Whither Statistics Education Research?

Jane Watson
University of Tasmania
<Jane.Watson@utas.edu.au>

This year marks the 25th anniversary of the publication of a *National Statement on Mathematics for Australian Schools*, which was the first curriculum statement this country had including "Chance and Data" as a significant component. It is hence an opportune time to survey the history of the related statistics education research, consider where we are at the moment, and speculate about where we should be going. Some of the issues to be considered along the way include the relationship of research to the curriculum and its implementation (or not), the rise of a research culture, the juxtaposition of statistical literacy and statistical reasoning, and the importance of context for learning about statistics. The continued interest in affective variables is considered, as are the recent trends related to the pedagogical content knowledge needs of teachers and the influence of advances in technology. Finally some views are canvassed on the new initiatives in STEM education and Mathematics by Inquiry, as well as the possible impact of statistics education researchers in the field.

Introduction

Anecdote

Recently I submitted a manuscript based on asking Grade 5 students to make predictions about a population from random samples. The fact that most students did not use the mean but other characteristics of the data to make their decisions was criticised by a reviewer, who subsequently suggested I should concentrate on teaching the mean and leave the sophisticated ideas of prediction until much later when confidence intervals were available. The referee was a university statistics lecturer. This view is totally opposite to the results of research in statistics education of the last 20 years. We want students to learn about the practice of statistical inquiry as a particular way of problem solving, and develop ever-more sophisticated tools along the way. There are still significant messages for school level statistics education researchers to pass on to their tertiary statistics colleagues and to the mathematics education community.

Premises

Statistical literacy is essential for all citizens and must be addressed during the school years as a *General Capability*.

The most meaningful way to approach statistical literacy and become statistically literate is to carry out the practice of statistics.

The practice of statistics at the school level involves informal inference and modelling.

Teachers need the pedagogical content knowledge to engage students in learning the practice of statistics.

The *Australian Curriculum: Mathematics* needs to present statistical concepts in a developmentally sound manner and include a framework for the practice of statistics.

Software should assist in developing concepts as well as computing algorithms.

2016. In White, B., Chinnappan, M. & Trenholm, S. (Eds.). Opening up mathematics education research (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia), pp. 33–58. Adelaide: MERGA.

Summary

In recognition of the 25th anniversary of a *National Statement on Mathematics for Australian Schools* (Australian Education Council [AEC], 1991) we consider the parallel history of statistics education research. After defining some of the terminology, we consider issues associated with the evolving curriculum, the rise of a research culture, the relationship of statistical literacy and statistical reasoning, the necessity for context, the continuing interest in affective variables, the education of teachers, the growing availability of technology, the emergence of STEM education, issues of education for sustainability, and the new initiative for Mathematics by Inquiry.

Terminology

Statistical Literacy

The phrase Statistical Literacy dates at least from 1992 when Wallman (1993) used it in her Presidential address to the American Statistical Association.

> 'Statistical literacy' is the ability to understand and critically evaluate statistical results that permeate our daily lives — coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions. (p. 1)

In 2002 Gal extended the description specifically for adults with the following two components:

> (a) people's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant,
>
> (b) their ability to discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions. (pp. 2–3)

At the school level Watson (2006) suggested the following:

> Statistical literacy is the meeting point of the data and chance curriculum and the everyday world, where encounters involve unrehearsed contexts and spontaneous decision making based on the ability to apply statistical tools, general contextual knowledge, and critical literacy skills. (p. 11)

The Practice of Statistics

The phrase Practice of Statistics goes back at least as far as the title of the famous introductory tertiary textbook by Moore and McCabe (1989). In their introduction they said it was their "intent to introduce readers to statistics as it is used in practice. Statistics in practice is concerned with gaining understanding from data; it is focused on problem-solving" (p. xi). The phrase used by the American Statistical Association in 2007 in its *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* (Franklin et al., 2007) was "statistical problem solving". It was defined as an investigative process involving four components:

- Formulate Questions - Anticipating Variability;
- Collect Data - Acknowledging Variability;
- Analyze Data - Accounting of Variability;
- Interpret Results - Allowing for Variability.

Other more detailed descriptions are found in Wild and Pfannkuch (1999) based on the analysis of their university colleagues' work and in Makar, Bakker, and Ben-Zvi (2011) translated to informal inference for the school level. Most recently the National Council of Teachers of Mathematics (NCTM) used the title *Statistics into Practice* (Crites & St. Laurent, 2015) in its series of books on developing essential understanding for the classroom.

The Curriculum

Statistics education was starting to gain attention in the 1980s (e.g., Shulte, 1981) but we can thank the NCTM in the United States for their *Curriculum and Evaluation Standards for School Mathematics* in 1989, which included Statistics and Probability across all grades. The *Standards* are still quotable. Standard 11 for Grades K-4 gives a precise description of the components of "practicing statistics" in the experiences students should have:

- collect, organize, and describe data;
- construct, read, and interpret displays of data;
- formulate and solve problems that involve collecting and analyzing data. (p. 54)

The *Common Core State Standards for Mathematics* (Common Core State Standards Initiative [CCSSI], 2010) is the most recent effort to create a national curriculum in the United States. It ignored these points for young children in not addressing Statistics and Probability until Grade 6, but it did apparently heed the rationale for 1989 NCTM Standard 10 for Grades 9-12:

> Statistical data, summaries, and inferences appear more frequently in the work and everyday lives of people than any other form of mathematical analysis. It is therefore essential that all high school graduates acquire, at the appropriate level, the capabilities identified in this standard. This expectation will require that statistics be given a more prominent position in the high school curriculum. (p. 170)

Although the *Common Core* Standards for Content from Grades 6 to 12 are quite dense for Statistics and Probability, they do cover most of the content from 1989 NCTM *Standards*.

Australia followed the NCTM in 1991 with *A National Statement on Mathematics for Australian Schools* (Australian Education Council [AEC], 1991), again including a serious treatment of Chance and Data across four bands of schooling, with similar goals for Data Handling and Statistical Inference to the NCTM *Standards* (1989). The national curriculum movement died before "National Statements" could be published in other subject areas, but *Mathematics – a curriculum profile for Australian schools* and *Mathematics – work samples* (AEC, 1994a, b) were produced, which provided some starting points for research with students.

New Zealand also had a *Mathematics* curriculum including a section on Statistics from 1992 (Ministry of Education). For each of "Statistical Investigations", "Interpreting Statistical Reports", and "Exploring Probability", there were Achievement objectives, Suggested learning experiences, Sample assessment activities, and Sample development band activities. The 2009 revision of the curriculum (Ministry of Education) was called *Mathematics and Statistics* acknowledging the claim of the NCTM (1989) of the importance of Statistics. The subsection on Statistical Investigations was based on the work of Wild and Pfannkuch (1999) and consistent with *GAISE* (Franklin et al., 2007). As well Interpreting Statistical Reports was renamed "Statistical Literacy", recognising the connections with Statistical Investigations but also differences in purpose. “Probability” remained a third subsection.

The current *Australian Curriculum: Mathematics* (e.g., Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015a) began evolving in 2010 and could have learned from the New Zealand experience in terms of a meaningful exposure to the practice of statistics. The Foundation to Year 10A document contains Statistics and Probability with two subsections: “Chance” and “Data representation and interpretation”. The curriculum suggests many tools for data analysis, although one could argue that the writers were not aware of the research pointing to the ages at which students would *be able* to use them; also the tools are not accompanied by a structured indication of *how* to use them. Words such as “interpret”, “explore”, and “investigate” are not accompanied by what they mean in a statistical investigation. To expect the Proficiencies as provided in the curriculum to fill the gap is disappointing as they are also vague, and for teachers with little background, virtually useless. At Grade 10, for example, *Understanding* includes “determining probabilities of two-and three-step experiments”, *Fluency* includes “using calculations to determine the shape of data sets”, *Problem solving* includes “investigating independent events”, and *Reasoning* includes “interpreting and comparing data-sets”. The curriculum could very easily have contained the overall framework to carry out a statistical investigation as found for example in *GAISE* (Franklin et al., 2007). In Australia, ironically, a suitable framework is found in the *Australian Curriculum: Science* (ACARA, 2015b) under “Science inquiry skills”. Further, the elements required to become statistically literate are found, not in the *Mathematics* curriculum itself, but in the *General Capabilities* (ACARA, 2013a) under Numeracy. A comparison of the links between the Science and Statistics curricula across New Zealand, the United States and Australia could point the way to future revision of the two curricula in all three countries (Watson, 2016). The current Descriptors for Statistics and Probability in the *Australian Curriculum: Mathematics* do not lay the foundation for meaningful classroom research.

Question: How can statistics education researchers work together to convince ACARA that the curriculum requires revision to capture the essence of the practice of statistics and the results of research?

A Culture of Research

Since 1999 there has been almost continuous discussion about the nature of statistical reasoning, thinking, and literacy (SRTL) with an SRTL forum held that year and every two years since to continue the discussion and advance research (<srtl.fos.auckland.ac.nz>). After the first forum, which considered the differences among the three terms, the next forum focussed on “reasoning” generally. From then, the focus turned to specific types of statistical reasoning: about variation, about distribution, and about statistical inference. Issues related to statistical inference and what it means before students reach formal theory led to development of the idea of “informal inference” (e.g., Makar, Bakker, & Ben-Zvi, 2011; Makar & Rubin, 2009) and the next four forums considered the role of context and evidence in informal inferential reasoning, the new approaches to samples and sampling, the implications for reasoning about uncertainty, and most recently (2015) the place of models and modelling within informal statistical inference.

The SRTL forums have been the catalyst for many significant publications, including books (Ben-Zvi & Garfield, 2004; Zieffler & Fry, 2015) and special issues of the *Statistical*

Education Research Journal (*SERJ*) on reasoning and variability (Volume 3, Number 2), on reasoning and distribution (Volume 5, Number 2) and on informal inference (Volume 7, Number 2), *Mathematical Thinking and Learning* (*MTL*) on context and informal inference (Volume 13, Numbers 1-2), and *Educational Studies in Mathematics* (*ESM*) on sampling (Volume 88, Issue 3). Although other authors have made significant contributions in the field outside of the SRTL forums, many have contributed to at least one forum, and looking at the topics covered suggests that the SRTL forums have addressed the significant and emerging learning issues in the field of statistics education research.

Twenty-five years ago, a review of available research to contribute to the writing of the *National Statement* (AEC, 1991) *could* have included reference to virtually all known relevant studies in the field (e.g., Green, 1983; Goodchild, 1988; Mevarech, 1983; Strauss & Bichler, 1988; Tversky & Kahneman, 1983). Since then the statistics education field has expanded greatly. Journals dedicated to statistics education and research have arisen: The *Journal of Statistics Education* (*JSE*) from 1993 and *SERJ* from 2002, being the most notable. *Teaching Statistics*, sponsored by The Royal Statistical Society Trust since 1979, although focussed on teaching, has reflected the research and promoted statistics education around the world. Before 2000, there were three edited books in statistics education (Gal & Garfield, 1997; Lajoie, 1998; Shulte, 1981) but since 2004, there have been nine besides the ones based on SRTL forums (Batanero, Burrill, & Reading, 2011; Ben-Zvi & Makar, 2016; Bidgood, Hunt, & Jolliffe, 2010; Burrill, 2006; Chernoff & Sriraman, 2014; Jones, 2005; Lovett & Shah, 2007; MacGillivray, Martin, & Phillips, 2014; Wassong, Frischemeier, Fischer, Hochmuth, & Bender, 2014) and a pending *International Handbook of Research in Statistics Education* (Ben-Zvi, Garfield, & Makar, 2016). Research studies have appeared individually in the top general mathematics education research journals, such as *MTL*, *ESM*, *Journal for Research in Mathematics Education* (*JRME*), *Mathematics Education Research Journal* (*MERJ*) and *Journal of Mathematical Behavior* (*JMB*), as well as in a special issue of *ZDM* (Volume 44, Number 7) on probability and reasoning about risk. Altogether, with these books, journal articles, and the many refereed research papers presented at the International Conferences on Teaching Statistics (ICOTS), Psychology of Mathematics Education conferences (PME), and other international and national conferences (such as MERGA), the number becomes amazing. It seems safe to claim that the literature available in statistics education generally, and in statistics education research in particular, has grown exponentially in recent years.

Question: What is left to research and how far can statistics education researchers push the boundaries of statistical inference at the school level?

Statistical Literacy and Statistical Reasoning

Judging by the recent titles for the SRTL gatherings, one is entitled to infer that "reasoning" is the key descriptor in the field of statistics education research. Gal (2002, 2005), however, has continued to argue for the importance of statistical literacy and probability literacy. Not all students will proceed to study formal statistics but all students need to leave school with the statistical literacy understanding to be able to analyse critically claims that are made in society, by the media, by politicians, by sporting commentators, etc. Being statistically literate is likely to involve some aspects of statistical reasoning but statistical reasoning is much broader, encompassing the ability to carry out the "practice of statistics". The language of statistical reasoning that has evolved from the

SRTL forums focuses on informal statistical inference (ISI) (Makar, Bakker, & Ben-Zvi, 2011). ISI is a generalisation beyond the collected data based on the evidence the data provide, acknowledging a degree of uncertainty in the generalisation reached (Makar & Rubin, 2009). Makar et al. then expand on five aspects of informal inferential reasoning (IIR) that underpin ISI: statistical knowledge (e.g., concepts, aggregate thinking), contextual knowledge, norms and habits of mind (e.g., collaboration, inquiry, critical, exploratory), inquiry drivers (e.g., expectation, conflict, explanation), and design elements (e.g., task, tools, scaffolds). This is a very comprehensive package of traits, not dissimilar to the earlier four-dimensional framework of Wild and Pfannkuch (1999), which includes an investigative cycle, an interrogative cycle, types of thinking, and dispositions. Best known is the investigative cycle, encompassing *p*roblem definition, *p*lanning for data, *d*ata collection, *a*nalysis of data, and *c*oncluding with an interpretation of results (PPDAC). Learning to carry out the practice of statistics is taken to encompass the fundamental ideas of these frameworks with various emphases over the school years.

Without the same complexity of language, the *GAISE* report (Franklin et al., 2007) sees the "practice of statistics" as "statistical problem solving", carrying out complete statistical investigations, and having an inquiry approach to learning statistics in context. Being statistically literate does not require the active involvement in investigation that the practice of statistics does. It does, however, require the critical thinking skills to make judgements about statistical claims made in actual social or scientific contexts. To apply the critical thinking skills, one needs the knowledge of the required statistical terminology and the ability to apply it in the context of the claim made (Watson, 2006). This three-step process (terminology, terminology in context, critical thinking) is tested daily when reading media reports presenting claims based on limited evidence, which students at all levels need to question (Watson, 1997; 1998b).

At the tertiary level the development of statistical reasoning has been linked to research and outlined for practice by Garfield and Ben-Zvi (2008) for introductory statistics courses. Despite their efforts and those of others, there continues to be concern about these courses fulfilling the statistical literacy needs of tertiary students. This has led to the development of specific courses for the latter purpose (e.g., Budgett & Pfannkuch, 2010; Finch & Gordon, 2014). Besides the concern about statistical literacy at this level, there is also the debate about whether students need formal inference or if simulation techniques such as bootstrapping or resampling (Cobb, 2007) will more easily introduce a meaningful understanding of the practice of statistics (delMas, Garfield, & Zieffler, 2014). *JSE*, as a mainly tertiary journal, includes papers on this topic (e.g., Watkins, Bargagliotti, & Franklin, 2014; Wood, 2005) and on issues associated with tertiary statistical literacy (e.g., Rumsey, 2002). Ziegler (2014) links these topics by suggesting tasks for measuring statistical literacy associated with randomisation and simulation be added to assessment instruments.

Although the statistically literate know that there is a *p*-value associated with a formal statistical decision, *p*-values are not presented to most school students and are not interpretable for many senior school, and even tertiary, students (Reaburn, 2014). For most primary and middle school students (at least) the mathematics they experience in the classroom is about certainty, from learning tables and solving equations to encountering proofs in algebra and geometry. Research has shown, however, that in many cases when students begin collecting data to make a decision, they want to "prove" the conjecture (e.g., Chick & Watson, 2002). One of the key ideas to be addressed across the school years is hence the “uncertainty” in decision-making in statistics. This is stressed in *GAISE* where the distinction is made between a mathematics question that anticipates a deterministic answer and a statistical question that anticipates an answer based on data that vary (Franklin et al., 2007, p. 11). Research is now addressing this topic (e.g., Ben-Zvi, Aridor, Makar, & Bakker, 2012).

How does modelling fit with statistical reasoning and the practice of statistics? Although SRTL forums first considered modelling in 2015, research from a data modelling perspective was first reported 20 years ago by Lehrer and Romberg (1996), who were interested in fifth grade students’ “construction of data” as a preamble to description and inference. Doerr and English (2003) considered middle school students’ development of reusable models for selecting, ranking, and modelling data based on multiple cycles of interpretation and re-interpretation of representations. Pushing the boundaries, English (2010, 2012) took data modelling to the first grade with two activities focussing on looking after the environment. The themes of this research in choosing attributes and ways to represent data from them are also reflected at the high school level with more complex activities of modelling data (Konold, Finzer, Kreetong, & Gaston, 2014) and visualising probability models (Budgett, Pfannkuch, & Franklin, in press).

The use of the term modelling varies across statistics education. In relation to putting statistics into practice, the NCTM (Crites & St. Laurent, 2015) builds statistical modelling from an assumption of understanding of mathematical modelling. There are three essential understandings:

a. Mathematical models describe structure.

b. Statistical models extend mathematical models by describing variability around the structure.

c. Statistical models are evaluated by how well they describe data and whether they are useful. (p.11)

This description does not appear to consider aspects of decision-making about what data to collect and how to put data in a manageable form, as suggested by English (2010) and Konold et al. (2014). The link of modelling to the practice of statistics may be better expressed the following way.

> Data modelling involves creating a description of a situation or system using statistical concepts and language. The description – model – is then used in making a decision about a meaningful question involving data.

Question: Is it necessary to undertake and understand the practice of statistics and data modelling in order to develop the critical skills necessary to become statistically literate? Would a really good textbook be sufficient? Designing a study to address this question, would surely be a challenging task.

Context

The recognition of the importance of context for statistics education is not a new discovery because as early as 1975 Rao was saying, “there is no statistics without context” (p. 151). Unfortunately this was forgotten for many years and textbooks gave students sets of numbers in order to calculate averages or plot histograms, with little if any meaning. Some would say that the *Australian Curriculum: Mathematics* (ACARA, 2015a) does little to mitigate this trend, reinforced by the type of items developed for NAPLAN (National Assessment Program Literacy and Numeracy, ACARA, 2013b) testing. Efforts to provide international data of relevance to social science study were laudable (e.g., Finlay & Lowe, 1993), as were the authentic contexts for Australian students in the *Chance and Data Investigations* of Lovitt and Lowe (1993). This has been continued in Australia with classroom suggestions, often associated with specific software (e.g., Day, 2013; Prodromou, 2015; Watson et al., 2011). Formal research studies now focus on context in various ways. On one hand, Pfannkuch (2011) considered three learning experience contexts: prior knowledge (historical), classroom interactions (social), and the motivation of the story in the data (task). On the other hand, Langrall, Nisbet, Mooney, and Jansem (2011) contrasted results for three countries using contexts relevant to 10-13-year-olds in the countries. The students made comparisons of data sets in the contexts based on their relevant expertise. Other contexts employed recently include 4-minute shower times (Fielding-Wells & Makar, 2012), teenagers' music listening habits (Manor & Ben-Zvi, 2015), and making sausages in a manufacturing scenario (Konold & Harradine, 2014). The re-emergence of context in research studies is an enormous motivator in terms of realism and relevance for students, as well as ownership if they collect the data themselves.

Although the *Australian Curriculum* (e.g., ACARA, 2015a, b, c) includes three cross-curriculum priorities—Aboriginal and Torres Strait Islander Histories and Cultures, Asia and Australia's Engagement with Asia, and Sustainability—all of which would provide meaningful contexts for statistical investigations, specific suggestions within the curriculum appear to be fairly sparse, found in the occasional elaboration of the content. Watson and English (2015) chose the context of sustainability in the Australian curriculum to introduce the practice of statistics by asking students if they were environmentally friendly, using data first from the students and then from the Australian Bureau of Statistics *CensusAtSchool* site (<http://www.abs.gov.au/censusatschool>). The work of Fielding-Wells (2014) and Fielding-Wells and Makar (2015), using an approach based on "inquiry" and "argumentation", fits well in a variety of contexts, as does that of English (2010) using a modelling approach.

The Chief Scientist of Australia's drive (Office of the Chief Scientist, 2013) to increase the number of engineers, technologists and scientists for Australia to be more competitive internationally has filtered down to the school level and provides an excellent opportunity to provide new contexts for students to collect and analyse data to provide evidence for a scientific theory or for the success of an engineered design. It is also likely that others researching applications of mathematics across the school curriculum, for example with a general focus on numeracy (e.g., Geiger, Forgasz, & Goos, 2015) will create similar opportunities to focus on contexts involving statistics.

Question: Can and should statistics education researchers form partnerships with researchers in other subject areas to trial authentic and practical activities, especially for high schools?

Affective Variables

Moving away from specific subject content, statistics education researchers have followed the lead of other mathematics education researchers to look at non-cognitive affective variables influencing learners and their teachers in relation to statistics. Psychologists have studied mathematics anxiety at all levels of learning for a long time and the Mathematics Anxiety Rating Scale (MARS) instrument was developed in 1972 (Richardson & Suinn). A statistics version called STARS (Cruise, Cash, & Bolton, 1985) was developed and then adapted for different levels of learners. STARS was one of the instruments used by Williams (2012) in a study of the complex relationship of variables affecting worry and anxiety for graduate students of statistics and by Chiesi and Primi (2010) in considering factors related to student achievement in an introductory statistics course. These studies represent the general focus of anxiety research and are mainly related to tertiary learners. The question of the relationship of anxiety and performance was recently the subject of a literature review by Carey, Hill, Devine, and Szucs (2016). It was focussed more generally on mathematics but is likely to apply equally to the study of statistics. Reviewing conflicting evidence for a Deficit theory suggesting poor performance causes anxiety, and a Debilitating anxiety theory suggesting anxiety causes poor performance, the authors also presented evidence for a complex Reciprocal theory that proposes a vicious cycle between anxiety and performance. This area offers challenges for research in statistics education, to find methods of alleviating anxiety about statistics at the same time as improving performance.

The non-cognitive area receiving the most attention by statistics education researchers, however, is attitudes. Gal and Ginsburg (1994) provided an early review of the research and outcomes to that point, and provided critical analysis and suggestions for the way forward. Although recognising the influence of beliefs, they found little evidence of research on beliefs and called for further research on both. Schau, Stevens, Daughinee, and Vecchio (1995) provided an updated instrument, Survey of Attitudes to Statistics (SATS), which has been utilised in many research projects and features in several papers in the *SERJ* special issue on Attitudes to Statistics (Volume 11, Number 2) considering teachers, students, relationship to achievement, and longitudinal change, as well as theoretical issues. Other studies have looked at previous experience and attitudes in relation to assessment (Dempster & McCorry, 2009), as well as the effect of student-designed data collection on their attitudes to statistics (Carnell, 2008); another comprehensive review is presented by Estrada, Batanero, and Lancaster (2011).

Closely linked to attitudes in many more recent research studies are beliefs and these are a focus of Pierce and Chick's (2011) review in relation to teachers of statistics. Alldredge and Brown (2006) considered beliefs in relation to performance of college students in a statistics course, taking account of gender and instructional software, and Zieffler, Park, Garfield, delMas, and Bjornsdottir (2012) introduced a comprehensive Statistics Teaching Inventory as an instrument to measure tertiary teachers' practices and beliefs. This inventory appears to offer promise for future research and interventions. Also linked to attitudes and beliefs is teacher confidence to teach statistics. At the tertiary level confidence for teaching statistics may be assumed (no research was found on this topic) but not necessarily at the school level. Three studies in Australia and New Zealand suggest that there is a need to focus on teachers' confidence as well as their content and pedagogical knowledge (Callingham & Watson, 2014; Callingham, Watson, Collis, & Moritz, 1995; Edwards, 1996).

Again following research in other areas of education, including mathematics, Carmichael developed the theme of "interest" of school students in statistics (e.g., Carmichael, Callingham, Hay, & Watson, 2010a, 2010b), measuring the construct and considering its relationship to self-efficacy and prior mathematics achievement. Very recently Sproesser, Engel, and Kuntze (in press) have considered statistics-specific learning environments for fostering both self-concept and interest.

Generally "belief" research is about non-cognitive beliefs but there is also literature on beliefs that students hold about aspects of statistics itself that can be detrimental to their learning of the subject. These beliefs often arise from cultural experiences, as well as life experiences such as playing games, and usually surface in research reporting conversations or responses to written tasks. They are most often found in approaches to problems associated with probability or in the use of contextual knowledge contrary to statistical evidence when drawing a conclusion (e.g., Amir & Williams, 1999; Sharma, 2014; Taylor, 1995; Watson & Callingham, 2015; Watson & Moritz, 2003). These are often classified as "misconceptions" alongside other arithmetic mistakes students make. Perhaps research could explore more closely what can be done to remedy such beliefs.

Question: Could a longitudinal research project explore whether an effective introduction to the practice of statistics in motivating contexts during the school years alleviates the anxiety created when students are introduced to a more complex set of tools to use in the practice in later years?

Educating Teachers

In its decadal plan for *The Mathematical Sciences in Australia*, the Australian Academy of Science (2016) claims that the term "mathematical sciences" encompasses "mathematics, statistics and the range of mathematics-based disciplines including teaching, teacher education and educational research" (p. 1). This description means that all members of MERGA are part of the mathematical sciences in Australia. The *first* objective of the decadal plan is "Giving all Australian school children access to outstanding mathematics teachers" (p. 26), with three recommendations for governments, schools, universities, and the teaching profession: Urgently increase professional development opportunities for out-of-field teachers, set national qualification standards for teachers and ensure universities can provide it, and provide rewarding career paths for primary and secondary teachers (p. 29). The evidence from statistics education research is that the objective is just as relevant for statistics as it is for the rest of mathematics.

As recognised in the 2008 ICMI/IASE Study on teaching statistics at the school level (Batanero et al., 2011), the part played by teachers in achieving the goals of the statistics curriculum is critical. When the curriculum was introduced, most mathematics teachers had no formal education in statistics or how to teach statistics. Research into teachers' knowledge has sometimes produced disturbing results, as when Jacobbe and Carvalho (2011) found that teachers' understanding of averages was "similar to that of students" (p. 207). Watson and Callingham (2013, 2014) also found deficiencies in the pedagogical content knowledge (PCK; Shulman, 1987) of teachers related to two-way table interpretation and to likelihood related to sample size; and Watson, Callingham, and Nathan (2009) found some teachers displayed low levels of PCK related to an inferential prediction from a pictograph and to anticipating students' responses to the questions. As well as reporting on teachers' beliefs, attitudes, and knowledge, Batanero et al. included chapters on issues associated with teacher education programs. Finding time in many preservice programs to provide indepth understanding (e.g., Pfannkuch & Ben-Zvi, 2011) is often difficult, as is finding the necessary time for statistics in professional learning in inservice programs. Appreciating the difficulties from the beginning Makar and Fielding-Wells (2011) gave advice to teacher educators from their research on moving teachers along to understanding and teaching statistical investigations. Convincing teachers that carrying out complete investigations is important, despite NAPLAN (ACARA, 2013b) ignoring them, is an ongoing challenge.

The emergence of NAPLAN testing has presented a further challenge for teachers over and above preparing their students for the tests. It is later necessary to interpret the detailed reports provided for schools and teachers about student performance. NAPLAN, as well as other state-wide testing, prompted a collaborative project in Victoria to develop a statistical literacy hierarchy for interpreting system data (Pierce, Chick, Watson, Les, & Dalton, 2014) and use it to improve teachers' professional statistical literacy (Pierce & Chick, 2013; Pierce, Chick, & Wander, 2014). Results indicating teachers' continuing problems with the interpretation of box plots reinforces the earlier concern of Bakker, Biehler and Konold (2005) for students.

Although there have been many professional development programs to improve teachers' content knowledge and PCK, the measurement of change has not been easy. Based on Shulman's (1987) seven types of knowledge required for teaching, Watson (2001) developed a profile to assess these for teachers of data and chance. Based on this profile Callingham and Watson (2011) developed a survey instrument to measure PCK in statistics of teachers in a program directly aimed at this goal. Over a year the change in PCK had an effect size of 0.59. The next more complex step is the measurement of change in the students of teachers in professional development programs.

Using data related to the same project, Callingham, Carmichael, and Watson (2015) found that students' improvement in statistical understanding was related to their teachers' initial PCK. Except for a study more widely in mathematics education by Baumert et al. (2010) showing an increase in student outcomes related to teacher PCK, there are no other known outcomes of this type showing outcomes from teacher learning to student learning. Although very complex to design, more research of this type is needed to justify and cement the importance of professional learning for teachers. The recent publication of *Statistical Education of Teachers* (*SET*) (Franklin et al., 2015) by the American Statistical Association provides an excellent benchmark for the understanding required by teachers at different levels. Basing professional learning for teachers around *SET* recommendations, students of these teachers can then be monitored to measure the change in their understanding. Killion (2003) provides an 8-step model for evaluating staff development programs, which includes within it the basic practice of statistics, preceded by assessing the ability of providers to evaluate, standards and indicators of success, how change will occur, and dissemination at the end. Such complex research design will require many resources and care in planning.

Following the release of *National Professional Standards for Teachers* (Australian Institute for Teaching and School Leadership [AITSL], 2011) the pressure is now on preservice programs to demonstrate their impact on graduate performance within the program and also on graduate outcomes when entering the workforce (AITSL, 2015). The need for evidence to support claims of impact provides a fertile ground for statistics education researchers to design studies that will collect the evidence needed. One aspect of this could be the adoption of a model such as Killion's (2003) for following the teacher education program to evaluate the preservice teacher growth followed by the growth of their students in the school classroom. Although developed in a different context, a setting such as described by Leavy and Hourigan (2015), using the Japanese Lesson Study model based on analysing data from games with primary preservice teachers and their students, could provide a starting point for planning.

Question: How can statistics education researchers convince funding agencies and school systems to invest adequate money and time into projects that will actually have the resources to follow the professional learning of inservice and preservice teachers through to the statistical outcomes of their students?

Technology

The relationship of technology to statistics education and statistics education research is three-fold. One aspect is technology's role in the dissemination of professional learning to teachers, the second is providing tools and a learning environment for carrying out the practice of statistics, and the third is its use for assessment.

Professional Learning

In terms of professional learning, the Australian Association of Mathematics Teachers (AAMT) ran the LUDDITE (Learning the Unlikely at Distance Delivered as an Information Technology Enterprise) program from 1994-1997, experimenting with various delivery modes, including narrowcast television in Victoria (Watson, Baxter, Olssen, & Lovitt, 1996); video-conferences across Australia (Watson, 1996); creation of a Chance-and-Data-in-the-News website with the Hobart *Mercury* newspaper (now converted to Numeracy at <http://www.tas-education.org/numeracy/>); and production of a comprehensive CD-ROM for professional learning called the C&D PDCD (Watson, 1998a; Watson & Moritz, 1997). This CD-ROM included extracts from the *National Statement* (AEC, 1991) and *Profile* (AEC, 1994a); video extracts from advertisements, ABC news, Media Watch, as well as students discussing their approaches to problems; links to David Moore's book (1991) and video series (1992); activities from the AAMT *Maths Works* project for the *National Statement* (Watson, 1994); and the software *ProbSim* (Konold & Miller, 1994a) and *DataScope* (Konold & Miller, 1994b) for Macintosh users.

The world of technology has changed massively since the AAMT LUDDITE project but the AAMT has moved forward to providing webinars throughout the year with experts on topics across the curriculum; on-line communities for teachers to share ideas, engage in activities, explore resources, and participate in on-line learning (<aamt.edu.au/communities>); resources related to Make-It-Count, a project to improve numeracy outcomes for indigenous learners (<mic.aamt.edu.au>); and Top Drawer Teachers, an on-line resource to support the *Australian Curriculum: Mathematics* (ACARA, 2015a), including a section on Statistics (<topdrawer.aamt.edu.au>). Top Drawer includes sections on "Big Ideas", "Misunderstandings", "Good Teaching", "Assessment", "Activities", and "Downloads". Currently the AAMT is developing an on-line portal to act as a clearinghouse for the five Australian Mathematics and Science Partnerships Projects (AMSPP) related to Mathematics. The AAMT is also a partner with the Australian Academy of Science in the new Mathematics by Inquiry Project (<http://www.science.org.au/learning/schools/mathematics-inquiry>), which will also be linked to the portal and offer many opportunities for research across the curriculum, including statistics.

Tools for Learning

For students to carry out the practice of statistics, new software of different types and appropriate for different levels of engagement is appearing regularly. Issues of downloading these to different devices as technological upgrades occur cause difficulty for longitudinal planning of research projects. Another issue is the expertise of the user when engaging with the software. Most statistics apps make the assumption that the user knows the purpose of the analysis to be performed, inputs the data, and can interpret the statistical output. Tabor and Franklin's (2013) textbook is an example at the senior secondary and lower tertiary level where instruction and theory are provided, as well as apps to complete the procedures associated with the topics covered. A larger program developed by Lock Morgan, Lock, Lock, Lock, and Lock (2014) called *StatKey* (<www.lock5stat.com/statkey>) aims to develop student understanding as well as implement randomisation tests for simulating confidence intervals and hypothesis tests. Software such as this is necessary to satisfy Cobb's (2007) appeal to drop theoretical statistics in first year university level and replace it with three Rs—randomise, repeat, reject—using the original data to see if the outcome is unusual or not. The work of Chance and Rossman (2006), Rossman and Chance (2014) and Tintle et al. (2014, 2015) illustrates Cobb's appeal.

At the upper primary and middle school level, however, I know of only one software that actually allows students to construct statistical understanding as they interact with it. When Konold and Miller found it was too expensive to convert their early Macintosh-based programs *DataScope* (1994a) and *ProbSim* (1994b) for PC users, they started on a new project with funding from the National Science Foundation to use the latest drag-and-drop technology and the model of the senior secondary software *Fathom* (Finzer, 2007) to create *TinkerPlots* (Konold & Miller, 2005, 2011). Version 1, for visualising data, grew from *DataScope*, whereas Version 2 (and later versions) added the data simulation capabilities evolved from *ProbSim*. Across the school years from about Grade 3 to senior secondary grades *TinkerPlots* and *Fathom* provide a dynamic data analysis environment for students to carry out statistical investigations. Many applications for the two programs are given in Biehler, Ben-Zvi, Bakker, and Makar (2013). Watson and Fitzallen (2016) review the learning affordances of *TinkerPlots* from many recent research projects based on the software. By Grade 10, for example, students can construct for themselves a simulation in *TinkerPlots* to collect 100 random samples from the original data to test a hypothesis about the difference in two groups (Watson, 2014), rather than enter data into an app that provides an instantaneous answer.

Assessment

Using computer technology for assessment in statistics education was initiated in the United Kingdom as part of the UK World Class Tests (World Class Arena, 2015). The Mathematics Assessment Resource Service at Durham University was a contributor, with Ridgway, McCuster, and Nicholson (2005) devising and evaluating tasks in the area of statistics learning in high school. Tasks, for example, give the students the opportunity to manipulate data to answer questions in context that require higher-order thinking. At the tertiary level, more traditional statistics examinations are administered via desktop computers (e.g., Van Duuren & Harvey, 2010) and individualised, automated assessments have been devised (e.g., Spencer, 2010), but more research is needed into the effectiveness, efficiency, and popularity of the innovations. Garfield, delMas, and Chance (2002) provide, through the ARTIST project, a wide array of on-line multiple-choice items assessing statistical thinking at the senior secondary and tertiary level that can be used within research projects for measuring understanding and change. Ziegler (2014) has suggested updating the ARTIST instrument acknowledging the importance of simulation techniques, which have not been previously included.

Callingham (2011) considers many of the issues associated with using computer technology for assessment at the middle school level, including its use within classroom projects such as the creation of student posters. Her main criticism, however, is of the assessment of procedural, rather than conceptual, understanding by multiple-choice items on international (e.g., Trends in International Mathematics and Science Study) and national (e.g., NAPLAN) tests. This is in contrast to her praise of the World Class Arena on-line tasks, which are not constrained by assessing official curricula, but ask students to participate in data analysis through the software.

Question: Can statistics education researchers help ensure that research projects involving new technology for professional learning in statistics, students' data analysis, or assessment of statistical goals are not restricted to a focus on algorithms and procedures but involve conceptual development and understanding?

Current Challenges for Statistics Education Research

STEM (Statistics, Technology, Engineering, Mathematics)

The implications of the current strong interest in STEM education are huge for statistics education researchers, providing a political and economic rationale for requests for funding (e.g., English, Watson, & Fitzallen, 2015). Although I am happy for Australia to have more scientists, technologists, and engineers to increase our scientific and economic stature in the world, which appears to be the public image portrayed by the Office of the Chief Scientist (2013), my selfish interest in STEM is two-fold. Not only do I see data modelling and the practice of statistics as a natural way to satisfy the M in STEM, but also all of the STEM disciplines, including other parts of mathematics, provide contexts to make statistics come alive. Almost any statistical investigation one can imagine involves the measurement part of the *Mathematics* curriculum for data collection as well as proportional reasoning for analysis (and algebra at higher levels). Through investigations in Science and Engineering as envisaged by the *Next Generation Science Standards* (National Research Council, 2013) the need for the collection of evidence to support a scientific principle or an engineering design is fundamental. Collecting data is a natural extension to activities in order to provide evidence to make a decision in the context of the investigation. The introduction of the *Australian Curriculum: Design and Technologies and Digital Technologies* (ACARA, 2015c) provides an excellent foundation for links to Engineering in the Design and Technologies strand and to Data in the Digital Technologies strand. For the Digital Technologies section, one strand is Data Representation, which has been created to link to Data Representation and Interpretation within the *Mathematics* curriculum. The understanding of how data are represented and structured symbolically is parallel to the planning of data collection in the practice of statistics, as well as data cleaning and deciding the type of representation to have data in a manageable form for analysis. There are hence many opportunities for research in the area of data modelling (e.g., Doerr & English, 2003; Konold et al., 2014).

The definition of STEM education has become quite controversial, ranging from a simple acronym for separate subjects linked together for political purposes, to an interdisciplinary approach acknowledging the dependence of each on the other to some degree. English (2015) provides a nice summary of the range of views, including that of Clarke (2014) that the four subjects have little in common. Having made this claim, however, Clarke goes on to discuss ways of crossing the boundaries of the four disciplines in the classroom: monitor the discourse including students, with teachers making the links across the boundaries clear to students; focus on artifacts used in the classroom for the implementation of a STEM-based activity; emphasise the reasoning that occurs across the boundaries for students either publicly or privately in the classroom; expect evidence to answer the question of "why", which should always be an expectation in the classroom. Evidence of course can refer to data collected to be analysed to support a conclusion reached. Clarke concludes, "STEM could be the vehicle for a new approach to disciplinary inclusivity and the transformation of the organising principles of the curriculum and of teacher expertise" (slide 94). This is an important point that fits well with the practice of statistics being a vehicle of inclusivity.

Fitzallen (2015) also considers the range of definitions of STEM and the debate about whether STEM should be a "discipline" in the Australian curriculum. Supporting the "no" case is the belief that teachers do not have the content knowledge and PCK across the four individual disciplines and their potential integration to teach a STEM discipline effectively (Sanders, 2009). Fitzallen suggests, however, that an approach to integration of mathematics and science (Treacy & O'Donoghue, 2014) could be applied when any two or more of the separate STEM disciplines are treated at the same time. This approach is based on four components: knowledge development, synthesis, and application; focused inquiry resulting in higher order learning; applicability to real world situations; and rich tasks. Statistics fits naturally in all four components.

The most recent initiative of the Statistical Society of Australia (SSA) was a colloquium on "STEMS: Putting Statistics into STEM in the Age of Data" (SSA, 2016). With keynote speakers and panelists from Harvard University, the Australian government, the Commonwealth Bank, ACARA, the Chief Scientist's Office, an ARC Centre for Excellence, AMSI, SSA, Qantas, and the NSW Board of Studies, none were from Statistics Education at the school level. With the aim of "initiating a transformation in the provision of Statistics education in Australia, from kindergarten to post-doctoral levels" (SSA, 2016), one wonders what will be initiated from kindergarten to Year 10.

Question: In this complex international milieu of STEM, are statistics education researchers willing to take on the challenge of collaborating with researchers in other STEM disciplines to develop and trial meaningful activities across the school years that show the power of statistics to make decisions in the other disciplines?

UNESCO Education for Sustainable Development

From 2005 to 2014, UNESCO sponsored the UN Decade of Education for Sustainable Development. One of the contributions of Australia to the Decade was *Living Sustainably: The Australian Government's Action Plan for Education for Sustainability* (Department of Environment, Water, Heritage, and the Arts, 2009). Among the six objectives of the Government for school education was embedding sustainability in the national curriculum (2.3.5, p. 24). This objective appears as a Cross-curriculum Priority in the current curriculum.

> Sustainability addresses the ongoing capacity of Earth to maintain all life… Education for sustainability develops the knowledge, skills, values and world views necessary for people to act in ways that contribute to more sustainable patterns of living… Sustainability education is futures-oriented, focusing on protecting environments and creating a more ecologically and socially just world through informed action. (<http://www.australiancurriculum.edu.au/crosscurriculumpriorities/sustainability/overview>)

Associated with the Decade has been the designation of UNESCO Regional Centres of Expertise in Education for Sustainable Development (EfSD) around the world (United Nations University, 2004), of which there are at least four regional communities in Australia, at Greater Western Sydney (UWS), Murray-Darling (Charles Sturt, Albury), Gippsland (Federation, Victoria), and most recently Tasmania (UTAS). These Centres are not just for universities, but for communities involving other interested organisations, school systems, local government, and industry. Among the aims of these Regional Centres of Expertise relevant to statistics education research is the acknowledged need for research on the implementation of EfSD in schools. Many areas offer promise for data modelling and statistical investigations at the school level, including Biodiversity, Climate change, Water, or Disaster Risk Reduction. Activities where evidence is collected are likely to be associated with one or more STEM disciplines and evidence is likely to include data.

In the light of EfSD, it is instructive to re-examine the five societal challenges for Australia identified by the Prime Minister's Science, Engineering and Innovation Council (Office of the Chief Scientist, 2012), which are quoted by the Office of the Chief Scientist (2013) as part of a strategic approach to STEM in the national interest. These are

- Living with a changing environment
- Promoting population health and wellbeing
- Managing our food and water assets
- Securing Australia's place in a changing world
- Lifting productivity and economic growth (p. 5)

Although not receiving as much publicity as the final two, the first three would seem to fit firmly within the auspices of EfSD. One wonders if the list was created in the order of importance. If "[s]ustainable economic growth is a vital strategy for the future of humanity" (p. 7), how will sustainability as envisaged by the UN and the *Australian Curriculum* be accomplished? In any case there seems to be considerable research opportunity to devise activities across the school years based on data modelling and statistical investigations that satisfy both EfSD and at least the first three of the Chief Scientist's societal challenges employing STEM disciplines (e.g., Watson & English, 2015).

Question: What is the intersection of EfSD and STEM? Can statistics education researchers work proactively to employ STEM disciplines to contribute more actively to Educating for Sustainable Development where data provide the evidence for decision-making?

Statistics and the Mathematics by Inquiry Project

The rationale for Mathematics by Inquiry, found in a *Desktop Review* (Stacey, Vincent, Stephens, & Holton, 2015), develops a case for

> a long-term Australian initiative of co-ordinated resource development and evaluation, research and professional learning for system-wide improvement in mathematics education. Within this program, mathematical inquiry, problem solving and reasoning are promoted in multiple ways. (p. 4)

"Long-term" is an important adjective and a plan for sustainability beyond the end of the project in the middle of 2018 needs to be included in the framework promised as a starting point. Other promises include inquiry experiences for every level, Foundation to Year 10. Further encouragement is found in the promise of teaching resources based on mathematical reasoning such as modelling, capitalising on emerging technologies and software. Accompanying this will be professional resources highlighting higher order thinking and mathematical inquiry in STEM contexts and hopefully the first three societal challenges as well. The most significant part of the dissemination, besides the AAMT portal, is the provision of 240 trained "champions" across Australia. The project must recognise the critical importance the professional development of teachers across the entire curriculum or the goals will not be achieved. The importance of problem solving and reasoning, two of the Proficiencies in the Mathematics curriculum, cannot be ignored, as well as the links to the *General Capabilities* (ACARA, 2013a), and Stacey et al. (2015) recommend that ACARA review the curriculum to improve the descriptions, particularly related to the Proficiencies.

Careful reading the Mathematics by Inquiry fact sheet (<https://www.science.org.au/files/userfiles/learning/documents/mbi-fact-sheet-2015.pdf>) or the *Desktop Review* (Stacey et al., 2015), however, shows that "inquiry" is taken as an undefined term. Fielding-Wells and Makar (2015) provide assistance, using other sources to describe mathematical inquiry as "an approach to teaching and learning where students address ill-structured problems that rely on mathematical (or statistical) evidence" (p. 5). They then go on to describe an ill-structured problem as "one in which the problem statement and/or pathway for solving the problem contain ambiguities that require negotiation" (p. 5). Instead of "inquiry", Stacey et al. use the phrase "inquiry-based pedagogy", acknowledging that it "is founded on the principle that students should be actively and socially engaged in the process of learning, constructing new concepts based on their current knowledge and understanding" (p. 11). Then, however, they claim that "this very open student-led interpretation of inquiry-based pedagogy has only a very small place in mathematics" (p. 11).

> Instead the best investigative pedagogies for mathematics use 'well engineered' mathematical problems, where engagement in the problem solving process individually and with others and supported by the teacher will assist in the development of targeted concepts, or strategic skills, or the ability to transfer knowledge. There is a spectrum of purposes and thus variations in the pedagogy are required. A few outcomes within the ACM can be well served with quite open inquiry-based learning (e.g. experience in statistical investigation). More topics will use a structured investigation spreading over say two lessons so that students explore a particular concept, or undertake a real world mathematical modelling task, to experience how to identify the relationships involved and make appropriate assumptions to apply their mathematical knowledge in real situations. (p. 11)

Except for a footnote on p. 6 saying "mathematics" includes statistics and a comment on advanced graphics calculators, the only place that a word related to statistics appears in the *Review* is in the above extract, in parenthesis. Putting these statements together one might infer that Statistics may have "only a very small place" in Mathematics by Inquiry.

Question: How much attention will Statistics receive as part of Mathematics by Inquiry? What will "inquiry" mean by the end?

Final Thoughts

Sustainability of outcomes from research and professional learning projects in statistics education is an issue for the future. An ARC project may be funded for three or four years and show significant outcomes for students or teachers but then the initiatives die because further funding is not available to continue them or they are not readily transferrable to new environments. It is to be hoped that new projects, such as Mathematics by Inquiry, will indeed, as promised, take advantage of the large amount of professional material already available for teachers through the AAMT, AMSPP, NCTM, and other resources, even related to statistics (e.g., Watson et al., 2011), and not "reinvent the wheel". Resources need to go into professional learning in order for the messages to reach teachers and students to produce change in outcomes.

Initiatives arising from STEM educational research, Education for Sustainable Development, and Mathematics by Inquiry run the risk of being ignored by schools and teachers in favour of the procedural approach that will achieve higher NAPLAN scores for their students. Somehow statistics education researchers, and dare I say other mathematics education researchers, need to make strong representation to ACARA both in terms of curriculum content and national testing. Inquiry and high level thinking across disciplines will not occur if not valued by the decision makers in the country who have the power to decide to fund changes in national testing.

Unlike some of our university colleagues who may choose to live in their ivory towers, we exist in the real world of education. We, whether statistics or other mathematics education researchers, must take our messages beyond the teachers or students in our research projects, to the wider education community: to classrooms, to principals, to systems, to authorities, and to governments. This means not only writing for high-ranking journals, but also writing for teachers and children, talking to parents about our messages, working professionally with those outside our research projects, speaking out in the media when necessary, and advocating to government agencies making decisions. Ultimately this is where our *impact* (and that of our research) will be felt, even if it is difficult to measure. Maybe one day this will be recognised by those who work out the rankings for Excellence in Research for Australia.

Question: Do we have the courage to take up the challenge to reach out beyond our research world?

Acknowledgements

Thanks to the following people who provided various advice in the preparation of this paper: Helen Chick, Noleine Fitzallen, Katie Makar, and Caroline Smith. The funding for the author's time to prepare the paper was provided by UniSuper Limited.

References

Alldredge, J. R., & Brown, G. R. (2006). Association of course performance with students beliefs: An analysis by gender and instructional software environment. *Statistics Education Research Journal, 5*(1), 64-77. Retrieved from http://iase-web.org/documents/SERJ/SERJ5(1)_Alldredge_Brown.pdf

Amir, G. S., & Williams, J. S. (1999). Cultural influences on children's probabilistic thinking. *Journal of Mathematical Behavior, 18*, 85-107.

Australian Academy of Science. (2016). *The mathematical sciences in Australia: A vision for 2025*. Canberra: Author.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013a). *General capabilities in the Australian Curriculum, January, 2013 (updated September 2014).* Sydney, NSW: ACARA.
Australian Curriculum, Assessment and Reporting Authority. (2013b). *NAPLAN: National Assessment Program Literacy and Numeracy*. Sydney: Author. Retrieved from http://www.nap.edu.au/naplan/naplan.html
Australian Curriculum, Assessment and Reporting Authority. (2015a). *Australian Curriculum: Mathematics, Version 8.1, 15 December 2015.* Sydney, NSW: ACARA.
Australian Curriculum, Assessment and Reporting Authority. (2015b). *Australian Curriculum: Science, Version 8.1, 15 December 2015.* Sydney, NSW: ACARA.
Australian Curriculum, Assessment and Reporting Authority. (2015c). *Australian Curriculum: Design and Technologies and Digital Technologies, Version 8.1, 15 December, 2015.* Sydney, NSW: ACARA.
Australian Education Council. (1991). *A national statement on mathematics for Australian schools.* Melbourne: Author.
Australian Education Council. (1994a). *Mathematics - A curriculum profile for Australian schools.* Melbourne: Curriculum Corporation.
Australian Education Council. (1994b). *Mathematics - Work samples*. Melbourne: Curriculum Corporation.
Australian Institute for Teaching and School Leadership. (2011). *National professional standards for teachers*. Carlton South, VIC: Education Services Australia.
Australian Institute for Teaching and School Leadership. (2015). *Classroom ready: Demonstrating the impact on student learning of initial teacher education programs.* Melbourne: Australian Institute for Teaching and School Leadership.
Bakker, A., Biehler, R., & Konold, C. (2005). Should young students learn about box plots? In G. Burrill & M. Camden (Eds.), *Curricular Development in Statistics Education. International Association for Statistical Education (IASE) Roundtable, Lund, Sweden, 2004* (pp. 163-189). Voorburg, The Netherlands: International Statistical Institute.
Batanero, C., Burrill, G., & Reading, C. (Eds.). (2011). *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study*. Dordrecht, The Netherlands: Springer.
Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133-180.
Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM Mathematics Education*, *44*, 913-925.
Ben-Zvi, D., & Garfield, J. (Eds.). (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht, The Netherlands: Kluwer.
Ben-Zvi, D., Garfield, J., & Makar, K. (Eds.). (2016). *International handbook of research in statistics education.* Manuscript in preparation. New York: Springer.
Ben-Zvi, D., & Makar, K. (Eds.). (2016). *The teaching and learning of statistics*. New York: Springer.
Bidgood, P., Hunt, N., & Jolliffe, F. (2010). *Assessment methods in statistical education: An international perspective*. Chichester, UK: Wiley.
Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. Clement, A. J. Bishop, C. Keitel, J. Kilpatrick, & A. Y. L. Leung (Eds.), *Third international handbook on mathematics education* (pp. 643-689). New York: Springer.
Budgett, S., & Pfannkuch, M. (2010). Assessing students' statistical literacy. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 103-121). Chichester, UK: Wiley.
Budgett, S., Pfannkuch, M., & Franklin, C. (in press). Building conceptual understanding of probability models: Visualizing chance. In C. Hirsch (Ed.), *Annual Perspectives in Mathematics Education 2016: Mathematical modeling and modeling mathematics*. Reston, VA: National Council of Teachers of Mathematics.
Burrill, G. F. (Ed.). (2006). *Thinking and reasoning with data and chance. Sixty-eighth yearbook*. Reston, VA: National Council of Teachers of Mathematics.
Callingham, R. (2011). Assessing statistical understanding in middle schools: Emerging issues in a technology-rich environment. *Technology Innovations in Statistics Education*, *5*(1). Retrieved from http://escholarship.org/uc/item/3qr2p70t

Callingham, R., Carmichael, C., & Watson, J. (2015). Explaining student achievement: The influence of teachers' pedagogical content knowledge in statistics. *International Journal of Science and Mathematics Education*. DOI 10.1007/s10763-015-9653-2

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. (2014). Teachers' confidence in teaching statistical ideas. In K. Makar, B. deSousa, & R. Gould (Eds.), *Sustainability in statistics education* (Proceedings of the 9th International Conference on the Teaching of Statistics, Flagstaff, Arizona, July 13-18). Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://icots.info/9/proceedings/pdfs/ICOTS9_C216_CALLINGHAM.pdf

Callingham, R. A., Watson, J. M., Collis, K. F., & Moritz, J. B. (1995). Teacher attitudes towards chance and data. In B. Atweh & S. Flavel (Eds.), *Galtha* (Proceedings of the 18th Annual Conference of the Mathematics Education Research Group of Australasia, pp. 143-150). Darwin, NT: MERGA.

Carey, E., Hill, F., Devine, A., & Szucs, D. (2016). The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance. *Frontiers in Psychology, 6*, Article 1987. DOI: 10.3389/fpsyg.2015.01987

Carmichael, C. S., Callingham, R., Hay, I., & Watson, J. M. (2010a). Statistical literacy in the middle school: The relationship between interest, self-efficacy and prior mathematics achievement. *Australian Journal of Educational and Developmental Psychology, 10*, 83-93.

Carmichael, C. S., Callingham, R., Hay, I., & Watson, J. M. (2010b). Measuring middle school students' interest in statistical literacy. *Mathematics Education Research Journal, 22*(3), 9-39.

Carnell, L. (2008). The effect of a student-designed data collection project on attitudes toward statistics. *Journal of Statistics Education, 16*(1). Retrieved from http://www.amstat.org/publications/jse/v16n1/carnell.html

Chance, B., & Rossman, A. (2006). Using simulation to teach and learn statistics. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education* (Proceedings of the 7th International Conference on Teaching of Statistics, Salvador, Bahai, Brazil. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://iase-web.org/documents/papers/icots7/7E1_CHAN.pdf

Chernoff, E., & Sriraman, B. (Eds.). (2014). *Probabilistic thinking: Presenting plural perspectives*. Amsterdam: Dordrecht.

Chick, H.L., & Watson, J. M. (2002). Collaborative influences on emergent statistical thinking – A case study. *Journal of Mathematical Behavior*, *21*, 371-400.

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal, 9*(1), 6-26. Retrieved from http://iase-web.org/documents/SERJ/SERJ9(1)_Chiesi_Primi.pdf

Clarke, D. (2014, July 12). *Disciplinary inclusivity in educational research design: Permeability and affordances in STEM education*. Keynote address presented to the conference: STEM Education and Our Planet: Making Connections Across Contexts. University of British Columbia, Vancouver, Canada.

Cobb, G. (2007). The introductory statistics course: A Ptolemaic curriculum? *Technology Innovations 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 http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Crites, T., & St. Laurent, R. (2015). *Putting essential understanding of statistics into practice in grades 9-12.* Reston, VA: National Council of Teachers of Mathematics.

Cruise, J. R., Cash, R. W., & Bolton, L. D. (1985). Development and validation of an instrument to measure statistical anxiety. In *American Statistical Association Proceedings of the Section on Statistical Education* (pp. 92-978). Alexandria, VA: American Statistical Association.

Day, L. (2013). Using statistics to explore cross-curricular and social issues opportunities. *The Australian Mathematics Teacher, 69*(4), 3-7.

delMas, R., Garfield, J., & Zieffler, A. (2014). Using *TinkerPlots* to develop tertiary students' statistical thinking in a modeling-based introductory statistics class. In T. Wassong, D. Frischemeier, P. R. Fischer, R. Hochmuth, & P. Bender, R. (Eds.), *Using tools for learning mathematics and statistics* (pp. 405-420). Heidelberg: Springer Spektrum.

Dempster, M., & McCorry, N. K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education*, *17*(2). Retrieved from http://www.amstat.org/publications/jse/v17n2/dempster.html

Department of the Environment, Water, Heritage and the Arts. (2009). *Living sustainably: The Australian Government's National Action Plan for Education for Sustainability*. Canberra: Commonwealth of Australia. Retrieved from http://www.environment.gov.au/system/files/resources/13887ab8-7e03-4b3e-82bb-139b2205a0af/files/national-action-plan.pdf

Doerr, H. M., & English, L. D. (2003). A modeling perspective on students' mathematical reasoning about data. *Journal for Research in Mathematics, 34*, 110-136.

Edwards, R. (1996). Teaching statistics: Teacher knowledge and confidence. In P.C. Clarkson (Ed.), *Technology in mathematics education* (Proceedings of the 19th Annual Conference of the Mathematics Education Research Group of Australasia, pp. 178-185). Melbourne, VIC: MERGA.

English, L. D. (2010). Young children's early modelling with data. *Mathematics Education Research Journal, 22*(2), 24-47.

English, L. D. (2012). Data modelling with first-grade students. *Educational Studies in Mathematics*, *81*, 15-30. DOI 10.1007/s10649-011-9377-3.

English, L. D. (2015). STEM: Challenges and opportunities for mathematics education. In K. Beswick, T. Muir, & J. Wells (Eds.), *Proceedings of the 39th annual conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 3-18). Hobart, Australia: PME.

English, L. D., Watson, J. M., & Fitzallen, N. (2015). *Modelling with data: Advancing STEM in the primary curriculum* [ARC Discovery Project 15010012, 2015-2018].

Estrada, A., Batanero, C., & Lancaster, S. (2011). Teachers' attitudes towards statistics. 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. 163-174). Dordrecht, The Netherlands: Springer.

Fielding-Wells, J. (2014). Where's your evidence? Challenging young students' equiprobability bias through argumentation. In K. Makar, B. deSousa, & R. Gould (Eds.), *Sustainability in statistics education* (Proceedings of the 9th International Conference on the Teaching of Statistics, Flagstaff, Arizona, July 13-18). Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_2B2_FIELDINGWELLS.pdf

Fielding-Wells, J., & Makar, K. (2012). Developing primary students' argumentation skills in inquiry-based mathematics classrooms. In K. T. Jan van Aalst, M. J. Jacobson, & P. Reimann (Ed.), *The Future of learning: Proceedings of the 10th International Conference of the Learning Sciences [ICLS 2012] - Volume 2 Short Papers, Symposia, and Abstracts* (pp. 149-153). Sydney, Australia: International Society of the Learning Sciences.

Fielding-Wells, J., & Makar, K. (2015). Inferring to a model: Using inquiry-based argumentation to challenge young children's expectations of equally likely outcomes. In A. Zieffler & E. Fry (Eds.), *Reasoning about uncertainty: Learning and teaching informal inferential reasoning* (pp. 1-27). Minneapolis, MN: Catalyst Press.

Finch, S., & Gordon, I. (2014). The development of a first course in statistical literacy for undergraduates. In H. MacGillivray, M. A. Martin, & B. Phillips (Eds.), *Topics from Australian conferences on teaching statistics: OZCOTS 2008-2012* (pp. 73-98). New York: Springer.

Finlay, E., & Lowe, I. (1993). *Chance and data: Exploring real data*. Melbourne: Curriculum Corporation.

Finzer, W. (2007). *Fathom*. [computer software] Emeryville. CA: Key Curriculum Press.

Fitzallen, N. (2015). STEM Education: What does mathematics have to offer? In M. Marshman, V. Geiger, & A. Bennison (Eds.), *Mathematics education in the margins* (Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia, pp. 237-244). Sunshine Coast: MERGA.

Franklin, C., Bargagliotti, A., Case, C., Kader, G., Scheaffer, R., & Spangler, D. (2015). *Statistics Education of Teachers*. Alexandria, VA: American Statistical Association. Retrieved from http://www.amstat.org/education/SET/SET.pdf

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 preK-12 curriculum framework.* Alexandria, VA: American Statistical Association. Retrieved from http://www.amstat.org/education/gaise/

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*, 1-51.

Gal, I. (2005). Towards "Probability Literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 39-63). New York: Springer.

Gal, I., & Garfield, J. (Eds.). (1997). *The assessment challenge in statistics education*. Amsterdam: IOS Press & The International Statistical Institute.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education, 2*(2). Retrieved from http://www.amstat.org/publications/jse/v2n2/gal.html

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. New York: Springer.

Garfield, J., delMas, R., & Chance, B. (2002). *The Assessment Resource Tools for Improving Statistical Thinking (ARTIST) Project. NSF CCLI grant ASA-0206571*. Retrieved from https://apps3.cehd.umn.edu/artist/index.html

Geiger, V., Forgasz, H., & Goos, M. (2015). A critical orientation to numeracy across the curriculum. *ZDM Mathematics Education, 47*, 611-624. DOI 10. 1007/s11858-014-0648-1

Goodchild, S. (1988). School pupils' understanding of average. *Teaching Statistics, 10*, 77-81.

Green, D. (1983). Shaking a six. *Mathematics in Schools, 12*(5), 29-32.

Jacobbe, T., & Carvalho, C. (2011). Teachers understanding of average. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study* (pp. 199-209). New York: Springer.

Jones, G. A. (Ed.). (2005). *Exploring probability in school: Challenges for teaching and learning.* New York: Springer.

Killion, J. (2003). 8 smooth steps: Solid footwork makes evaluation of staff development programs a song. *JSD The Learning Forward Journal*, *24*(4), 14-26.

Konold, C., Finzer, W., Kreetong, K., & Gaston, R. (2014, April). *Modeling as a core component of structuring data*. Paper presented at the annual meeting of the Research Conference of the National Council of Teachers of Mathematics, New Orleans, April, 2014.

Konold, C., & Harradine, A. (2014). Contexts for highlighting signal and noise. In P. Bender, R. Hochmuth, P. R. Fischer, D. Frischemeier, & T. Wassong (Eds.), *Using tools for learning mathematics and statistics* (pp. 237-250). Heidelberg: Springer Spektrum.

Konold, C., & Miller, C. (1994a). *ProbSim*. [Computer software] Amherst, MA: University of Massachusetts.

Konold, C., & Miller, C. (1994b). *DataScope*. [Computer software]. Santa Barbara, CA: Intellimation Library for the Macintosh.

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

Konold, C., & Miller, C. D. (2011). *TinkerPlots: Dynamic data exploration* [Computer software, Version 2.2]. Emeryville, CA: Key Curriculum Press.

Lajoie, S. P. (Ed.). (1998). *Reflections on statistics: Learning, teaching and assessment in grades K-12*. Mahwah, NJ: Lawrence Erlbaum.

Langrall, C., Nisbet, S., Mooney, E., & Jansem, S. (2011). The role of context expertise when comparing data. *Mathematical Thinking and Learning*, *13*, 47-67.

Leavy, A., & Hourigan, M. (2015). Motivating inquiry in statistics and probability in the primary classroom. *Teaching Statistics, 37*, 41-47.

Lehrer, R., & Romberg, T. (1996). Exploring children's data modeling. *Cognition and Instruction, 14*(1), 69-108.

Lock Morgan, K., Lock, R. H., Lock, P. F., Lock, E. F., & Lock, D. F. (2014). Statkey: Online tools for bookstrap intervals and randomization tests. In K. Makar, B. deSousa, & R. Gould (Eds.), *Sustainability in statistics education* (Proceedings of the 9th International Conference on the Teaching of Statistics, Flagstaff, Arizona, July 13-18). Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_9B2_LOCKMORGAN.pdf

Lovett, M. C., & Shah, P. (Eds.). (2007). *Thinking with data*. NY: Lawrence Erlbaum.

Lovitt, C., & Lowe, I. (1993). *Chance and Data Investigations, Volumes 1 and 2*. Melbourne: Curriculum Corporation.

MacGillivray, H., Martin, M. A., & Phillips, B. (Eds.). (2014). *Topics from Australian conferences on teaching statistics: OZCOTS 2008-2012.* New York: Springer.

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, *13*, 152-173.

Makar, K., & Fielding-Wells, J. (2011). Teaching teachers to teach statistical investigations. 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. 347-358). Dordrecht, The Netherlands: Springer.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82-105. Retrieved from http://iase-web.org/documents/SERJ/SERJ8(1)_Makar_Rubin.pdf

Manor, H., & Ben-Zvi, D. (2015). Students' articulations of uncertainty in informally exploring sampling distributions. In A. Zieffler & E. Fry (Eds.), *Reasoning about uncertainty: Learning and teaching informal inferential reasoning* (pp. 57-94). Minneapolis, MN: Catalyst Press.

Mevarech, Z. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics, 14*, 415-429.

Ministry of Education. (1992). *Mathematics in the New Zealand curriculum*. Wellington, NZ: Author.

Ministry of Education. (2009). *The New Zealand Curriculum: Mathematics standards for years 1-8*. Wellington, NZ: Author. Retrieved from http://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Learning-areas/Mathematics-and-statistics

Moore, D. S. (1991). *Statistics: Concepts and controversies* (3rd ed.). New York: W. H. Freeman and Co.

Moore, D. S. (1992). *Statistics: Decisions through data*. [Video tape] Lexington, MA: CoMap.

Moore, D. S., & McCabe, G. P. (1989). *Introduction to the practice of statistics*. New York: W.H. Freeman.

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

National Research Council. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Office of the Chief Scientist. (2012). *Strategic research priority setting: Key messages*. Canberra: Australian Government. Retrieved from http://www.chiefscientist.gov.au/wp-content/uploads/STRATEGIC-RESEARCH-PRIORITY-SETTINGS-powerpoint-AC2-4.pdf

Office of the Chief Scientist. (2013). *Science, technology, engineering and mathematics in the national interest: A strategic approach*. Canberra: Australian Government.

Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: A classroom study. *Mathematical Thinking and Learning*, *13*, 27-46.

Pfannkuch, M., & Ben-Zvi, D. (2011). Developing teachers' statistical thinking. 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. 323-333). Dordrecht, The Netherlands: Springer.

Pierce, R., & Chick, H. (2011). Teachers' beliefs about statistics education. 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. 151-162). Dordrecht, The Netherlands: Springer.

Pierce, R., & Chick, H. (2013). Workplace statistical literacy for teachers: Interpreting box plots. *Mathematics Education Research Journal, 25*, 189-205.

Pierce, R., Chick, H., & Wander, R. (2014). Improving teachers' professional statistical literacy. In H. MacGillivray, M. A. Martin, & B. Phillips (Eds.), *Topics from Australian conferences on teaching statistics: OZCOTS 2008-2012* (pp. 295-309). New York: Springer.

Pierce, R., Chick, H., Watson, J., Les, M., & Dalton, M. (2014). A statistical literacy hierarchy for interpreting educational system data. *Australian Journal of Education*, *58*, 195-217. DOI 10.1177/0004944114530067

Prodromou, T. (2015). Teaching statistics with technology. *The Australian Mathematics Teacher, 71*(3), 32-40.

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

Reaburn, R. (2014). Introductory statistics course tertiary students' understanding of p-values. *Statistics Education Research Journal, 13*(1), 53-65. Retrieved from http://iase-web.org/documents/SERJ/SERJ13(1)_Reaburn.pdf

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology, 19*, 551-554.

Ridgway, J., McCuster, S., & Nicholson, J. (2005). Uncovering and developing student statistical competences vie new interfaces. In G. Burrill & M. Camden (Eds.), *Curricular development in statistics education. International Association for Statistical Education (IASE) Roundtable, Lund, Sweden, 2004* (pp. 75-89). Voorburg, The Netherlands: International Statistical Institute.

Rossman, A. J., & Chance, B. L. (2014). Using simulation-based inference for learning introductory statistics. *WIREs Computational Statistics, 6*(4), 211-221.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10*(3). Retrieved from https://www.amstat.org/publications/jse/v10n3/rumsey2.html

Sanders, M. (2009). STEM, STEM education, STEMmania. *The Technology Teacher, 68*(4), 20-26.

Schau, C., Stevens, J., Daughinee, T. L., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes toward Statistics. *Educational and Psychological Measurement, 55*, 868-875.

Sharma, S. (2014). Cultural influences in probabilistic thinking. In E. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 657-681). Amsterdam: Dordrecht.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*, 1-22.

Shulte, A. P. (Ed.). (1981). *Teaching statistics and probability. 1981 yearbook.* Reston, VA: National Council of Teachers of Mathematics.

Spencer, N. (2010). An adaptive, automated, individualised assessment system for introductory statistics. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 211-221). Chichester, UK: Wiley.

Sproesser, U., Engel, J., & Kuntze, S. (in press). Fostering self-concept and interest for statistics through statistics-specific learning environments. *Statistics Education Research Journal*.

Stacey, K., Vincent, J., Stephens, M., & Holton, D. (2015). *Desktop review of mathematics school education pedagogical approaches and learning resources*. Canberra: Australian Academy of Science. Retrieved from https://docs.education.gov.au/system/files/doc/other/trim_review_paper_2_-_aas_-_final.pdf

Statistical Society of Australia. (2016). *STEMS: Putting statistics into STEM in the Age of Data*. Sydney; SSA. Retrieved from https://stems2016.com/

Strauss, S., & Bichler, E. (1988). The development of children's concept of the arithmetic average. *Journal for Research in Mathematics Education, 19*, 64-80.

Tabor, J., & Franklin, C. (2013). *Statistical reasoning in sports*. New York: W. H. Freeman and Co.

Taylor, M. (1995). Probability education: Can primary children cope? In B. Atweh & S. Flavel (Eds.), *Galtha* (Proceedings of the 18th Annual Conference of the Mathematics Education Research Group of Australasia, pp. 505-510). Darwin, NT: MERGA.

Tintle, N., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., & VanderStoep, J. (2015). *Introduction to statistical investigations.* Hoboken, NJ: Wiley.

Tintle, N. L., Rogers, A., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., & VanderStoep, J. (2014). Quantitative evidence for the use of simulation and randomization in the introductory statistics course. In K. Makar, B. deSousa, & R. Gould (Eds.), *Sustainability in statistics education* (Proceedings of the 9th International Conference on the Teaching of Statistics, Flagstaff, Arizona, July 13-18). Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_8A3_TINTLE.pdf

Treacy, P., & O'Donoghue, J. (2014). Authentic integration: A model for integrating mathematics and science in the classroom. *International Journal of Mathematical Education in Science and Technology, 45,* 703-718.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review, 90*, 293-315.

United Nations University. (2004). *Regional Centers of Expertise on Education for Sustainable development: Concept paper*. Tokyo: United Nations University. Retrieved from http://portal.unesco.org/education/en/files/37599/11039012553RCE_Concept_Paper.pdf/RCE_Concept_Paper.pdf

Van Duuren, M., & Harvey, A. (2010). A computer-based approach to statistics teaching and assessment in psychology. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 57-67). Chichester, UK: Wiley.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*, No. 421, 1-8.

Wassong, T., Frischemeier, D., Fischer, P. R., Hochmuth, R., & Bender, P. (Eds.). (2014). *Using tools for learning mathematics and statistics*. Heidelberg: Springer Spektrum.

Watkins, A. E., Bargagliotti, A., & Franklin, C. (2014). Simulation of the sampling distribution can mislead. *Journal of Statistics Education, 22(*3). Retrieved from http://www.amstat.org/publications/jse/v22n3/watkins.pdf

Watson, J. (1994). *Maths Works: Teaching and learning chance and data*. (assisted by W. Ransley) Adelaide: Australian Association of Mathematics Teachers. [Presenter's Handbook, 123 pp.; Participant's Handbook, 186 pp.; Overhead Transparency and Handout Masters, 158 pp.]

Watson, J. M. (1996). Reflections on videoconferencing and hypertext as media for professional development. In R. Zevenbergen (Ed.), *Mathematics education in changing times: Reactive or proactive* (pp. 165-176). Melbourne, VIC: Mathematics Education Lecturers' Association.

Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). Amsterdam: IOS Press and The International Statistical Institute.

Watson, J. M. (1998a). Professional development of teachers using CD-ROM technology. In L. Pereira-Mendoza (Ed.), *Statistical education – Expanding the network* (Proceedings of the 5th International Conference on Teaching Statistics, Vol. 2, pp. 921-927). Voorburg: International Statistical Institute.

Watson, J. M. (1998b). The role of statistical literacy in decisions about risk: Where to start. *For the Learning of Mathematics*, *18*(3), 25-27.

Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education*, *4*, 305-337.

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

Watson, J. M. (2014). *TinkerPlots* as an interactive tool for learning about resampling. In T. Wassong, D. Frischemeier, P.R. Fischer, R. Hochmuth, & P. Bender (Eds.), *Using tools for learning mathematics and statistics* (pp. 421-436). Heidelberg: Springer Spektrum.

Watson, J. M. (2016). Linking science and statistics: Curriculum expectations in three countries. *International Journal of Science and Mathematics Education*. DOI 10.1007/s10763-015-9673-y

Watson, J. M., Baxter, J. P., Olssen, K. H., & Lovitt, C. (1996). Professional development at distance as an information technology enterprise. *Asia Pacific Journal of Teacher Education*, *24*, 139-146.

Watson, J., Beswick, K., Brown, N., Callingham, R., Muir, T., & Wright, S. (2011). *Digging into Australian data with TinkerPlots*. Melbourne: Objective Learning Materials.

Watson, J., & Callingham, R. (2013). Likelihood and sample size: The understandings of students and their teachers. *Journal of Mathematical Behavior*, *32*, 660-672.

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*, 254-284.

Watson, J., & Callingham, R. (2015). Getting out of bed: Students' beliefs. In M. Marshman, V. Geiger, & A. Bennison (Eds.), *Mathematics education in the margins* (Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia, pp. 619-626). Adelaide, SA: MERGA.

Watson, J., Callingham, R., & Nathan, E. (2009). Probing teachers' pedagogical content knowledge in statistics: "How will Tom get to school tomorrow?" In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides* (Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia, Wellington, NZ, Vol. 2, pp. 563-570), Palmerston North, NZ: MERGA.

Watson, J., & English, L. (2015). Introducing the practice of statistics: Are we environmentally friendly? *Mathematics Education Research Journal*, *27*, 585-613. DOI 10.1007/s13394-015-0153-z

Watson, J., & Fitzallen, N. (2016). Statistical software and mathematics education: Affordances for learning. In L. English & D. Kirshner (Eds.), *Handbook of International Research in Mathematics Education* (3rd ed.) (pp. 563-594). New York: Taylor and Francis.

Watson, J. M., & Moritz, J. B. (1997). The C&D PD CD: Professional development in chance and data in the technological age. In N. Scott & H. Hollingsworth (Eds.), *Mathematics creating the future* (Proceedings of the 16th Biennial Conference of the Australian Association of Mathematics Teachers Inc., pp. 442-450). Adelaide, SA: AAMT.

Watson, J. M., & Moritz, J. B. (2003). Fairness of dice: A longitudinal study of students' beliefs and strategies for making judgments. *Journal for Research in Mathematics Education*, *34*, 270-304.

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

Williams, A. S. (2012). Statistics anxiety and worry: The roles of worry beliefs, negative problem orientation, and cognitive avoidance. *Statistics Education Research Journal, 14*(2), 53-75. Retrieved from http://iase-web.org/documents/SERJ/SERJ14(2)_Williams.pdf

Wood, M. (2005). The role of simulation approaches in statistics. *Journal of Statistics Education, 13*(3), available at www.amstat.org/publications/jse/v13n3/wood.html

World Class Arena. (2015). *World Class tests.* UK: World Class Arena Ltd. Retrieved from http://www.worldclassarena.org/en/about/aboutWCT.htm

Zieffler, A., & Fry, E. (Eds.). (2015). *Reasoning about uncertainty: Learning and teaching informal inferential reasoning*. Minneapolis, MN: Catalyst Press.

Zieffler, A., Park, J., Garfield, J., delMas, R., & Bjornsdottir, A. (2012). The Statistics Teaching Inventory: A survey on statistics teacher's classroom practices and beliefs. *Journal of Statistics Education*, *20*(1). Retrieved from http://www.amstat.org/publications/jse/v20n1/zieffler.pdf

Ziegler, L. (2014). Reconceptualizing statistical literacy: Developing an assessment for the modern introductory statistics course. In K. Makar, B. deSousa, & R. Gould (Eds.), *Sustainability in statistics education* (Proceedings of the 9th International Conference on the Teaching of Statistics, Flagstaff, Arizona, July 13-18). Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_7F1_ZIEGLER.pdf