MAKING THE CONNECTIONS: FROM STATISTICS EDUCATION RESEARCH TO TEACHING ADVICE

Rosemary Callingham, Jane Watson and Dianne Siemon
University of Tasmania, Australia
RMIT University, Australia
Rosemary.Callingham@utas.edu.au

Although statistics has now been in the Australian curriculum for two decades, much of the advice to teachers focuses on computational statistics rather than developing reasoning about statistical situations. A recent project focused on developing a Learning Progression of Statistical Reasoning and then providing explicit guidance for teachers linked to the Learning Progression. The intended outcome is a package of assessment tasks that teachers can use to identify their students' current zones of statistical reasoning and suitable activities to provide the foundation for further developing students' statistical reasoning. This paper explores the processes used to move from the research-based validated Learning Progression to teaching advice.

INTRODUCTION

It will soon be 30 years since Statistics had an acknowledged place in the curricula of many countries, with the National Council of Teachers of Mathematics (NCTM) leading the way with its 1989 *Standards*. Until then school students mainly experienced statistics as finding the arithmetic mean (e.g., Denbow & Goedicke, 1959). From the 1990s, research about learning and teaching statistics with school students in the classroom has developed into a field in its own right. Much of this research has been summarized and reported to the academic community (e.g., Shaughnessy, 2007) but the influence on teachers in the classroom is less obvious. Curriculum documents outline a proposed progress in the levels of understanding students should acquire but fail to give advice to teachers on the difficulties students experience in achieving the prescribed goals and how to overcome them.

Early research at the school level often produced examples of potential outcomes for innovative ideas (e.g., Ben-Zvi & Arcavi, 2001). Despite the focus on classroom experiences, this body of work had limited impact on classroom teachers and did not offer suggestions for how to deal with all levels of cognitive development in the classroom. Some of the early work, such as the *Used Numbers* series of books (e.g., Russell & Corwin, 1990) suggested meaningful activities across the primary years but without specific links to cognition except for hypothesized grade level. Later work, such as that incorporating the Problem, Plan, Data, Analysis, Conclusion (PPDAC) cycle (Wild & Pfannkuch, 1999) enlivened statistics teaching and learning and was taken up in some places, notably New Zealand, but again did not provide teachers with a framework for monitoring their students' growth in statistical understanding other than the stages of the curriculum.

The advent of Learning Progressions

The notion of a learning progression is relatively recent (e.g., Clements & Sarama, 2006), although ideas about hierarchical cognitive development have a much longer history (e.g., Piaget & Inhelder, 1975). Clements and Sarama (2006) identified three elements appear common to all learning progressions or trajectories: Goals for learning; Appropriate tasks; Hypotheses about the progression of learning along the continuum (Baroody, Cibuliskis, Lai, & Li, 2006). Baroody et al. went on to warn that however well developed and validated, any specific learning progression should not be considered as the only route to development of understanding.

In terms of school level statistics, early work in the context of learning progressions occurred when Watson and Callingham (2003) and Callingham and Watson (2005) used Rasch analyses (Bond & Fox, 2015) to suggest a hierarchical structure for statistical literacy. This construct was described in six levels of understanding and was later shown to be stable over time by repeating the original analysis with a new data set some ten years after the original project (Callingham & Watson, 2017). Despite work with teachers (Callingham & Watson, 2011), these analyses were not used to make specific suggestions for teaching to raise levels of students' outcomes or for specific professional

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

development of the theoretical construct. The *Reframing Mathematical Futures Project* (RMF2) aimed to fill this gap. The focus of this paper is to exemplify the process used to link theory and practice.

LINKING LEARNING PROGRESSIONS TO TEACHING ADVICE

The RMF2 project intends to "build a sustainable, evidence-based, integrated learning and teaching resource to support the development of mathematical reasoning in Years 7 to 10" (Royal Melbourne Institute of Technology (RMIT), 2017). *Reasoning* appears in the *Australian Curriculum: Mathematics* (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015) as one of four mathematical proficiencies. In the RMF2 project statistical reasoning was one of three focus strands, along with algebraic and geometric reasoning. The focus on reasoning meant that teachers and researchers had to go beyond the content described in the curriculum to the underlying structure and development of thinking in each of the strands.

The approach taken in the RMF2 project to linking research to teaching uses multiple tactics. As well as the development of learning progressions that is at the heart of the project, teachers have been provided with teaching advice and professional development delivered at meetings and online. The final outcomes of the project are not yet available. The learning progression for statistical reasoning makes use of quantitative data from Rasch analysis (Bond & Fox, 2015) based on carefully designed items to develop a learning progression. Items were developed following a comprehensive review of the literature to ensure that they covered key conceptual aspects of statistics (e.g., Watson, 2006) and content appropriate to the middle school years that were the target of the project (e.g., ACARA, 2015; NCTM, 2000). Questions often had several parts, termed items, that generally allowed for answers to be coded using a rubric that recognized several levels of quality responses. For example, SHAT8 is a single item question addressing chance (see Figure 1). The coding rubric applied allows for incorrect and correct reasoning to be accepted with gradations of meaning reflecting an increasing quality and complexity of response.

<table>
<tr><th>Question</th><th colspan="2">Coding rubric</th></tr>
<tr><td rowspan="6">A mathematics class has 13 boys and 16 girls in it. Each pupil's name is written on a piece of paper. All the names are put in a hat. The teacher picks out one name without looking. Tick the box to show which outcome is more likely

☐ the name is a boy or
☐ the name is a girl or
☐ the name could be a boy or a girl

Please explain your answer using as much mathematics as you can.</td><td>Score code</td><td>Rubric description</td></tr>
<tr><td>0</td><td>No response/irrelevant response</td></tr>
<tr><td>1</td><td>Incorrect, little/no reasoning (e.g., it's just luck)</td></tr>
<tr><td>2</td><td>Incorrect (e.g., name is a boy or girl) but reasoning that recognises variation in some way (e.g., depends on mix, same chance, could be anything)</td></tr>
<tr><td>3</td><td>Correct (name is a girl) with either no explanation or explanation does not reference total (e.g., 16 is bigger than 13)</td></tr>
<tr><td>4</td><td>Correct, fraction included in explanation (e.g., 16/29 chance)</td></tr>
</table>

Figure 1. Question and coding for SHAT8.

These items were organized using a three-part framework: variation in expectation that included central tendency and chance; variation in distribution that included graphs and tables; and variation in inference that addressed drawing conclusions and making justifications based on data. These categories were deliberately broad, acknowledging variation as underpinning statistics, but ensured that the assessments covered key ideas and concepts beyond mathematical statistics. By the conclusion of the project, the items developed will be organized into assessment tasks that can be used as pre- and post-assessments with scoring rubrics and score bands that teachers can use to identify an approximate zone for every student in their class, and hence target their teaching more effectively.

The initial task was to develop a learning progression for statistical reasoning. Using questions from previous projects (Callingham & Watson, 2005, 2017; Watson and Callingham, 2003) test forms were developed and given to nearly 1500 students from Year 7 to Year 10. Using the same methodology as in previous projects, the responses were Rasch analyzed and a continuum developed (Callingham & Watson, 2005). The item clusters along the continuum were examined for evidence of statistical reasoning using a strategy termed "segmenting the variable" (Wilson & Sloane, 2000), and an eight-zone hierarchy was developed (Watson & Callingham, 2017). The choice of eight zones was made to link the statistical reasoning hierarchy more closely to the other project domains. Eight zones were also considered useful for teachers in that they provided an overview of where students had come from as well as where they would most likely develop next, although in any one class most students would probably be located across about three zones. The advantage to teachers of describing zones within which students could be located means that rather than atomized skills, teachers can consider students' statistical reasoning more holistically, both to anticipate possible cognitive demands of a new task and to predict students' likely responses. This approach was taken to give teachers a "road map" within which they could address the curriculum, which has a focus on statistical skills. For example, the Year 7 curriculum statement that would be addressed by SHAT8 is *Assign probabilities to the outcomes of events and determine probabilities for events* (ACARA, 2015) but this statement does not give teachers ways of recognizing or addressing the range of understanding that might be present in their classes.

Characteristics of the zones

The first step in providing teaching advice was to identify characteristic behaviors within each zone by considering carefully the knowledge, skills, and understanding needed to gain the requisite score on each item located within that zone. From this analysis, a broad descriptor of behaviors was developed. These descriptors are summarized in Table 1.

Table 1. Characteristics of Statistical Reasoning Zones

Zone	Characteristic behaviors	Examples
Zone 1	Idiosyncratic response or single procedural focus	Uncertainty expressed as 50%; Reads single value on graph; Ignores context
Zone 2	Considers aggregated information but without recognizing value	Anything can happen; Describes isolated features of a graph; One characteristic of a sample
Zone 3	Emerging statistical appreciation but without explanation	Claims for average without justification; Elaborated physical description of graphs; Choose "all" for sample
Zone 4	Recognizes influence of variation but interprets inappropriately	Rejects "luck"; suggests unlikely; Does not distinguish scale in graph reading; Recognizes sample but not its bias
Zone 5	Straightforward explanation and simple numerical justification	Orders chance phrases correctly; Appropriate attention to graph details; Partial recognition of sample requirements
Zone 6	Informal appreciation of uncertainty and variation in chance	Recognizes outlier; Recognizes correct variation in graphs; Suggests random sampling
Zone 7	Makes inferences across ideas using proportional reasoning	Creates appropriate probability distribution; Creates hypothesis based on data; Criticizes sample size and bias;
Zone 8	Integrates proportional, statistical, and contextual reasoning	Correct association in 2-way tables; Conclusion with both positives and negatives; Includes human/psychological component

The next step in providing teaching advice was to consider the question "If students located in this zone are doing …, what is needed to help them move to the next higher zone?" This advice was framed within two aspects: *Consolidate and Establish*, and *Introduce and Develop*. The first, *Consolidate and Establish,* attempted to identify knowledge, skills, and understanding that were still borderline and to suggest activities that might help students develop deeper understanding. The second, *Introduce and Develop,* was forward looking, and aimed to set up activities and ideas that would be needed in the next zone of development. Figure 2 shows an example of this teaching advice for Zone 4. The text in brackets (e.g., SHWKB.2) indicates an item—in this case an item that used a curvilinear graph showing time spend on homework against test score as a stimulus—with the score code (2) allocated to the response. Most items can be found in the appendix to Callingham and Watson (2005). The italicized text in the Teaching Implications column gives the titles of activities available to teachers via a drop box or from indicated websites.

ZONE 4 BEHAVIORS	TEACHING IMPLICATIONS
Compares data in two graphs but focuses on single elements only (SHSE2.1). Can associate two variables with single value (SHWKB.2) and provides descriptive explanations (SHWKC.1). Recognises variability and expectation in more complex random situations (e.g. SD12B.2) but explanation refers to uncertainty in general terms and is not quantified or is based on strict probability (expectation) (e.g., SD12B.3). May not recognize the importance of equal likelihood (e.g., STATS.2). Recognizes relative order in language of uncertainty (WORD.1) but does not appreciate some subtleties. Reasons quantitatively in familiar situations involving related comparison and in the context of uncertainty (e.g., SBOX9). Relies on additive thinking in situations involving measures of central tendency (SAMEA.2), and is unlikely to question the quality of data (SAOUT1). Critiques sampling approaches using single aspects only (i.e., size or method) in an evaluative situation (e.g., SMV11.1; SMV12.1; SMV13.1). Falls back on personal beliefs in more complex situations when asked for an explanation (e.g., SCON3.1).	**Consolidate and Establish**: Compare whole distributions, e.g., *Balancing Act Comparing Groups*. *Manufacturing Licorice* activity. Extend the language of uncertainty, e.g., Read *Pigs Might Fly* by Emily Rodda, discuss the language, ask students to write a short story using words related to chance and uncertainty. Research the meaning and origins of common chance phrases such as "Once in a Blue Moon," "Pigs Might Fly," etc. Discuss which sayings have a quantitative basis and which are just about luck. *Friday 13th* (https://nrich.maths.org/610) Develop the relationship between uncertainty and statistical information, e.g., *The Birth Month Problem*. *Mystery Bag* investigation. *Fairness of Dice*. *Two Coins.* Provide opportunities for quantitative reasoning in statistical and probabilistic contexts, e.g., *Mystery Bag*. Telling stories from graphs. *Balancing Act*. *Fairness of Dice*. **Introduce and Develop:** Make hypotheses that can be tested by collecting data, e.g., *Balancing Act*. Explore fairness through different sampling methods, e.g., *Winning the Lottery* (http://nrich.maths.org/7244. Compare outcomes from dice and spinners. Fair games, e.g., *Odds or Sixes* (http://nrich.maths.org/2859) Explore the role of context with a focus on language, questioning, and the meaning of numbers relative to context, e.g., "The average family has 2.3 children. What does this mean?"

Figure 2. Example of teaching advice for a particular zone.

Professional learning

In addition to the advice provided directly to teachers, professional learning was delivered to the project teachers, both face-to-face and through web-based interactive sessions. The online interactive professional learning sessions were also recorded and made available to teachers for later use. The aim was for teachers to use these recorded sessions, either individually or preferably in a group, as a basis for discussion about how their own students might respond. At the face-to-face professional learning sessions teachers also used similar rich tasks and discussed how these could be implemented in classrooms, as well as possible responses from their students. They also provided feedback on trial materials and tasks. One example of using the results of the research is the *Lung Disease* problem (Batanero, Estepa, Godino, & Green, 1996) shown in Figure 3, which had been used in earlier research with both students and teachers (Watson & Callingham, 2014). This example was used in a face-to-face workshop with 45 teachers and triggered a great deal of discussion among the groups of teachers around the room, with summary suggestions reported to the entire workshop. The medium was a series of cartoons with students giving answers typical of different zones, and the teacher responding.

Lung Disease Task

The following information is from a survey about smoking and lung disease among 250 people.

	Lung disease	No lung disease	Total
Smoking	90	60	150
No smoking	60	40	100
Total	150	100	250

Using this information, do you think that for this sample of people lung disease depended on smoking? Explain your answer.

Figure 3. *Lung disease* task (from Batanero et al., 1996).

In Figure 4, Lucas is giving a Zone 5 response by comparing cells in one column. It is, however, inadequate because it does not consider enough of the evidence in the table. Mia shows a Zone 6 response, recognizing that there is more information in the Row Totals that needs to be taken into account. She does not go on, however, to show the proportional reasoning of a Zone 7 response that would draw the appropriate conclusion of independence of the two conditions. At this point the teacher refers back to the issue of sample size raised earlier by another student and leads other members of the class to explain several Zone 7 responses using percentages, fractions, and ratios. The complete sequence of exchanges in the cartoon classroom explicitly addresses the difficulties that students have moving to using proportional reasoning to compare groups (Watson & Callingham, 2014).

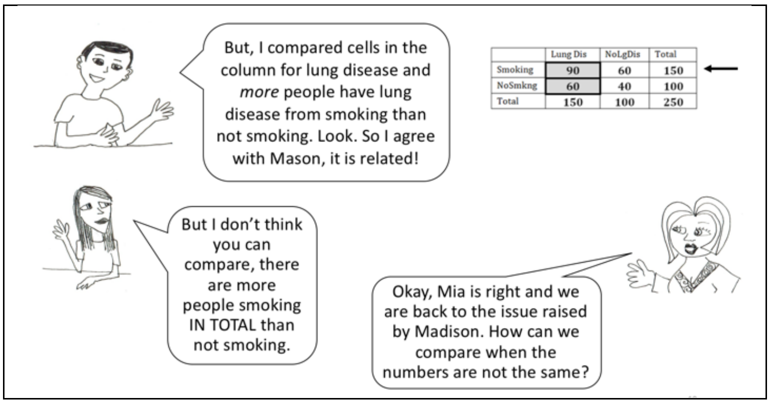

Figure 4. Part of the Long Disease cartoon showing two different responses.

Packages of materials including suggested tasks, and assessment activities based on the items developed and piloted in the early stages of the project were sent to project schools. Teachers chose a focus area (some teachers chose two) and undertook to do a pre- and post-assessment with at least one class to determine how successful the approach was for improving students' outcomes. The final data from this component of the project is not yet available because some teachers chose to leave doing this until 2018. Feedback from teachers will be incorporated into the final products.

DISCUSSION

There is frequent comment that educational research does not usefully inform teachers. The RMF2 project set out explicitly both to involve teachers in the research and to provide useful material for teachers. The starting point was the historical research base that provided the theory. Working with teachers who have to deliver the curriculum provided links to practice, through feedback on trial tasks, suggestions for activities, and discussion at professional development sessions delivered in different modes.

The hypothetical learning trajectory initially developed was refined following initial task trialing, including modifying coding rubrics so that they were more teacher-friendly, allowing the practices and understandings of real teachers to be included. The content of the tasks and items developed addressed the curriculum, but the Statistical Reasoning Learning Progression went beyond curriculum statements to give teachers deeper insights into their students' statistical reasoning.

Linked to targeted teaching advice, the progression should help teachers focus their teaching to allow all students to progress in understanding, providing a practical outcome from a sound research base.

REFERENCES

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2015). *Australian Curriculum: Mathematics, Version 8.1, 15 December 2015.* Sydney, NSW: Author.

Baroody, A., Cibulskis, M., Lai, M-l., & Li, X. (2006). Comments on the use of learning trajectories in curriculum development and research, *Mathematical Thinking and Learning, 6(*2), 227-260.

Batanero, C., Estepa, A., Godino, J.D., & Green, D.R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27,* 151-169.

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

Bond, T., & Fox, C. (2015), *Applying the Rasch Model: Fundamental measurement in the human sciences* (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum

Callingham, R., & Watson, J. M. (2005). Measuring statistical literacy. *Journal of Applied Measurement, 6*(1), 29, 19-47.

Callingham, R., & Watson, J. (2011). Measuring levels of statistical pedagogical content knowledge. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study* (pp. 283-293). Dordrecht, The Netherlands: Springer.

Callingham, R., & Watson, J. (2017). Appropriate goals for statistical literacy at school. *Statistics Education Research Journal 16*(1), 181-201.

Clements, D. & Sarama, J. (2004) Learning trajectories in mathematics education, *Mathematical Thinking and Learning, 6*(2), 81-89

Denbow, C.H., & Goedicke, V. (1959). *Foundations of mathematics*. New York: Harper & Row.

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

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Piaget, J., & Inhelder, B. (1975) *The origin of the idea of chance in children*. London: Routledge & Kegan Paul. (Original work published 1951).

Royal Melbourne Institute of Technology (RMIT) (2017). *Reframing Mathematical Futures II.* Retrieved from www.rmit.edu.au/about/our-education/academic-schools/education/research/projects/reframing-mathematical-futures-ii

Russell, S.J., & Corwin, R.B. (1990). *Sorting: Groups and graphs: A unit of study for grades 2-3*. [Used numbers: Real data in the classroom] Palo Alto, CA: Dale Seymour.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957–1009). Charlotte, NC: Information Age Pub.

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

Watson, J., & Callingham, R., (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3–46.

Watson, J., & Callingham, R., (2014). Two-way tables: Issues at the heart of statistics and probability for students and teachers. *Mathematical Thinking and Learning, 16*(4), 254-284.

Watson, J.M., & Callingham, R., (2017). Developing learning progressions to support mathematical reasoning in the middle years – Statistical reasoning. In A. Downton, S. Livy, & J. Hall (Eds.), *40 years on: We are still learning!* (Proceedings of the 40th annual conference of the Mathematics Education Research Group of Australasia, pp. 663-666). Melbourne: MERGA.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-248.

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208.