INFERENTIAL REASONING: LEARNING TO "MAKE A CALL" IN THEORY

Chris J. Wild[1], Maxine Pfannkuch[1], Matt Regan[1] and Nicholas J. Horton[2]
[1]Department of Statistics, The University of Auckland, New Zealand
[2]Smith College, Northampton, United States of America
c.wild@auckland.ac.nz

Drawing on recent research in statistics education, the new New Zealand Statistics curriculum plans for three years of instruction in informal statistical inference to lay conceptual foundations for instruction in formal statistical inference. We discuss issues involved in formulating beginning versions of statistical inference and present some specific and highly visual proposals. These are built upon simple metaphors and novel ways of experiencing sampling variation. They are designed to give students practically useful tools for data analysis as well as underpinning more advanced and formal methods of making inferences to be encountered later. Our proposal uses visual comparisons to enable the inferential step to be made without taking the eyes off relevant graphs of the data. This allows the time and conceptual distances between questions, data and conclusions to be minimized, so that the most critical linkages can be made.

INTRODUCTION

School-level statistics for the bulk of students has suffered and stagnated for many years under a computational mentality pejoratively termed “meanmedianmode” and the “construct a graph” syndrome (Bakker, 2004). The focus on construction and the tools of statistics rather than statistical reasoning processes has resulted in a discipline that is perceived by many students and teachers as boring with little intellectual substance. Often descriptive statistics has been the only diet for students up to the penultimate year of high school, to be followed by an attempt to force-feed statistical inference, with its mathematical underpinnings, concepts, and reasoning in the final year. Recent research (e.g., Chance et al., 2004) on students’ reasoning about sampling distributions has demonstrated that their grasp of statistical inference will continue to be limited unless there are changes to the curriculum and teaching in the earlier years. Introducing statistical inference concepts in a single course is overwhelming and such research points to a need to provide a learning pathway over the years that introduces some of the “big ideas” behind inference. The challenge is to stage students’ reasoning in a way that will allow them to make sound informal or simplified inferences about populations from samples but at the same time will be gradually building an integrated conceptual mental schema which provides a solid foundation for statistical inference.

Technology is now opening up new opportunities to build a substantive pedagogical and cognitive base for statistical concept development such as inferential reasoning even at the school level. Furthermore, in the last decade EDA and new data visualization tools have led to a reassessment of current school-level curricula which, as Shaughnessy (1997) exhorted, should move away from “univariatitis” and concentrate on comparing distributions in order for students to learn how statisticians make decisions based on data. Such a reform in statistics education is promoted by the school-level GAISE report (Franklin et al., 2007, p. 6):

> A major objective of statistics education is to help students develop statistical thinking. Statistical thinking, in large part, must deal with this omnipresence of variability; statistical problem solving and decision making depend on understanding, explaining, and quantifying the variability in the data.

All of this, of course, leads to a need to devise and develop types of educational experiences that are very different from those that have been relied on in the past. Joan Garfield and Dani Ben-Zvi responded to the calls to focus on thinking skills from 1999 by organising the SRTL series of biennial international research forums with the aim of encouraging research focussed on students’ thinking and reasoning; SRTL stands for Statistical Reasoning, Thinking and Literacy. Initially there was a focus on different types of statistical reasoning, (SRTL-1, 1999; SRTL-2, 2001), then reasoning about variation (SRTL-3, 2003), followed by reasoning about distributions (SRTL-4, 2005). At the SRTL-4 Forum, Wild (2006) discussed conceptions of distribution, including different conceptions of populations and how samples relate to populations, and where these ideas were leading to. Discussions at SRTL-4 led to the realization that students

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

and teachers were having difficulties determining whether two groups were different. Part of the problem was that the students and teachers being studied did not know whether they were reasoning about the data as if it were the whole population or about an underlying population from which the data were a sample (Pratt et al., 2008). The researchers at the SRTL-4 Forum came to a consensus that students should be playing Game 2 and that the situation could be rectified by providing students with informal inference understandings that acted as a bridge between exploratory data analysis and formal statistical inference (Ben-Zvi et al., 2007). Therefore informal inferential reasoning became the subject of the next Forum (SRTL-5, August 2007). Konold and Kazak (2008) commented on the importance of this insight by the Forum, believing that the recognition that students needed deeper understandings of inference to make comparisons had the potential to result in an acceptance that chance or sampling behavior must be addressed.

RESEARCH ON INFORMAL INFERENCE

Research on informal inference dates back at least to Curcio (1987) who described three levels of graphical comprehension; read the data, read between the data, read beyond the data. Reading beyond the data was interpreted as making a prediction or inference. Watson (2006) asked the Grade 3 students she was studying to predict the travel mode of a student who was absent from the class that day, after they had graphed the mode of travel to school of students in a class. Bakker (2004) experimented with "growing a sample" activities for middle school students in order for them to develop intuitive ideas about sample and population without specifically referring to these terms. Ben-Zvi (2006) built on this work by providing a learning framework for young students, which started with the subjects a child interviewed, expanded to also include the subjects interviewed by a neighboring child and so on to ever increasing supersets. Thus these researchers were actively creating conceptual links between sample and population for young students.

Particularly for SRTL-5 in 2007, statistics education researchers have actively sought to research the nature of informal inferential reasoning in exploratory studies. The Forum collectively characterized informal inferential reasoning as:

> the cognitive activities involved in drawing conclusions with some degree of uncertainty that go beyond the data and having empirical evidence for them. Three principles appear to be essential to informal inference: (1) generalizations (including predictions, parameter estimates, and conclusions) that go beyond describing the given data; (2) the use of data as evidence for those generalizations; and (3) conclusions that express a degree of uncertainty, whether or not quantified, accounting for the variability or uncertainty that is unavoidable when generalizing beyond the immediate data to a population or a process. (SRTL-6, 2008, p. 3).

Zieffler et al. (2008, p. 44) proposed an informal inferential reasoning framework that combines several research perspectives into a working definition: "the way in which students use their informal statistical knowledge to make arguments to support inferences about unknown populations based on observed samples". According to Makar and Rubin (2009, p. 102), building a bridge to formal statistical inference via informal inferential reasoning provides "new opportunities to infuse powerful statistical concepts very early in the school curriculum and return the focus of statistics on a tool for insight into understanding problems."

In their exploration of how students might draw informal inferences, the SRTL-5 researchers took varied teaching and learning approaches. Ben-Zvi et al. (2007), Watson (2008), and Zieffler et al. (2007) assessed how students compared two distributions. The way their various students judged whether the observed difference in the centers was real or due to chance or sampling variability was quite different. Watson's Grade 7 students, who were using the hat feature in Tinkerplots, appeared to making a judgment about the magnitude of the difference. That is, if the difference was intuitively "significant" or "meaningful" to them, a notion more akin to practical significance than statistical significance. Working at the tertiary level, Zieffler at al. (2007, p. 23) immersed their students in activities that were designed to form conceptions about distribution, sample, sampling variability, including with and between group variability, and so forth. However, they "were disappointed when even the top student in the class failed to consider the variability within samples when comparing two samples, focusing only on differences in center". They conjectured that part of the problem might have been that the student experience was that of a "collection of separate tasks and methods" rather than a connected experience that could "lead to a unified understanding" (p. 26). However, all that students may see when they compare two sample distributions is the within sample variability. They may know intuitively that there will be variability if they take more samples but they may not appreciate the extent of the between sample variability, and that the difference of the observed direction could be reversed. Therefore, another underlying problem for students gaining a unified understanding may be that they have no basis, other than the difference between the centers, on which to make a decision when comparing groups even though they may be aware of sampling variability.

Across the studies of SRTL-5, student reasoning drawing upon the problem context seemed to override attention to the probabilistic nature of the data. The lack of attention to sampling variability may not be surprising when considering students' initial probabilistic understandings. Pratt et al. (2008, p. 126) described their Grade 6 students' naive informal inferences as "emergence-related (rather than based on chance), focused on the local, and made on small amounts of data." However, Tarr et al. (2006) believe that middle school students can recognize the importance of using larger samples in drawing inferences, if they are provided with carefully designed instructional materials.

Since 2003 members of our group have been concerned about how students in Grade 10 (aged 15) were drawing inferences from the comparison of box plots and the justifications that they were giving for their conclusions (Pfannkuch 2006). In exploratory research focused on how a teacher was reasoning from box plots, it was noted that the teacher did not make the distinction between samples and populations. Some of her statements were made about the sample, and some about the population from the sample. That is, she and her students were confused about what game they were playing: the Description Game (about the sample), or the Inference Game (about what is happening back in the population). Furthermore, when considering sampling variability she talked about it hypothetically rather than allowing students to experience sampling behavior (Pfannkuch, 2006). Pfannkuch believed that lack of appreciation of sampling variability was limiting the students' informal inferential reasoning. Therefore in the next two years she experimented with developing Grade 9 students' concepts of sampling variability using two web applications. These seemed to stimulate in students some awareness of variability both in categorical and numerical data, of the effect of sample size, and of a connection between sample and population. Dynamic visual imagery, tracking the history of the variability, and the discourse of the teacher were considered important in stimulating these concepts. But the problem of judging whether A tended to have bigger values than B still remained as the students and teacher continued to draw conclusions based on the difference between the medians without considering sampling variability. Discussions with teachers and assessment moderators have revealed that this type of reasoning is common, at least in New Zealand schools.

We believe the work of SRTL-5 and other researchers needs to be further developed by building students' awareness of sampling variability more purposefully. When students reason comparatively from the centers of sample distributions, they need to take into account sampling variability and the sample size. Simultaneously learning and understanding these three factors is very challenging, particularly for novices.

We now briefly describe some of the issues in our design of a conceptual pathway across the curriculum (ages 14 to 17) towards formal inferential reasoning, and our dynamic visualizations. For a fuller discussion on the issues and technical aspects of our approach see Wild et al. (to appear). Also what we present in this paper are *the final steps* in a multi-step learning process that carefully introduces the underpinning concepts (see Arnold & Pfannkuch, 2010).

A NEW VISUAL APPROACH TO SAMPLING VARIATION

Our goal is build the *concept* of sampling variability through a dynamic visual approach according to the following principles. The approach:

1. Encourages students not to take their eyes off their plots (see Figure 1).
2. Has connections to more formal statistical inferential methods.
3. Works from a minimal set of inferential ideas that can gradually be built on.
4. Uses principles of learning with dynamic visualizations such as color coding, prior knowledge, and pace control (e.g., Mayer, 2009).

5. Recognizes that students do not see what experts see and that prior knowledge such as hands-on learning experiences are essential for understanding abstract dynamic imagery.
6. Recognizes that the visual approach is developed in conjunction with a verbal approach that communicates the essence of what is being captured in the visuals.
7. Allows students to make a decision or call about whether one group tends to have bigger values than another group by providing a pathway of increasingly more complex heuristics that seek to deepen students' understanding of sampling variability (see Figure 2).
8. Recognizes the current technology constraints of the New Zealand learning environment where most secondary classrooms have access to only one computer and a data projector.

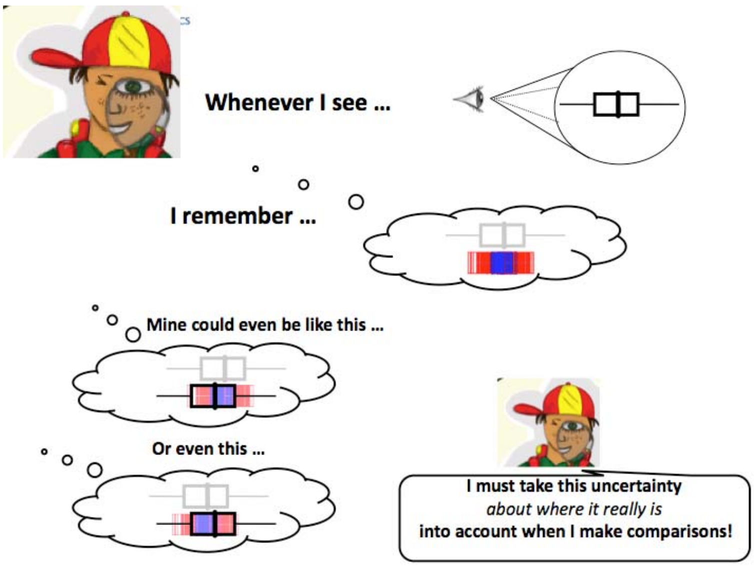

Figure 1. The desired habit of mind

Our aim is that whenever students see a box plot they will visualize a vibrating box plot (Figure 1), recall the words "what I see is not really the way it is back in the populations" and remember to take sampling variability into account before making a call. The desire is for this embedded visual imagery in their minds to be associated with concepts of sample, population, sampling variability, sample size effect, and with a heuristic that is appropriate for their age or developmental level for making a call (Figure 2). In order to embed these images and concepts in students' memories we have developed a range of dynamic visualizations that become increasingly more complex. We use a pdf format, which allows the user or teacher to control the pace of the visualizations in order to give students time to understand the changing representations and to relate to previous learning experiences. Some of the first dynamic visualizations that students experience directly relate to their hands-on experiences of drawing samples from two populations and each time plotting the resultant comparison. The dynamic visualizations show the resultant two box plots each time the samples are drawn and students' attention is focused on the direction of the medians. Other visualizations track the history of the variability of the medians and the quartiles by leaving footprints (Figure 1), demonstrate the sample size effect, show dot plots and box plots using a variety of sample sizes, feature the unseen world of the population distributions, which then fades with the reminder that unfortunately we do not get to see the populations but only one frame from the "movie" and so forth. (Some animations are at: www.censusatschool.org.nz/2009/informal-inference/WPRH/).

CONCLUSION

Our proposed visual approach to sampling variation enables students to make sound and simplified inferences about populations from samples. In this way they will be enculturated into statistical practice by learning how and why statisticians make decisions based on data and at the same time develop a conceptual basis for inference. Our hope is that when students investigate data within the spirit of EDA they will be able to make inferences based on their developing understandings of sampling variability. The approach we envisage to *making a call* is based on *theoretical considerations*. The question is whether the proposed conceptual pathway and dynamic visualizations for developing inferential reasoning will work in practice. Initial explorations with students and teachers in classrooms suggest the theory is promising.

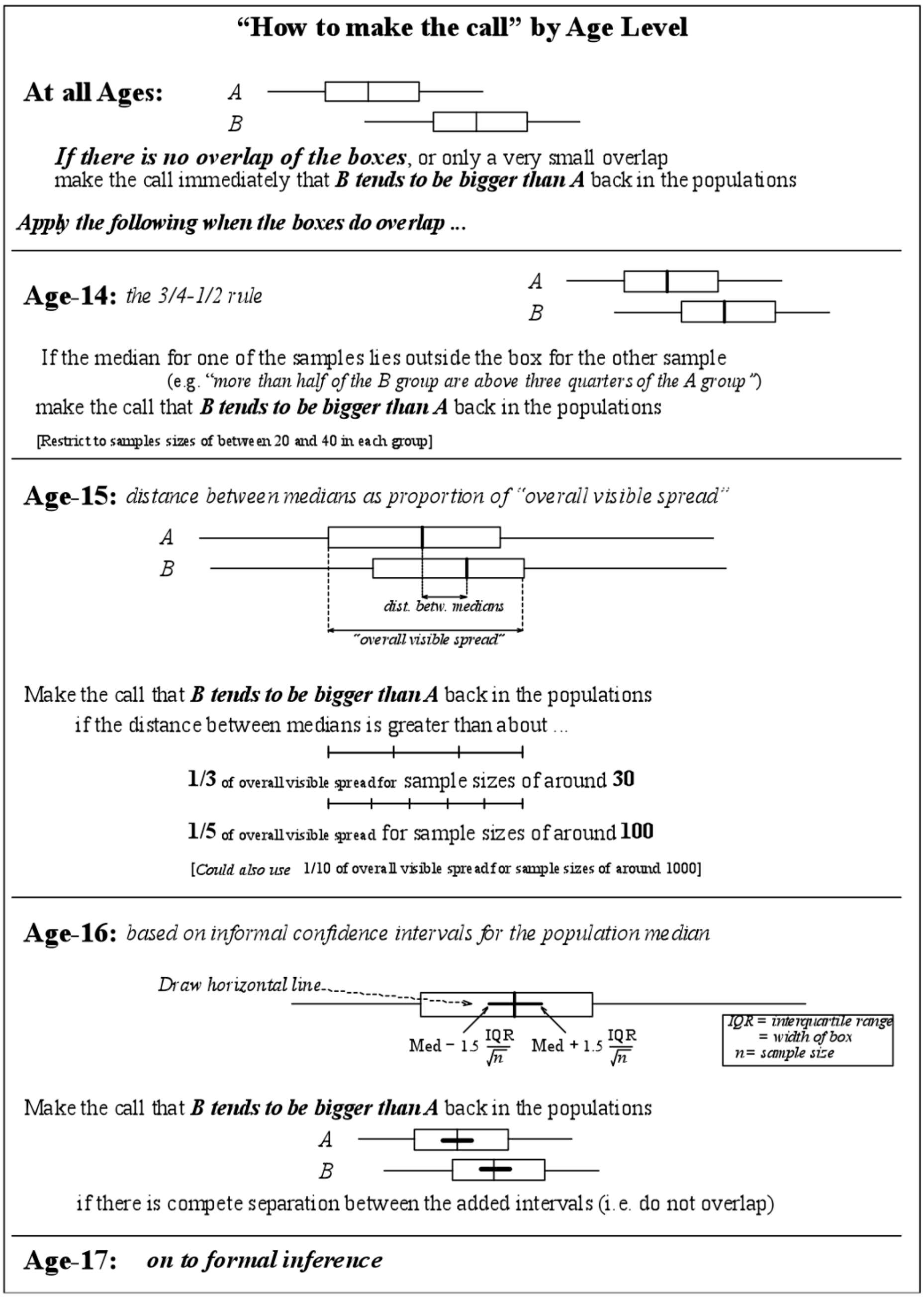

Figure 2. How to make a call at each age or developmental level

ACKNOWLEDGEMENT
This work is supported by a grant from the Teaching and Learning Research Initiative (www.tlri.org.nz).

REFERENCES

Arnold, P., & Pfannkuch, M. (2010). Enhancing students' inferential reasoning: From hands on to "movie snapshots". In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics, Ljubljana, Slovenia, July 2010*. Voorburg, The Netherlands: International Statistical Institute. Online: www.stat.auckland.ac.nz/~iase/publications.php

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, *3*(2), 64-83.

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education. Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Brazil, July 2006*. Voorburg, The Netherlands: International Statistical Institute. Online: www.stat.auckland.ac.nz/~iase/publications.php.

Ben-Zvi, D., Gil, E., & Apel, N. (2007). What is hidden beyond the data? Helping young students to reason and argue about some wider universe. In J. Ainley & D. Pratt (Eds.), *Proceedings of the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy: Reasoning about Statistical Inference: Innovative Ways of Connecting Chance and Data*. The University of Warwick, (pp. 1-26). 11–17 November 2007.

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-324). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, *18*(5), 382-393.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: a pre-k-12 curriculum framework*. Alexandria, VA: American Statistical Association.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education*, *2*(1). http://repositories.cdlib.org/uclastat/cts/tise/vol2/iss1/art1/

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

Mayer, R.E. (2009). *Multimedia Learning*. New York: Cambridge University Press.

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

Pratt, D., Johnston-Wilder, P., Ainley, J., & Mason, J. (2008). Local and global thinking in statistical inference. *Statistics Education Research Journal*, *7*(2), 107-129.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.), *People in mathematics education*. *Proceedings of the 20th Annual Conference of the Mathematics Education Research Group of Australasia, Rotorua, New Zealand, July, 1997.* (Vol.1, pp. 6–22). Sydney: MERGA.

SRTL-6. (2008). *SRTL-6 first announcement*. Retrieved April 2, 2008 from http://srtl.stat.auckland.ac.nz/srtl6/announcements

Tarr, J., Lee, H., & Rider, R. (2006). When data and chance collide: drawing inferences from empirical data. In G. Burill & P. Elliott (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth Yearbook* (pp. 139–149). Reston, VA: National Council of Teachers of Mathematics.

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

Watson, J. M. (2008). Exploring beginning inference with novice grade 7 students. *Statistics Education Research Journal*, *7*(2), 59–82.

Wild, C. J. (2006). The concept of distribution. *Statistics Education Research Journal*, *5*(2), 10-26.

Wild, C. J., Pfannkuch, M., Regan, M., & Horton, N. (to appear). Towards more accessible conceptions of statistical inference. *Journal of the Royal Statistical Society Series A*.

Zieffler, A., Garfield, J., delMas, R., & Gould, R. (2007). Studying the development of college students' informal reasoning about statistical inference. In J. Ainley & D. Pratt (Eds.), *Proceedings of the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy: Reasoning about Statistical Inference: Innovative Ways of Connecting Chance and Data*. The University of Warwick, (1–42). 11-17 November 2007.

Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, *7*, 40–58.