be critical of what was represented in the graphs.

Research questions

With reference to the questions asked at the beginning of this paper and the underlying model, the data for the 12 students were analysed to answer the following research questions.

1. Within a grade 5/6 classroom teaching setting, what skills for Graph Creation for Multiple Attributes leading to the use of contextual understanding for Graph Interpretation are developed?
2. Within a one-to-one computer-based interview setting employing software, what skills for Graph Creation for Multiple Attributes leading to the use of contextual understanding for Graph Interpretation are transferred and applied?

Methodology

The subjects in the study were 12 students from a grade 5/6 class at a rural primary school (K–6) in Tasmania, Australia. The 12 were members of a class of 26 who participated in four lessons with the first author using *TinkerPlots* to analyse the class' resting and active heart rates. Throughout the experience, the students were introduced to bins, stacked dot plots, reference lines, hats, scaling, and scatterplots. Students were asked to save the graphs they produced and the text boxes in which they wrote their interpretations and conclusions from their graphs.

A month later the students were videotaped as they were interviewed by the second author or another researcher in a study to explore the effectiveness of paper-and-pencil versus on-line assessment of students' statistical understanding (Watson & Donne, 2009); these students were among the 24 on-line subjects using *TinkerPlots* as the software in that study. Transcripts of the videotapes of the students' interviews and *TinkerPlots* files were used as the data for this study. Earlier analyses of related data from these students are found in Fitzallen and Watson (2010).

Student data

In the classroom, students had access to their own heart rate data both before and after a jump-rope activity. They could hence identify the data cards and icons with their fellow students or themselves. The other attribute in the classroom was gender. For the interview the data set consisted of data for 16 students, named (but unknown), along with their age, eye colour, weight, favourite activity, and number of fast food meals eaten per week (data found in Chick & Watson, 2001).

Analysis

All students in this study satisfied the criteria for the first cycle of the model of Watson and Fitzallen (2010), understanding the Concept of Graph, a basic element of the second cycle, creating a Graph for Multiple Attributes. The other elements of the second cycle all contribute to the forms of graphs that were created by the students. The types of attributes, categorical or numerical, determined the types of plots that could be created in *TinkerPlots*, either with "bins" or "scaled" axes. Bins are used for categorical data (such as gender) or for numerical data that are grouped together (such as age ranges of 7–13 years and 14–20 years). A scaled axis can only be used with numerical data.

In considering two attributes, it was hence possible depending on the type, to create six different plots in *TinkerPlots*. These are shown in Figure 1 on the following page. Students' skills in graph creation and interpretation are displayed as they create these different plots with the software and discuss what they mean.

For the third cycle of Graph Interpretation the most significant element beyond using the second cycle of creating Graphs for Multiple Attributes for this study is the element of Context. In relation to the tasks provided to the students during the classroom sessions and interviews, there were three aspects of student introduction of context. In the first case either the individual icons on the plot, or the data cards themselves, were used to provide context for the individual people represented in the data set. In the second case, the graph was used to provide a contextual description about trends observable in the graph (not individuals). In the third case, students went beyond the information in the graph to create stories about the people in the data set, to speculate about why relationships were seen, or to give advice about how to mitigate a trend seen in the graph.

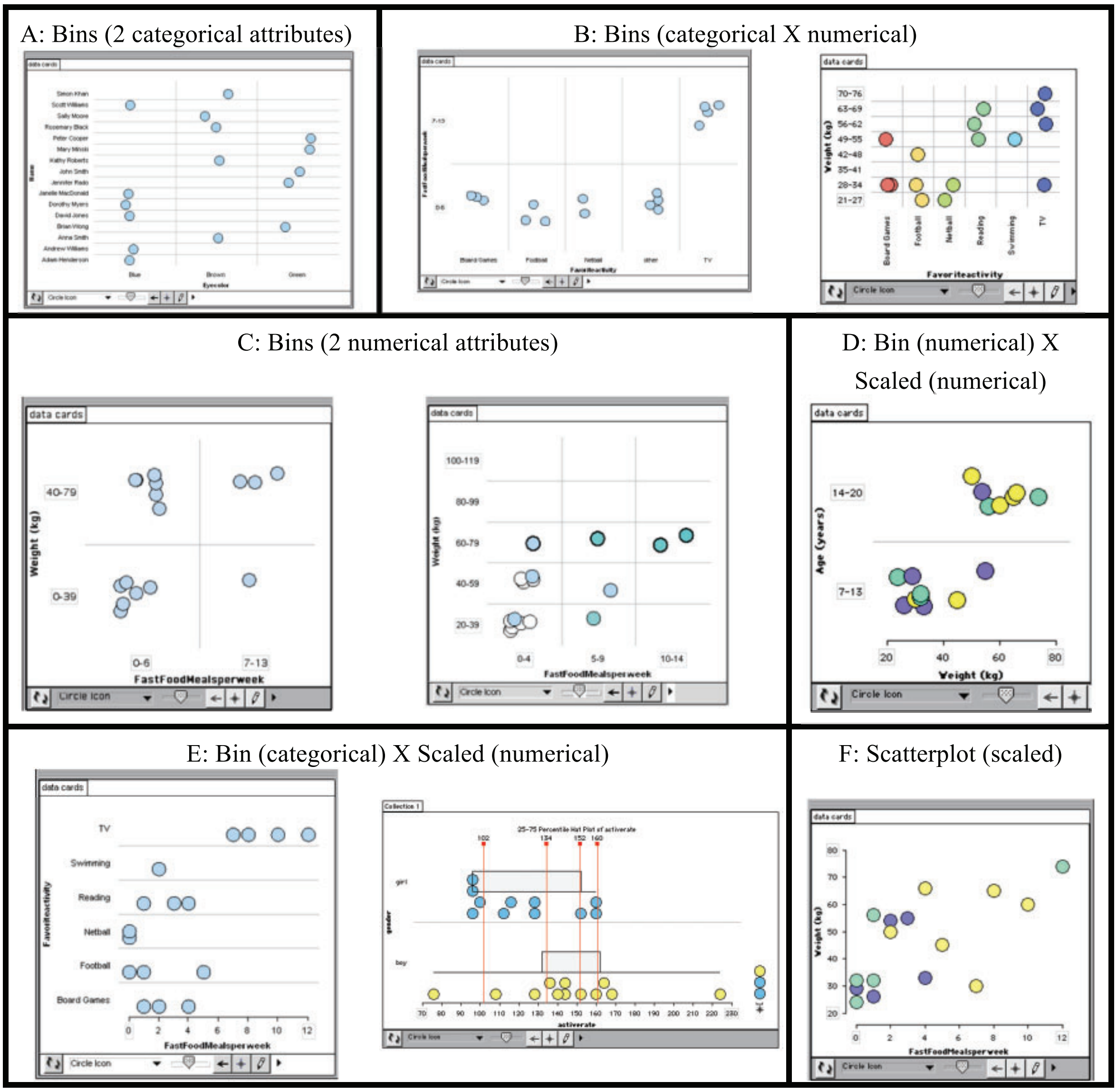

Figure 1. Six different types of graphs for two attributes in TinkerPlots.

The two groupings for skills of Graph Creation and use of context for Graph Interpretation were used to categorise the responses of students in the two situations: first the work created by the students as part of their four classroom lessons and second as part of their follow-up interviews.

Results

In the classroom (Research Question 1)

The students created four of the types shown in Figure 1 (B, C, E, and F) as well as using bins or stacked dot plots for single attributes. It was not possible to create bins for two categorical attributes (A) because only gender was available and because of their concurrent experience with scaled stacked dot plots, it would not be expected that students would combine bins with scaled axes for the numerical attributes (D). The text boxes saved in the files provided the material to assess the extent to which context was used in the interpretation of the graphs created.

Although there is no doubt that to some extent the files represented interaction with the teacher and perhaps other students, they present a picture of the students' beginning appreciation and understanding of the data handling task and its purpose relating to heart rates under different conditions. Even graphs of the same type had different appearances in different students' files because of the flexibility of plot creation in *TinkerPlots*.

Except for one student who forgot to save his file, the other 11 students created between 3 and 6 graphs, for a total of 48. Fifteen graphs were based on a single attribute, with two displaying categorical attributes (name or gender), three displaying the numerical attributes in bins (hence showing ranges), and 10 stacked dot plots of one or other of the heart rate data sets (with scaled axis). Thirty-three of the graphs produced showed the relationship between two attributes. Four graphs used bins for gender and one of the heart rate sets (B), 2 used bins for numerical data on both axes (C), 16 used bins for gender and scaled the heart rate data (E), and 11 were variations of scatterplots (F).

The students' interpretations of the graphs recorded in the text boxes in *TinkerPlots* were placed in three categories. From their graphs, three students reported individual values for interesting participants; e.g., the highest or lowest rate, or the rates for their teachers. Ten of the students wrote more generally about what the graph portrayed about the data set, e.g., there is a larger range for the boys, and 50% (or most) of the people are between 80 and 110. These comments were based specifically on reading trends or clusters in the graph. Going further than the graphs, using their personal knowledge, six students added other interpretations to the data. Some created explanatory text, such as the heart pumps more blood for the active rate, and "less average" for the girls so girls are more healthy. In contrast to this declaratory text, some students speculated on causes, e.g., the difference might be due to boys being into video games or Mrs. M. might be skinnier than everyone. Some of the students made more than one type of comment for a graph.

From classroom to interview (Research Question 2)

In the interviews, the six attributes (three categorical and three numerical) provided the opportunity for students to choose up to six types of plots. No suggestions of graph types were made by the interviewers, so all were "spontaneous" from the students. The 12 students each produced between 2 and 5 plots for a total of 42 graphs. Only one student plotted two categorical attributes in bins whereas 10 used bins to consider one attribute of each type (e.g., favourite activity and weight) (a total of 16). Six students produced 8 plots with bins using two numerical attributes. Seven students also created 8 plots with a categorical attribute and a scaled numerical attribute. Four students produced a mix of bins and scaled plots for two numerical attributes (a total of 5) but only one went on to produce a scatterplot. Two other students produced scatterplots having elsewhere created a scaled numerical and categorical plot (a total of 3).

In the interview, 10 students provided context from the individual data values, such as the highest and lowest values for an attribute. One student clicked on icons in a scatterplot to discuss values for a third attribute found on the data card. Nine of the twelve students made more general comments about the data set specifically from the relationships portrayed in the graphs they produced. These included trends such as more sport means less weight, younger are more active, and those who weigh more eat more fast food. Six students went on to add more to their explanations from their life experience. Some created stories to match what they observed, such as those who watch TV see junk food, and those who watch TV are lazy. Others speculated on reasons for the trends, rather than making declarative statements, such as maybe they are lazing around and watching TV and not getting up and doing anything. Two students gave advice to students in the data set, e.g., they need to go out more and they should cut down on fast foods and eat vitamins.

The performance of the students in the interviews where they encountered twice as many attributes as in the classroom indicated that they had internalised many of the graphing skills presented there. They showed flexibility by creating plot types that they had not used in the classroom. At times students were content to use a graph with bins rather than scaled axes for numerical attributes (type C in Figure 1) to interpret the data because the particular trend could be seen easily there. Although perhaps more experienced graph creators might always use scaled axes for numerical attributes, these students showed that such conventions are not always needed.

The students' use of their contextual knowledge was more varied in the interviews. They made more comments on individual cases than in the text boxes created in the classroom. This was most likely because (i) there would have been initial oral discussion of classroom data among students as they were entered; and (ii) the data in the interview were all new and hence it was interesting to check out the characteristics of the people represented. The greater number of attributes in the interview data may have contributed to more extensive discussions and the students may have also found it easier to articulate their thinking verbally than in text boxes.

Implications for the curriculum and teaching

This study shows that students in grade 5/6 are able to pick up graphing skills and application of contextual understanding in meaningful ways. Of particular interest is the ability to deal with two attributes in relationships, with bins or scatterplots, when scatterplots are not specifically mentioned in the Australian Curriculum – Mathematics until Year 10 (ACARA, 2010, p. 46). This seems to be an oversight. The first mention of considering two attributes, by creating side-by-side column graphs in Year 6, appears to be mentioned later and be more basic than is necessary. From the results of this study, it is seen that students can deal with two attributes and create and interpret relatively sophisticated representations by this time. The mention of considering two attributes in the curriculum is often done with no suggestion of the types of graphs that should be employed (e.g., in Year 7, where "investigating issues" has elaborations about two attributes but with no specific graph types suitable for the level).

The fact that the curriculum mentions creating graphs with (and without) digital technology from Year 3 is appropriate and encouraging. After the initial classroom introduction, the students in this study required virtually no assistance in working with *TinkerPlots* to create graphs.

The first specific mention of context in relation to Statistics and Probability in the Australian Curriculum – Mathematics is in Year 5, with reference to interpreting different graphs (ACARA, 2010, p. 29).

Context is implicit, however, in all grades in the elaborations, often from Year 4 in relation to examples presented from the media. The evidence from this study supports this recommendation—but teachers need to be aware that discussion of context needs to be developed from two standpoints: that of interpreting contextual information specifically available within the graph itself (e.g., trends observed) and that of the extra real-world understanding brought to the data set by the students' life experiences. Both aspects of context are essential for students to become critical statistical thinkers (Mooney, 2010). The link to other areas of the curriculum is crucial to provide the background for students to be able to go beyond the information provided in a graph and perhaps ask further questions.

The students in this study are typical of students from working class and rural schools in Australia. They hence challenge the writers of the Australian Curriculum—Mathematics to acknowledge that with creative software, it is possible to progress rapidly to meaningful representations of two attributes and make sense of these in real-world contexts.

Particularly at the primary school level, the opportunities to introduce cross-curricular contexts that link to the statistics component of the Australian Curriculum—Mathematics abound, especially if constructivist software is available for students to use. After developing the basic concept of graph, software can assist in the rapid consideration of various representations of data sets, in order for students to choose the one that best suits the story they wish to tell. By the time students reach secondary school, they should have a repertoire of graph types that can be used in various subject areas, and the ability to use them. It is then up to teachers across the curriculum to collaborate so that all teachers, not just teachers of mathematics, are aware of the possibilities that can be further developed with their students.

The Australian Curriculum—Mathematics mentions the use of media extracts in several places to motivate critical thinking in statistics. These opportunities not only may be based on unusual or inappropriate use of statistics but also may lead to in-depth cross-curriculum collaboration involving current real-world contexts. This will happen most effectively if all teachers are aware of the possibilities, including the application of "digital technologies."

Acknowledgements

Key Curriculum Press supplied *TinkerPlots*: Dynamic Data Exploration to the school and Australian Research Council grants LP0669106 and LP0560543 supported the research. The authors thank Julie Donne who conducted half of the interviews.

References

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2010). *The Australian Curriculum—Mathematics, Version 1.1 (13 December, 2010).* Sydney, NSW: ACARA.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Biggs, J. B., & Collis, K. F. (1991). Multimodal learning and the quality of intelligent behaviour. In H. A. H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-76). Hillsdale, NJ: Lawrence Erlbaum.

Chick, H. L., & Watson, J. M. (2001). Data representation and interpretation by primary school students working in groups. *Mathematics Education Research Journal*, *13*, 91–111.

Fitzallen, N., & Watson, J. (2010). Developing statistical reasoning facilitated by *TinkerPlots*. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society*. *Proceedings of the 8th International Conference on the Teaching of Statistics, Ljubljana, Slovenia, July* [CDRom]. Voorburg, The Netherlands: International Statistical Institute.

Konold, C., & Miller, C.D. (2005). *TinkerPlots: Dynamic data exploration*. [Computer software] Emeryville, CA: Key Curriculum Press.

Mooney, G. (2010). Preparing students for a data rich world: The case for statistical literacy. *Curriculum Perspectives, 30*(1), 25–29.

National Curriculum Board. (2009). *The shape of the Australian Curriculum*. Barton, ACT: Commonwealth of Australia.

Rao, C. R. (1975). Teaching of statistics at the secondary level: An interdisciplinary approach. *International Journal of Mathematical Education in Science and Technology*, *6*, 151–162.

Watson, J. M., & Donne, J. (2009). *TinkerPlots* as a research tool to explore student understanding. *Technology Innovations in Statistics Education*, *3* (1), 1–35. Retrieved February 15, 2011, from http://repositories.cdlib.org/uclastat/cts/tise/vol3/iss1/art1/

Watson, J., & Fitzallen, N. (2010). *The development of graph understanding in the mathematics curriculum: Report for the NSW Department of Education and Training.* Sydney: State of New South Wales, through the Department of Education and Training. Retrieved February 15, 2011, from http://www.curriculumsupport.education.nsw.gov.au/primary/mathematics/assets/pdf/dev_graph_undstdmaths.pdf