EXTENDING THE CURRICULUM WITH TINKERPLOTS: OPPORTUNITIES FOR EARLY DEVELOPMENT OF INFORMAL INFERENCE

Noleine Fitzallen and Jane Watson
University of Tasmania, Australia
noleine.fitzallen@utas.edu.au

Although increasing in recent years, the statistics content in the curriculum does not always align with students' ability to develop an understanding of those key concepts. In this paper, two examples are presented that challenge the school curriculum by introducing students to activities that focus on decision making under uncertainty earlier than when first acknowledged in the curriculum. The first study investigated Grade 5 and 6 students' understanding of covariation; the second investigated Grade 10 students' understanding of resampling. Common to both studies were the emphases on opportunities for the development of informal inference and the application of the software package, TinkerPlots. In both cases, many students based their confidence in conclusions on visual aspects of the strength of an association or shape of a simulated distribution. Besides making suggestions for future curriculum improvement, this presentation draws attention to pointers for development of student understanding.

BACKGROUND

In light of interest in revised statistics curricula, the newly developed *The Australian Curriculum: Mathematics (ACM)* (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2013) provides the setting within which the two case studies reported in this paper are situated. The curriculum has three strands: Number and Algebra, Statistics and Probability, and Measurement and Geometry. In previous iterations of the curriculum statistical concepts were included in the strand of Chance and Data. The change to the Statistics and Probability strand reflects the world view of statistics as expressed by the National Council of Teachers of Mathematics (1989, p. 170) and acknowledges the important role that statistics plays in contributing to the overarching ideals of the Australian Curriculum, which are to "develop successful learners, confident and creative individuals, and active and informed citizens" (Ministerial Council on Education, Employment, Training and Youth Affairs, 2008, p. 7).

One of the dilemmas for writers of curriculum documents at the school level in the area of statistics is how far to go in decision making with data and how uncertainty is catered for in decision-making. It is possible to cover the skills of data collection, data representation, data reduction, and ask students to make a decision solely for the data collected. A typical early question would be, "What fruit does our class like best?" The uncertainty associated with such an investigation could be related to whether the data were recorded properly, whether the students would answer the same way tomorrow, how to cater for a tie in the data, or whether a plot showed the data properly. A much more complex and uncertain situation arises when the question is "larger" than the data set available to answer it. For example, based on students in our class, "What is the favorite fruit of the students in our school, or in our city, or in our country?" Uncertainty multiplies enormously: from the interpretation of the question, through the same issues about the data from the class and their analysis, to the confidence we have in making a decision for the larger set based on the data collected from the class sample. At the school level, this type of decision-making is often called informal inference, to distinguish it from the formal inference carried out by statisticians based on theoretical distributions and perhaps hypothesis tests. Makar and Rubin (2009) describe informal inference as decision-making where generalizations are made beyond the data collected based on the evidence presented, acknowledging the uncertainty resulting from the process. It is not possible with certainty to predict the favorite fruit of children in the school or country based on the data collected from a class. Issues important for students to discuss include possibly increased confidence (less uncertainty) if a larger or more representative (perhaps random) sample were used (Shaughnessy, Chance, & Kranendonk, 2009).

In Australia, the *ACM* (ACARA, 2013) section on Data Representation and Interpretation provides descriptors that cover asking questions, collecting data, producing graphs, and calculating summary measures. Although the introduction to the document states that "Students recognise and analyse data and draw inferences" (p. 5), the word is not used again and descriptors include

describing, comparing, interpreting, and evaluating data sets and data displays. Nowhere in the content of the curriculum is uncertainty noted as an aspect of an investigation. Variation, one of the major sources of uncertainty in decision-making, is mentioned in two descriptors in this section of the curriculum. There is no reference to uncertainty in decision-making associated with variation.

Reading the content descriptors for Data Representation and Interpretation teachers would not be encouraged to plan investigations that asked students to generalize beyond the data they collected or were given. This paper reports on activities at two grade levels in Australia that challenged this reading of the Australian curriculum. Both investigations included the use of the software, *TinkerPlots: Dynamic Data Exploration* (Konold & Miller, 2011), as a learning and discovery tool. The software provided an efficient, user-friendly environment that allowed timely interactions impossible in a non-technological setting (Konold, 2007; Konold & Lehrer, 2008).

CASE 1. STUDENTS' DEVELOPMENT OF UNDERSTANDING OF COVARIATION

The students in this study either had not used *TinkerPlots* previously or those that had were considered novice users. Hence, the study was set up to build their knowledge of how to use the technology and how to interpret the graphical representations at the same time as learning about the statistical concepts that underpin covariation (Fitzallen, 2008). The learning sequence the students worked through was designed to introduce the features of the software together with the statistical ideas of variation, distribution, covariation, and informal inference, in order to develop the skills necessary to use the statistical ideas to make and justify informal inferences evidenced from the graphical representations created. It was delivered over a period of six weeks, which included two 45-minute sessions each week with the students working in pairs with the teacher-researcher.

Upon completion of the learning sequence, 12 Grade 5 and 6 (11-12 year old) students were interviewed as they worked individually through an activity set up in *TinkerPlots*. The activity involved using a data set (n=200) to investigate the relationship between the physical attributes of the cases, such as height, foot length, belly button height, and hand span. The students were given the freedom to create the graphical representations of their choosing to determine if there was a relationship evident. Throughout the interview the students used the features of the software both to create the graphical representations and to explain the relationships they identified in the graphs. The data gathered from the interviews were captured by on-screen video, which were analyzed using thematic analysis to determine the students' levels of understanding of covariation (Fitzallen, 2012) and the strategies the students employed as they engaged with the software learning environment to work through a statistical investigation (Fitzallen, 2013).

Covariation is introduced in the *ACM* (ACARA, 2013) in Grade 10 when students "Use scatterplots to investigate and comment on relationships between two numerical variables" (p. 71). In the preceding years, the curriculum states that students are to use a variety of graph types when investigating data but the ones suggested are limited to column charts, dot plots, and picture graphs, which are utilized to display univariate data. Somewhat more encouraging is the observation that in the U.S. the *Common Core State Standards: Mathematics (CCSSM)* (Common Core State Standards Initiative [CCSSI], 2010) introduces scatter plots in Grade 8, "for bivariate measurement data to investigate patterns of association between two quantities" (p. 56). The decision to provide the opportunity for Grade 5 and 6 students to explore the notions of covariation was based primarily on the work of Moritz (2004), whose results suggested young students were able to describe covariation and identify trends in data. This study varies from that of Moritz because student learning was situated within a data analysis software package, whereas the work undertaken by the students in the Moritz study was paper-based.

As expected after a teaching intervention of this kind, all the students were able to describe the data in terms of covariation and make informal inferences about the data from the evidence in the graphical representations created (Fitzallen, 2012). They did, however, demonstrate three levels of understanding of covariation, which were determined using the developmental levels of the SOLO taxonomy (Biggs & Collis, 1982). At the uni-structural level, the students identified correctly when there was a trend evident in a graph but only described the relationship in terms of one variable at a time and did not use information from the graphical representations to support their conclusions about the data. At the multi-structural level, the students identified the trend in the data and identified the variation within the trend. They were, however, unable to explain how the

variation identified informed their overall conjectures and inferences. Students categorized to be operating at the relational level described the relationship between two attributes from both local and global perspectives. They brought together information about specific data points, the variation within the graphs and the trend identified, as well as their knowledge of the context to justify their informal inferences.

It was also determined that the students employed three dominant strategies when creating and interpreting graphical representations with *TinkerPlots*. The strategies were Snatch and Grab, Proceed and Falter, and Explore and Complete (Fitzallen, 2013). Students who adopted the Snatch and Grab strategy worked aimlessly in *TinkerPlots*. They clicked on buttons and moved things around in anticipation that something helpful would appear on the screen. Often, they did not stop to evaluate the differences made by their actions. The students who adopted the Proceed and Falter strategy used pre-established patterns of behavior to create graphs they were familiar with and then hesitated when they could not use the graphs to answer the questions about the data. They appeared not to be able to make the link between what was produced and what was needed to be produced to answer the questions. The Explore and Complete strategy was employed by students who were purposeful in their actions. They made correct decisions about which graph type would assist in answering the questions and added additional features to the graphs to determine if more information could be gleaned. When they did this, they took the time to evaluate the changes to determine if they were helpful before exploring the data further.

When the students' levels of understanding of covariation were compared with the interaction strategy each student exhibited, the results suggested that the three strategies adopted had varying levels of productivity in terms of potential learning outcomes for covariation (Table 1). There is also the possibility that there is a relationship between the development of understanding of covariation and the development of strategies students employ when working within data analysis software environments. Further research is required to determine the validity of that conjecture.

Table 1. Students' levels of understanding of covariation and interaction strategy.

Level of understanding of covariation	Interaction strategies		
	Snatch and Grab	Proceed and Falter	Explore and Complete
Uni-structural	Natalie, Natasha, Kimberley, Johnty	Jake, Rory	
Multi-structural		Shaun, Blaire, Jessica	
Relational			James, William, Mitchell

CASE 2. RESAMPLING WITH *TINKERPLOTS*

The second study focuses on the description of classroom experiences of students in a Grade 10 mathematics class. The examples provided are indicative of the classroom interactions observed by the researcher and the student work produced by the class. The students were introduced to *TinkerPlots* as part of a 5-week unit on probability and statistics. Probability problems were introduced to illustrate the simulation features of the Sampler in *TinkerPlots* and how they could be used to collect repeated samples (with replacement) and fixed-size random samples (without replacement) from a large finite population to confirm students' predictions of models to answer questions (Stack & Watson, 2013). These preliminary activities led to the informal inference investigation where students collected their class data to answer the question of whether it is easier to memorize nonsense or meaningful words. The activity was based on the work of Shaughnessy et al. (2009), which included the modeling of classroom discussion led by the teacher. The *ACM* (ACARA, 2013) provides expectations for Grades 9 and 10 that anticipate the class collecting and analyzing the data for the class. For example, "Compare data displays using mean, median and range to describe and interpret numerical data sets in terms of location (center) and spread" (Grade 9, p. 66) and "Construct and interpret box plots and use them to compare data sets" (Grade 10, p. 71). The purpose of the unit, however, was to extend the question to "Using the data from our class can we conclude that Grade 10 students in Australia would find it easier to memorize meaningful or nonsense words?" It would have been possible to extend the curriculum

expectation to interpret and construct box plots for the two sets of data students collected on their numbers of words memorized in the two conditions and use criteria devised by Wild, Pfannkuch, Regan, and Horton (2011) for the degree of overlap of the box plots for the class sample size to generalize beyond their class. Instead, based on the arguments of Cobb (2007) and others, it was decided to use the randomization method of resampling, which when carried out in *TinkerPlots* becomes a constructivist activity with students able to "see" each step (e.g., Watson & Chance, 2012).

The proposition was set in terms of an understanding of random samples, which had been in the curriculum in Grade 8 and was reviewed using examples from probability and the Sampler in *TinkerPlots*. The question the students were asked to consider was whether the difference of 4 that their class found for the medians of the two kinds of words memorized was just a random result or whether it was genuine, leading to the conclusion that meaningful words are easier to memorize than nonsense words. Initially they concluded that meaningful words were easier for their class and they were a "typical" Grade 10 class. Raising the question of whether their result might just be random led to the suggestion that they do the tests again. They agreed, however, that this was "no good" because they had seen the words and practiced. This led to the possibility of assuming that if the data were random, what would happen if they randomly reassigned the data values to the two conditions, meaningful and nonsense? Would there be no difference in the medians or by chance would the difference be the same or greater than the difference of 4 observed for the class? First the students wrote their scores on sticky notes, which were randomly picked out of a box by two students, one for meaningful words and one for nonsense words. These notes were then stuck on a white board on the appropriate plot. The difference of medians for this random redistribution of values was 2, favoring meaningful, but less than the 4 observed in the class.

The question then arose about getting a feel for how unusual the class data were compared to many random reallocations for the data. At this point students were assisted to put their class data into the software and set up a way to reallocate randomly, or "resample," the data, placing values into the two groups, meaningful and nonsense (Watson, 2013; Watson & Chance, 2012). In *TinkerPlots* the Sampler resampled the data, which were then plotted in the same fashion as the original data and the medians marked. The software has a tool called a Ruler that was then set up to measure the distance between the two medians. For each run of the Sampler, creating new allocations to the two groups, a new measure was calculated. Finally there is a History button in *TinkerPlots*, which kept track of the measures from many resamples. These measures of difference were then plotted in order to find how many were equal to 4 or more. In other words, how unusual was the class difference for memorizing the two types of words, compared with random allocations of the data? If none or very few random resamples had differences of 4 or more, then the class difference value was unlikely to be random and likely to be due to a difference in the ease of memorizing meaningful or nonsense words. Three plots of this process are shown in Figure 1. On the left are the original numbers of words remembered split by type of word; in the middle is one random resample with the ruler measuring a difference of -4; on the right is a plot of 300 such random differences showing that only 6 (2%) were equal to 4 and none were greater than 4. There was still a decision for students to make as to whether this percentage was so small that indeed it was more difficult to memorize nonsense than meaningful words, or whether the class result was one of the 2% expected by chance to have a median difference of 4 in the long term. The students chose the former, based on their beliefs about the experiment as well as the data before them!
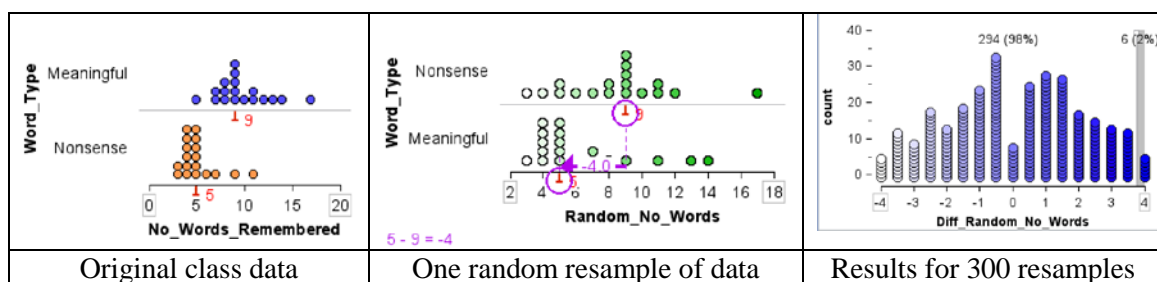

| Original class data | One random resample of data | Results for 300 resamples |

Figure 1. Sample of typical graphs produced.

DISCUSSION AND CONCLUSION

The two case studies presented in this paper illustrate that students are able to engage in thinking about complex statistical ideas, such as informal inference, that arise from investigations at a much younger age than expected by the current curriculum. Being able to explore the data in multiple ways with representations of their choosing allowed them to focus on the deeper, richer ideas of statistics. The ease with which students made changes to the software interface relinquished them of the burden of conducting multiple resampling operations or drawing new graphs each time the data were rearranged. The students were not stifled by the laborious task of generating multiple data representations. This suggests that there is the opportunity to capitalize on the students' ability to use the technologies available to develop intuitive ideas and foundation knowledge about statistical ideas. New educational technologies, such as *TinkerPlots*, provide the opportunity to include complex graphical representations and operations earlier in the learning progression than they are currently staged. Hence, there is no longer the need to introduce concepts in the curriculum in the same order as previously organized.

Returning to the Australian curriculum it is disappointing that there is no recognition of decision making under uncertainty as experienced in informal inference (Makar & Rubin, 2009). It is interesting that by comparison the *CCSSM* (CCSSI, 2010) have as two of the three headings under Statistics and Probability for Grade 7, "Use random sampling to draw informal inferences about a population" and "draw informal comparative inferences about two populations" (p. 50). For Grade 6, box plots are introduced as one of the ways of displaying numerical data (p. 45). The second activity described here, using box plots (not resampling) would satisfy the *CCSSM* long before Grade 10. Further, in the high school section of the *CCSSM* one of the examples presented under Making Inferences and Justifying Conclusions is "Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between two parameters are significant" (p. 82). Although specific year levels for high school are not mentioned for the topics in the *CCSSM*, there is recognition of this method for drawing conclusions. (Grade 11/12 curricula are not yet confirmed for Australia, hence no claims are made in comparison.)

Although the other section of the Statistics and Probability part of the Mathematics curriculum in Australia is called Chance, there is no attempt to connect the two sections. The only place the word uncertainty appears in the curriculum is in a section on Ethical understanding in the General Capabilities (ACARA, 2013). As well, there is no indication of a model for carrying out a complete statistical investigation, from asking a question in context to reaching a considered decision. The stages of an investigation are presented in a developmental fashion across the years for asking questions, collecting data, representing data, and summarizing data but hints for evaluating and interpreting do not help students to acknowledge the uncertainty that is involved at every stage of an investigation. The *CCSSM* (CCSSI, 2010) suffers in the same regard. Mathematics curriculum developers could learn from the *Australian Curriculum: Science* (ACARA, 2013) in its structure, if not all of the details, for the section Science Inquiry Skills. The subheadings are Questioning and predicting, Planning and conducting, Processing and analyzing data and information, Evaluating, and Communicating. Across the years in Mathematics it would be possible to provide descriptors indicating the stages in a statistical investigation, using the skills developed to that point.

In addition, embedding the use of technologies within an inquiry framework to target the development of specific statistical ideas will acknowledge the capacity students have to engage with those ideas as well as exploit the affordances of the technologies that facilitate students to conduct data explorations through an inquiry perspective. This has, however, implications for assessment practices. Questions about assessment need to be asked, such as, "Can and should assessment of student understanding be situated within the learning environment in which the statistical ideas were developed?" The first case study suggests there is a synergy between the development of ideas and the way in which students work within the technological learning environment and that meaningful assessment of the interaction can take place. Based on the work of Watson and Donne (2009), comparing assessment of tasks on paper and in a data analysis software environment, the second case study would appear to offer the same opportunity for assessing the interaction of the synergies as the first. It would be therefore prudent to investigate further the influence technologies have on learning in order to develop assessment practices that

acknowledge the way in which technologies empower students to explore statistical ideas and the different levels of thinking they may evoke, particularly when the software allows students to select and construct representations of their choosing.

REFERENCES
Australian Curriculum, Assessment and Reporting Authority. (2013). *The Australian Curriculum. Version 5.1, August 5, 2013.* Sydney, NSW: Author.

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

Cobb, G. (2007). The introductory statistics course: A Ptolemaic curriculum? *Technology in Statistics Education, 1*(1), Article 1. Retrieved from http://escholarship.org/uc/item/6hb3k0nz

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association for Best Practices and the Council of Chief State School Officers. Retrieved from www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Council on Education, Employment, Training and Youth Affairs. (2008). *The Melbourne declaration on educational goals for young Australians.* Canberra, ACT: Author.

Fitzallen, N. (2008). Validation of an assessment instrument developed for eliciting student prior learning in graphing and data analysis. In M. Goos, R. Brown, & K. Makar (Eds.), *Navigating Currents and Charting Directions* (Proceedings of the 31st annual conference of the Mathematics Education Research Group of Australasia, Brisbane, pp. 203-209). Sydney, NSW: MERGA.

Fitzallen, N. (2012). Interpreting graphs: Students developing an understanding of covariation. In J. Dindyal, L. P. Cheng, & S. F. Ng (Eds.), *Mathematics Education: Expanding Horizons* (Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia, Singapore, pp. 290-297). Sydney, NSW: MERGA.

Fitzallen, N. (2013). Characterising students' interaction with *TinkerPlots. Technology Innovations in Statistics Education, 7*(1), Article 2. [Online]

Konold, C. (2007). Designing a data analysis tool for learners. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 267-291). New York: Taylor and Francis.

Konold, C., & Lehrer, R. (2008). Technology and mathematics education: An essay in honor of Jim Kaput. In L. English (Ed.), *Handbook of International Research in Mathematics Education* (2nd ed.) (pp. 49-72). New York: Routledge.

Konold, C., & Miller, C.D. (2011). *TinkerPlots: Dynamic data exploration* [computer software, Version 2.0]. Emeryville, CA: Key Curriculum Press.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82-105.

Moritz, J. B. (2004). Reasoning about covariation. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227-255). Dordrecht: Kluwer.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Shaughnessy, J. M., Chance, B., & Kranendonk, H. (2009). *Focus on high school mathematics: Reasoning and sense making in statistics and probability.* Reston, VA: National Council of Teachers of Mathematics.

Stack, S., & Watson, J. (2013). Randomness, sample size, imagination and metacognition: Making judgments about differences in data sets. *Australian Mathematics Teacher, 69*(4), 23-30.

Watson, J. M. (2013). Resampling with *TinkerPlots. Teaching Statistics, 35*(1), 32-36.

Watson, J., & Chance, B. (2012). Building intuitions about statistical inference based on resampling. *Australian Senior Mathematics Journal*, *26*(1), 6-18.

Watson, J. M., & Donne, J. (2009). *TinkerPlots* as a research tool to explore student understanding. *Technology Innovations in Statistics Education*, *3*(1), Article 1. [Online]

Wild, C. J., Pfannkuch, M., Regan, M., & Horton, N. J. (2011). Towards more accessible conceptions of statistical inference (with Discussion). *Journal of the Royal Statistical Society A, 174*(2), 247-295.