DEVELOPING STATISTICAL REASONING FACILITATED BY TINKERPLOTS

Noleine Fitzallen and Jane Watson
Faculty of Education, University of Tasmania, Australia
Jane.Watson@utas.edu.au

Of interest in this study, based in a class of 26 Year 5/6 students, is the way TinkerPlots facilitated the development of statistical reasoning for students with no previous data handling experience. The class completed four lessons related to data collection, data representation, data summary, and data inference based on a sporting activity, where they had recorded their heart rates before and after the event and used the software TinkerPlots to analyse the data. A month later 12 of the students were individually interviewed with a three-part protocol using TinkerPlots to assess their reasoning in relation to comparing two data sets, to hypothesising relationships and providing evidence, and to interpreting differences in large data sets. The consolidation and transfer of reasoning evidenced in the student interviews demonstrated the value of employing TinkerPlots in the classroom.

BACKGROUND

The development of school students' reasoning in various areas of statistics has been the focus of much research in recent years. Watson (2006) for example describes developmental research across all aspects of the curriculum from sampling to beginning inference, including data representation, averages, chance, and variation. Following the development of benchmarks, how to facilitate and enhance learning with classroom intervention has become the next focus of research. Classroom research has included the use of concrete materials (Jones, Langrall, Thornton & Mogill, 1999), the collaborative grouping of students (Chick & Watson, 2001), the use of particular teaching sequences (Pfannkuch, 2006), the use of meaningful contexts (Watson, Fitzallen, Wilson & Creed, 2008), and the use of dedicated software (Ben-Zvi & Arcavi, 2001; Rubin & Hammerman, 2006; Watson et al., 2008), as well as various combinations of these.

This informal study follows through on developing statistical reasoning employing the software TinkerPlots (Konold & Miller, 2003). The popularity of TinkerPlots as a teaching tool is seen in the reports of Ben-Zvi and Arcavi (2001) and Konold, Harradine and Kazak (2007). As well, recently, TinkerPlots has been viewed as a research tool to explore student understanding after students have developed sufficient familiarity with the tool to be able to use it to display their analysis of a problem (Watson & Donne, 2009). This report further combines these uses of the software by describing how it was used as a teaching tool in the classroom and then how it was used as a research tool to evaluate the development of student understanding.

METHODOLOGY

Sample and procedure

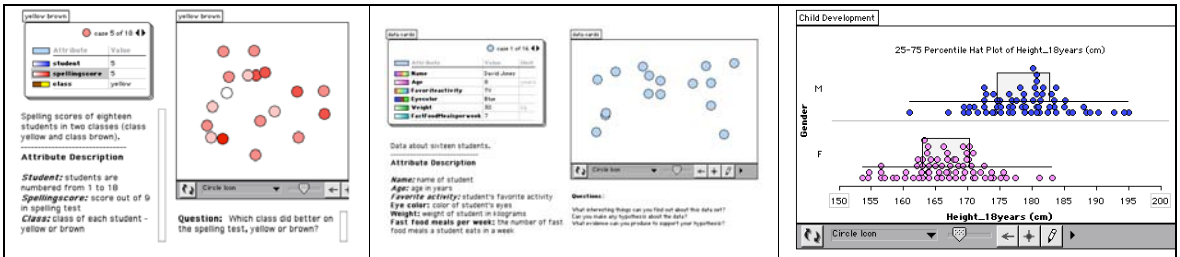
A class of 26 grade 5/6 students (aged 10-12 years) was chosen in a school involved in a professional learning program through which it was given TinkerPlots. These students were taught a series of four lessons by the first author with the aims of developing an appreciation of basic statistical analysis especially related to comparing two data sets, and introducing the fundamental features of TinkerPlots. With their usual classroom teacher students had collected data on their resting and active heart rates. The four lessons began with brain-storming about taking a person's pulse rate and discussing why heart rates might change at different times. This was followed by introducing the students to TinkerPlots and helping them enter their class heart rate data. Students were encouraged to create a plot as they entered the data in order to watch the distribution grow. After some time to explore the software on their own, they were shown how to create stacked dot plots and place a hat plot over the top of the data, with a discussion about the middle 50% of the data being under the crown of the hat and the bottom and top 25% being under the brims.

After discussing their expectation of a difference in heart rates, various ways of providing evidence for a difference were suggested. Students were asked to insert text boxes in their TinkerPlots files and write their responses to a set of questions about changing heart rates. There was much teacher interaction with the students as they worked, individually, in groups, and as a whole class. Of the eight elements of Pfannkuch's Beginning Inference Framework (2006), seven

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

were covered repeatedly throughout the lessons: hypothesis generation, summary (using graphs), shift (comparing positions of hats), signal (the middle 50% of the data), spread, making sense of context, and considering individual cases. The eighth element, sample and population relationship, was not a focus with this class. Toward the end of the lessons, the scatterplot was introduced with a discussion of whether students expected there to be an association between resting and active heart rates. Students were asked to save their TinkerPlots files periodically and most students' work over the lessons was available for analysis.

A month later, with no exposure to TinkerPlots in the intervening time, 12 of the students (5 males and 7 females) from the class were chosen by their teacher based on ability to converse with the researchers and parental permission. They were interviewed by the second author and another researcher, using a protocol based entirely within TinkerPlots (Watson & Donne, 2009). The protocol consisted of three scenarios that students were expected to analyse within TinkerPlots. The first, called Comparing Groups, provided four different pairs of data sets of spelling scores for classes, distinguished by colour, e.g., the "red" class and the "blue" class. Students were asked to determine for each pair which class had done better; the class sizes ranged from 6 students in each to 21 and 36 in two classes. The second scenario, called Data Cards, presented students with data on 16 young people: name, age, eye colour, weight, favourite activity, and number of fast food meals eaten per week. Students were asked to form hypotheses about the data and produce plots in TinkerPlots to support or refute them. The third scenario, called Child Development, presented students with three stacked dot plots, including hats, that contained the heights of 136 male and female children at ages 2, 9, and 18 years. Each stacked dot plot was separated by Gender, making it possible to compare male and female heights. Students were asked to consider the differences between the boys and girls and changes over time, and to suggest hypotheses about growth. They were further asked if the graphs supported their hypotheses and could change the plots if they desired. Three extracts from the scenarios are shown in Figure 1. The student interviews were video-taped, transcribed, and the TinkerPlots files created by the students were saved throughout the interview to provide evidence of the student discussion observed.



Scenario 1: Comparing Groups	Scenario 2: Data Cards	Scenario 3: Child Development

Figure 1. Extracts from research interview protocol

Analysis

The students' engagement with the classroom activities was assessed by viewing their saved TinkerPlots files. One of the twelve students had deleted rather than saved his files and it was surmised in two cases that a plot associated with a description in a text box had been over-ridden. The criteria for assessing engagement were whether the students had created a stacked dot plot, had written in a text box to comment on it, had created a scatter plot, and had written comments on it in a text box. Each student's level of engagement was given a rating of low, medium, or high based on the extent to which these tasks were completed. Some students used hats and reference lines to add detail to their comments and some also used bins in their responses.

With the two objectives in mind, of documenting each student's level of understanding of the statistical content and of suggesting the level of support provided by the software, the following schema were adopted for the interviews of the students. Transcripts and saved TinkerPlots files were read in relation to Pfannkuch's criteria for developing informal inference. The number of elements observed and how they were integrated to create coherent arguments within and across the three scenarios resulted in the observation of three levels of developing understanding; these are summarized in Table 1. The use of TinkerPlots representations was integral to all responses and hence influenced the levels of response achieved. Examples are provided of the different levels and types of support generated from the software.

Table 1. Interview Criteria for Levels of Statistical Understanding

	Level 1	Level 2	Level 3
Scenario 1: Comparing Groups	Examine individual data cards and/or only use bins	Include stacked dot plots and recognition of unequal groups	Use stacked dot plots and proportional reasoning for unequal groups
Scenario 2: Data Cards	Focus only on individuals and one variable at a time	Consider pairs of attributes in bins, scaled or scatter plots	Consider 3 attributes and develop hypotheses with evidence
Scenario 3: Child Development	Look at individual icons, e.g., the extremes	Look at regions of the plots and perhaps the hats	Develop hypotheses and find evidence, e.g., change the axes
Overall Protocol	Focus on individual elements of Pfannkuch's model, often including strong interest in individual cases	Take in multiple elements in a linear fashion, with ability to consider 2 attributes at once	Consider most of Pfannkuch's elements providing evidence for hypotheses, attempting to consider 3 attributes at once

RESULTS

Classroom exposure and engagement. The files saved by students from their classroom work with TinkerPlots and the heart rate data provided the evidence for the level of engagement of the students with the features of TinkerPlots in relation to the questions set about differences between males and females and the relationship of resting and active heart rates. Table 2 indicates the observed features (X) and a "T" in the cell indicates that text was used to explain the feature. A question mark indicates a text box describing a stacked dot plot and it was assumed that the plot had been modified without being saved. Although the material suggested low, medium, and high engagement, except for S2 with missing files, it was concluded students had the prerequisites for understanding the interview protocol.

Table 2. In-class engagement with TinkerPlots

Student	Bins	Stacked Plots	Hat	Reference lines	Scaled by bins/groups	Scatterplots	Engagement Overall
S1		X T	X	X	X		Low
S2	—	—	—	—	—	—	NA
S3	X	? T	X	X	X	X T	Medium
S4		X T	X	X	X	X T	Medium
S5		X T	X	X	X		Medium
S6		X T	X	X	X	X T	High
S7	X	? T				X	Low
S8		X T	X	X	X	X T	Medium
S9	X	X T	X	X		X T	High
S10	X	X T	X	X	X		Low
S11		X T	X	X	X	X T	High
S12	X	X T	X	X	X		Low

Interviews

Table 3 contains information for the students from the interview roughly parallel to that in Table 2. An Individual Case column was added to reflect the focus on the individual cases in the data cards or the table throughout the interview. Because the students had not entered the data themselves they were more interested in interrogating the information in the cards than would have appeared from the classroom files. Also there are no T's in Table 3 because the students answered questions orally for the interviewer. Comparing the two tables it is seen that whereas nine students used reference lines in the classroom only one did in the interview and the lines were used in the third scenario with the large data set presented in stacked dot plots. On the other hand, whereas five students used bins in the classroom all did in the interview. The reverse occurred for hats and scatterplots, with more students using them in the class than the interview.

Table 3. Interview engagement with TinkerPlots

Student	Individual Cases	Bins	Stacked Plots	Hat	Reference lines	Scaled by bins/groups	Scatterplots
S1		X	X	X		X	
S2	X	X	X			X	
S3		X	X	X		X	X
S4	X	X	X	X			
S5	X	X					
S6	X	X	X		X	X	X
S7	X	X	X	X		X	
S8	X	X	X	X		X	
S9	X	X	X			X	
S10	X	X	X			X	
S11	X	X				X	X
S12	X	X	X			X	

To consider the students' levels of performance in the three scenarios and overall, and how TinkerPlots may have assisted or diminished the understanding of the statistical concepts in Pfannkuch's (2006) framework, Table 4 contains the level assessed for each scenario and overall (cf. Table 1) for each student. The association with the assessed degree of engagement in the classroom (Table 2), where each variable has three hierarchical categories, is not strong; however, 50% of the students achieved the equivalent degree of engagement as level of understanding.

Table 4. Levels of Understanding within and across the Interview Protocol

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Sc 1: Comparing Groups	3	1	3	3	1	2	2	1	2	1	1	1
Sc 2: Data Cards	2	1	3	2	2	3	1	1	2	2	3	2
Sc 3: Child Development	2	1	3	1	2	2	2	1	3	1	2	1
Overall	3	1	3	2	2	3	2	1	3	1	2	2

For each student one example is presented in Figure 2 to indicate the type of support supplied by TinkerPlots as part of the justification for the answer, suggestion or hypothesis presented in the student's response. Whether the feature was a help or hindrance is discussed. Not always did the plots support the highest level of response displayed but are shown to demonstrate the variety of plots used by students. Figures 2a to 2c are related to the first scenario, Comparing Groups. In Figure 2a, the data in four bins led S7 to suggest appropriately that the Green class did better on the spelling test. In Figure 2b however there are several conclusions that fit the data, with the total scores for the two classes the same and the distributions symmetrical, hence assuming the Brown class did better based on 4 bins, as done by S2, is not adequate. Figure 2c was used by S1 to say "Black [did better], because [Pink] have more people down here. They're the same at the top… [for] Black, most of the people are higher on the hat."

Figures 2d to 2k are associated with responses to the second scenario, Data Cards. In all three scenarios S9 was interested in individual cases and for the Data Cards he began by exploring the weight of each student in the data set and creating imaginative stories. Later, however, S9 used the individual cases for evidence of the relationship of a third attribute to the two that had been plotted. S4 constantly changed the attributes on the plots, highlighting particular values (as shown in Figure 2e) and making up stories with reasons for the values being in the cells. There was considerable interest in students' favourite activities and their weight (Figures 2f, 2g, 2h). S10, S5 and S11 displayed this in different ways, the first two with bins and the last with scaled values for each activity. Creating the plot in 2f, S10 struggled to reconcile the heavy and light people who watched TV. S5 concluded from Figure 2g, "These people over here who play sport aren't going to weigh as much as these over here who watch TV," not worrying about the lighter case. S11 hypothesised, "That more people [who] sit around and watch TV and stuff are heavier… than people who do sport. I think it might depend on their age," adding a third attribute, which was then checked by clicking on icons to read cases. S12 began with a plot similar to S8 but then separated the data on the axis. Both provided a similar interpretation about fast food meals being related to

TV-watching, perhaps because of the ads on TV. Figure 2k is the scatterplot S3 used to conclude, “This shows me that the people who don’t have as many fast food meals a week weigh less, which would probably tell me that they are the younger children, more likely because they are around 20kg.”

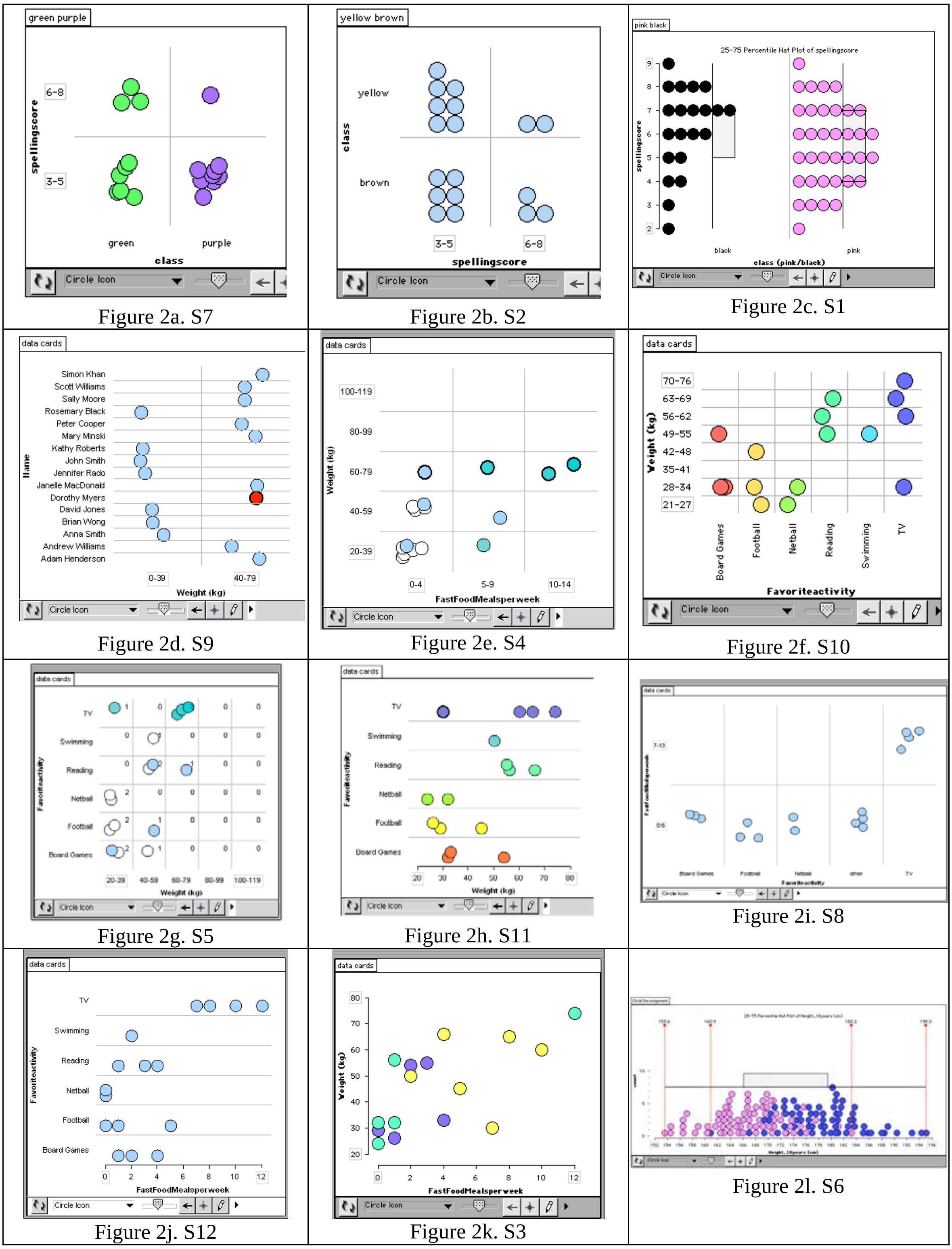


Figure 2a. S7 — Figure 2b. S2 — Figure 2c. S1

Figure 2d. S9 — Figure 2e. S4 — Figure 2f. S10

Figure 2g. S5 — Figure 2h. S11 — Figure 2i. S8

Figure 2j. S12 — Figure 2k. S3 — Figure 2l. S6

Figure 2. TinkerPlots plots from student interviews

Finally the only use of reference lines was in the third scenario, the final graph of which is in Figure 2l. S6 recombined the data (cf. Figure 1), coloured by Gender, and used the reference lines to discuss the differences and similarities, the spread in the extremes with no overlap, concluding overall, “Boys taller at some times and not others.”

DISCUSSION

The plots in Figure 2 were chosen to show the variety of representation selected by the students as well as the support provided for conclusions drawn. The nature of the Data Cards scenario with a small data set led to students using bins more in the interview than the classroom. The fact that bins were as helpful as fully separated data in some cases illustrates the flexibility of TinkerPlots in allowing students to have choice in what appealed to them visually. Not always, e.g., S2 in Figure 2b, did using certain bin configurations help in answering a question but only in the Comparing Groups scenario did it at times appear a disadvantage. In the Data Cards scenario students were more likely to have problems because of not appreciating all of the attributes in relation to the context than because of difficulty with the plots. Many, however, did have the background to suggest multiple associations, even if sometimes frustrated at lacking the skill to display three attributes easily. This led some to click on the icons to check the data cards for confirmatory data for their hypotheses. This was a more sophisticated purpose than the curiosity of some students in initially clicking through the cards. The nature of the tasks set would appear to encourage different TinkerPlots features to be used to support different arguments.

In a very short period of time, the students had developed the ability to be independent users of TinkerPlots. They had shifted from following the procedures taught in the classroom to using TinkerPlots creatively, discriminating appropriately for their needs. They demonstrated that they were able to transfer the data analysis skills developed through exploring a small data set in the classroom to reasoning about a variety of data representations in the interview protocol. In doing so, they created plots that made sense to them and used the plots effectively to support their thinking about the data. In addition, they were able to move back and forth between hypothesis creation and plot creation to make sense of the data–a strong affordance of TinkerPlots.

ACKNOWLEDGEMENTS

Key Curriculum Press supplied TinkerPlots 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

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics, 45*, 35-65.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30*, 487-519.

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.

Rubin, A., & Hammerman, J. K. (2006). Understanding data through new software representations. In G. F. Burrill (Ed.), *Thinking and reasoning with data and chance* (pp. 241-272). Reston, VA: National Council of Teachers of Mathematics.

Konold, C., Harradine, A., & Kazak, S. (2007). Understanding distributions by modelling them. *International Journal of Computers for Mathematical Learning, 12*, 217-230.

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

Pfannkuch, M. (2006). Informal inferential reasoning. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics: Working cooperatively in statistics education,* Salvador, Brazil. [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Erlbaum.

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

Watson, J. M., Fitzallen, N. E., Wilson, K. G., & Creed, J. F. (2008). The representational value of hats. *Mathematics Teaching in the Middle School*, *14*(1), 4-10.