The challenges of teaching probability in school

Graham A. Jones

It is now more than 15 years since probability and statistics became a main stream strand across the curriculum commencing in the early elementary years (e.g., Australian Education Council [AEC], 1991; Department of Education and Science and the Welsh Office [DES], 1989; National Council of Teachers of Mathematics [NCTM], 1989). Although research and classroom experience during the ensuing period have produced a prodigious knowledge base on the teaching and learning of probability, there are still significant challenges regarding what probability should be taught and how it should be taught in order to foster understanding.

The continuing challenge of what to teach

Prior to probability and statistics becoming a mainstream strand, probability was only taught, if at all, at the high school level. The high school curriculum in probability was part of the mathematics program and largely comprised "balls in urns problems" that had a major emphasis on counting and combinatorics. As Scheaffer, Watkins, and Landwehr (1998) remarked in retrospect, "Counting is beautiful and useful mathematics, but it is not probability" (p. 17). In addition to these veiled combinatorics problems there was often some treatment of the binomial and normal distributions and ipso facto of statistical inference.

Mainstream probability: a beginning

With the advent of the aforementioned curriculum reforms in the late eighties and early nineties, probability not only commenced in the early elementary school, it focused on key themes that were intended to spiral across the grade levels. In analyzing emerging probability curricula of this period (AEC, 1991; DES, 1989; NCTM, 1989), Jones, Langrall and Mooney (2007) identified key themes or big ideas that percolated through the elementary, middle, and high school curricula. In the **elementary school** curriculum, key ideas included the exploration and description of *random phenomena* arising socially and from random generators, and the *representation and ordering of random outcomes*. This ordering process involved both experimental estimation and symmetry-based measures of probability, albeit, informal ones. In the **middle school**, the key ideas of the elementary school curriculum were treated more formally with more precise representations of the *sample space* (both one- and two-stage experiments), *empirical estimations and theoretical measures of the probabilities of events*, and some consideration of *compound events and independent events*. At the **high school** level, the measurement of probability was extended to *random variables*, *discrete and continuous distributions* (e.g., binomial, normal and geometric), and the use of these distributions in *modeling random phenomena*. Probability distributions were also used to support work in statistical inference.

Along with the identification of these key themes in content, educators at the cutting edge of curriculum development invariably highlighted the accompanying need for language and concept building (e.g., chance, certainty, odds, and independence) and, for the first time, advocated a highly visible emphasis on experimentation and simulation across all grade levels. The emphasis on experimentation, advocated strongly in the writing of Steinbring (e.g., 1991), highlighted the importance of empirical estimation of probability (frequentist orientation) and its connection with theoretical probability (classical orientation). It also made possible the simulation and construction of various kinds of distributions (e.g., Burrill & Romberg, 1998; Scheaffer et al., 1998). This experimental orientation required technological support and was increasingly girded by graphics calculator and computer technologies that provided short- and long-run experimental data. Finally there was a recognition that students of all ages would bring to the classroom misconceptions in probability (e.g., Konold, 1989, Shaughnessy, 1992, Tversky & Kahnemann, 1974), such as *representativeness*, *availability*, and the *outcome approach*, and that the content of these misconceptions should be confronted through experimental activity and reflection (NCTM, 1989, p. 110).

Mainstream probability: further refinements
By the arrival of the new millennium, the knowledge base in probability learning had burgeoned as a consequence of research undertaken during the 10-year period since probability and statistics became a mainstream strand (see Jones, 2005; Jones et al., 2007). This emerging research on students' understanding of probability highlighted the need for greater curriculum emphasis on fundamental elements like variation and randomization.

Variation is the *sine qua non* of both statistics and probability. In statistics, variation refers to the differences exhibited by a sample of data values around some centre such as a mean or median. For example, even though a group of children all belong to the same grade level, their heights will vary around an average. More pertinently, if we weigh ourselves daily for a week, there will be small variations in mass about the weekly mean. With respect to probability, variation is a key element of randomization in that outcomes vary about some expected value.

Recent research (Shaughnessy & Ciancetta, 2002; Watson & Kelly, 2003) has revealed a great deal of instability and misunderstanding in students' conceptions of variation in both statistical and probabilistic environments. In responding to this research, Shaughnessy (2003, 2006) has suggested that the study of variation should be a prime curriculum element in launching probability learning. As a corollary to this curriculum direction, there is also a growing tendency to have students explore variation in statistical environments before studying variation in probabilistic environments. In essence, the study of statistical variation is seen to be pivotal in providing a natural impetus for the study of randomization; however, as the next paragraph reveals, randomization incorporates a unique type of variation.

Although the process of randomization had been given lip service in earlier curriculum documents, contemporary research (e.g., Batanero & Serrano, 1999) has revealed that high school students, even after instruction in probability, demonstrated serious limitations in their understanding of randomization. These authors noted that "students overemphasized unpredictability and luck to justify their attributions of randomness and this tendency seemed greater in older students…" (p. 565). From a curriculum perspective, Batanero and Serrano emphasized the need for randomness to be considered in terms of its multiple properties: local variability, long-term stability, the length and frequency of runs, and the proportion of alternations. Other researchers (e.g., Amir & Williams, 1999; Watson & Moritz, 2003) have found similar issues vis-a-vis students' thinking about luck, fairness and bias.

Accordingly, in current thinking about the probability curriculum, variation and randomization are perceived as the launching and anchoring elements; in essence, they set the scene for key elements such as the representation of the sample space, empirical estimation and theoretical measures of probability, and probability distributions. Several researchers (Gal, 2005, Watson, 2006, Watson & Callingham, 2003) have moved the content of the contemporary probability curriculum even further through their discourses on probability literacy and statistical literacy. In referring to the *knowledge* *elements* of his probability literacy framework, Gal (p. 46) not only includes familiar elements such as *big ideas* (variation, randomness, independence, predictability and uncertainty), *figuring probabilities* (finding or estimating the probability of an event), and *language* (the terms and methods used to communicate about chance), he also introduces ideas like context and critical questions. Context deals with "the role of probabilistic processes and communications in the world" (p. 52), and critical questions focus on the need for people to become familiar with the flaws, problems, and biases in probabilistic information. An example of a critical question in probability is presented below:

> Sam buys a Gold Lotto ticket every week. He never selects the first six numbers (1, 2, 3, 4, 5, and 6) because he believes his chances are better by selecting a number from each line of the entry form. What do you think about Sam's reasoning?

Watson (1997, 2006) recommends that children should examine critical questions beginning in the primary grades. She also underlines an important curriculum link between critical questions and context, in that critical questions can be readily extracted from media, financial, and scientific reports.

In summary, the issue of what to teach in probability has undergone several changes in the more than 15 years since probability became part of a mainstream curriculum strand. These changes have been subtle rather than spectacular but they have forged, from the early grades, stronger links between statistics and probability through key concepts like variation. The on-going changes have also produced a stronger focus on experimental probability and its connection with theoretical probability. Moreover, as a consequence of these changes, the powerful and historical connection between probability and statistical inference has been made more transparent (see Pfannkuch, 2005).

The continuing challenge of how to teach

Even though probability was not a regular part of elementary and middle school curricula until the late 1980s, early research on children's probabilistic thinking did draw implications for the teaching of probability. Most noteworthy was the research of Efraim Fischbein (e.g., Fischbein, 1975, Fischbein & Gazit, 1984). In these studies Fischbein attempted to use experimental activities, involving probability generators, to transform children's *primary intuitions* (cognitive beliefs derived only from experience) into restructured cognitive beliefs or *secondary intuitions*. This research set the scene for a more comprehensive instructional theory generated by Steinbring (1991) and for further research on instructional environments in the last 15 years (see Jones, 2005).

Instructional theory in probability: a beginning

Utilizing an epistemological analysis of the nature of stochastics (statistics and probability), Steinbring (1991) examined probability from both its frequentist (empirical) and classical (theoretical) forms. He claimed that neither the empirical situation nor the theoretical model can act alone to express the meaning of probability; rather, there is a need to develop these two dependent ideas in unison. For Steinbring, learning involved the following sequence: (a) personal judgments or predictions about the random phenomenon in question; (b) comparisons between the empirical data and various conjectured theoretical models; and (c) the creation of generalizations and more precise characterizations of the random phenomenon, based on an evaluation of the comparisons described in (b). As an example of Steinbring's instructional process, let us investigate, as a random process, the first service performance of world number one tennis player, Roger Federer. A possible instructional sequence might go as follows: make some personal judgments about the percentage of Federer first serves that go in; collect actual data on Federer's first serve and make comparisons between the first-serve data and probability models of his first service (e.g., binomial, geometric); and finally, produce a more precise description of the key statistical features of his first service that is based on one or more of these models.

Computer environments
Computers became an important tool in supporting Steinbring's instructional theory in that they provided powerful classroom tools for creating more realistic simulations and for generating data. The advent of computer environments also enabled teachers and students to deal with more realistic probability problems, ones where simulation provided a genuine alternative to complex computations (Biehler, 1991).

While the initial development of computer environments and simulations was directed largely at forging the gap between frequentist (empirical) and classical (theoretical) probability, recent research with computer microworlds (Pratt, 2000, Stohl, 1999-2002) has resulted in new pedagogical advances in probability. For example, working from the assumption that students hold multiple, even competing, intuitions, Pratt's *Chance-Maker* microworld enabled students to engage in a task that refined their understanding of randomness by enabling them to generate and evaluate the long-term behavior of gadgets (probability generators that sometimes required fixing). The microworld helped them to focus their evaluation of random behavior by examining the sample space and data distribution of the gadgets and this in turn supported and enriched their understanding of randomness. Similar theoretical assumptions characterize Stohl's (1999-2002) *Probability Explorer* microworld.

From an instructional perspective, Pratt (2005) recommended that teachers need to produce probabilistic tasks that are linked to teaching objectives and are also rewarding for children. In particular, he advocated the design of tasks that incorporate *purpose* (have a meaningful outcome for the student) and *utility* (enable children to appreciate the applicability of the idea). Finally, he claimed that these tasks should involve the student in testing personal conjectures, performing large-scale experiments, and in experiencing the systematic variation of the context. These latter characteristics, although geared to a computer environment, have a salient link with Steinbring's (1991) instructional sequence.

Cognitive frameworks and instructional theory
Fischbein (1975), Steinbring (1991) and Pratt (2005) all highlighted the need for instruction to build on students' existing notions of probability, whether they be immature intuitions or more formal understandings. Given the importance of this link between instruction and students' probabilistic thinking, recent research on cognitive frameworks (e.g., Jones, Langrall, Thornton, & Mogill, 1997; Polaki, 2005; Tarr & Jones, 1997, Watson, Collis, & Moritz, 1997) adds a further dimension to the development of instructional theory. These frameworks (see Fig. 1) characterize and describe students' probabilistic thinking according to various levels of cognitive maturity and across various probabilistic concepts such as sample space, probability of an event, conditional probability, and independence. As such the frameworks provide a coherent overview of the kinds of probabilistic reasoning that students can be expected to bring to the classroom. They also equip teachers with domain specific knowledge that can be used in the design, implementation, and assessment of instruction in probability.

	LEVEL 1 (Subjective)	**LEVEL 2** (Transitional)	**LEVEL 3** (Informal quantitative)	**LEVEL 4** (Numerical)
INDEPENDENCE	• Predisposition to consider that consecutive events are always related. • Pervasive belief that they can control the outcome of an event. • Uses subjective reasoning which precludes any meaningful focus on independence. • Exhibits unwarranted confidence in predicting successive outcomes.	• Shows some recognition as to whether consecutive events are related or unrelated. • Frequently uses a "representativeness" strategy, either a positive or negative recency orientation. • May also revert to subjective reasoning.	• Recognizes when the outcome of the first event does or does not influence the outcome of the second event. In replacement situations, sees the sample space as restored. • Can differentiate, albeit imprecisely, independent and dependent events in "with" and "without" replacement situations. • Some reversion to representativeness.	• Distinguishes dependent and independent events in replacement and non-replacement situations, using numerical probabilities to justify their reasoning. • Observes outcomes of successive trials but rejects a representativeness strategy. • Reluctance or refusal to predict outcomes when events are equally-likely.

Figure 1. An extract of a framework describing middle-school students' thinking in independence (Tarr & Jones, 1997)

In terms of the *design of instruction*, cognitive frameworks highlight the diversity of students' probabilistic reasoning and they assist teachers in the formulation of learning goals, in the development of learning activities and in conjecturing how the learning process might go. In essence, they enable teachers to plan meaningful instructional sequences similar to those advocated by Steinbring (1991) and Pratt (2005). With respect to *implementation of instruction*, cognitive frameworks can act as a filter for analyzing students' oral and written responses and for building questions that will stimulate students to acquire more mature probabilistic thinking. Finally cognitive frameworks can assist teachers in assessing students' performance over time and in evaluating the effectiveness of their own instruction.

Instruction that connects with statistics

In discussing what aspects of probability should be taught, I made reference to recent developments that encourage a stronger link between probability and statistics, such as those involving the concept of variation. Shaughnessy (2003) took this link even further by advocating an instructional process that introduces the study of probability through data.

In explicating this data-based instruction, Shaughnessy (2003) built on the work of Gigerenzer (1994) who suggested that problems should be framed in terms of frequencies rather than probabilities. The problem below provides an illustration of how this might occur.

> The price of regular unleaded gas in Australia was below $1.20 on only 15 of the last 100 days. How many days in the next 20 would you expect gasoline to fall below $1.20?

This data-inspired use of frequencies is seen to provide a more contextual and transparent approach for students, at least in the initial stages of probability learning. Shaughnessy wrote, "The importance of probability questions in the context of real data supercedes past approaches to probability that started with counting problems and games of chance" (p. 224). In furthering instructional theory, Shaughnessy has also highlighted the importance of maintaining an ongoing connection between statistics and probability and he especially emphasized the connection between the concept of variation in statistics and the concept of sample space in probability. On the one hand, students need to be able to identify the set of possible outcomes in a probability experiment; on the other hand they need to realize that while the outcomes vary they occur predominantly in a "likely range" (p. 224). The connection between variation and sample space needs to be re-examined regularly: in diverse contexts and in various aspects of probability content culminating in the powerful bond between probability distribution and confidence intervals.

Like the issue of what to teach, the challenge of how to teach probability has also undergone substantial change during the past 15 years, with the main thrust being on the development of instructional theory that unifies statistics and probability. In this kind of instruction *data* are used to presage probability through frequency, to link sample space and variation, to connect empirical and theoretical probability, and to build notions of probability distributions. Such an instructional theory also recognizes the importance of learning environments, both computer and otherwise, that utilize cognitive frameworks of students' probabilistic thinking to inform instruction that incorporates both purpose and utility.

Conclusion

The teaching of probability is in its infancy, especially as part of a mainstream curriculum strand that traverses elementary, middle-school, and high-school education. In spite of the fact, that it was mathematically desirable for statistics and probability to be unified in a single strand, early curriculum documents (AEC, 1991; DSE, 1989; NCTM, 1989), especially at the elementary and middle-school levels, generally treated the two areas as separate topics. Research and instructional experiences over the past 15 years have resulted in a concerted effort to link statistics and probability in addressing the challenges of "what to teach" and "how to teach". Given this stronger connection, it is appropriate in a paper like this to advocate more universal usage of the term "stochastics". In my mind the use of stochastics would provide a more valid curriculum indicator of where the *probability and statistics strand* is headed and would also send a robust signal to researchers interested in investigating the learning and teaching of probability and statistics.

References

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

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Carlton, VIC, Australia: Curriculum Corporation.

Batanero, C., & Serrano, L. (1999). The meaning of randomness for secondary students. *Journal for Research in Mathematics Education*, 30, 558-567.

Biehler, R. (1991). Computers in probability education. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 169-211). Amsterdam: Kluwer.

Burrill, G., & Romberg, T.A. (1998). Statistics and probability for the middle grades: examples from mathematics in context. In S.P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K-12* (pp. 33-62). Mahwah, NJ: Lawrence Erlbaum.

Department of Education and Science and the Welsh Office. (1989). *National curriculum: Mathematics for ages 5 to 16*. York, UK: Central Office of Information.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children.* Dordrecht, The Netherlands: Reidel.

Fischbein, E. & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics*, 15, 1-24.

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. 39-64). New York: Springer.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129-161). Chichester, England: Wiley.

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

Jones, G.A., Langrall, C.W., & Mooney, E.S. (2007). Research in probability: responding to classroom realities. In F.K. Lester Jr. (Ed.), *The second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 909-955). Charlotte, NC: Information Age Publishing.

Jones, G.A., Langrall, C.W., Thornton, C.A., & Mogill, A.T. (1997). Framework for assessing and nurturing children’s thinking in probability. *Educational Studies in Mathematics*, 32, 101-125.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6, 59-98.

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

Pfannkuch, M. (2005). Probability and statistical inference: how can teachers enable learners to make the connection? In G.A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 267-294). New York: Springer.

Polaki, M.V. (2005). Dealing with compound events. In G.A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 191-214). New York: Springer.

Pratt, D. (2000). Making sense of the total of two dice. *Journal for Research in Mathematics Education*, 31, 602-625.

Pratt, D. (2005). How do teachers foster understanding of probability? In G.A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 171-190). New York: Springer.

Scheaffer, R.L., Watkins, A.E., & Landwehr, J.M. (1998). What every high-school graduate should know about statistics. In S.P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in Grades K-12* (pp. 3-32). Mahwah, NJ: Lawrence Erlbaum.

Shaughnessy, J.M. (1992). Research in probability and statistics. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

Shaughnessy, J.M. (2003). Research on students’ understanding of probability. In J. Kilpatrick, W.G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 216-226). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J.M. (2006). Student work and student thinking: an invaluable Resource for teaching and research. In A. Rossman & B. Chance, (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics* (On CD). Salvador, Bahia, Brazil: International Statistical Institute.

Shaughnessy, J.M., & Ciancetta, M. (2002). Students’ understanding of variability in a probability environment. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics, Cape Town* [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Steinbring, H. (1991). The theoretical nature of probability in the classroom. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 135-168). Dordrecht, The Netherlands: Kluwer.

Stohl, H. (1999-2002). *Probability Explorer* [Computer Software]. Software application distributed by the author. Retrieved, February 8, 2005, from http://www. Probexplorer.com.

Tarr, J.E., & Jones, G.A. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal*, 9, 39-59.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124-1131.

Watson, J.M. (1997). Assessing statistical literacy through the use of media surveys. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistical education* (pp. 107-121). Amsterdam: IOS Press.

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

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

Watson, J.M., Collis, K.F., & Moritz, J.B. (1997). The development of chance measurement. *Mathematics Education Research Journal*, 9, 60-82.

Watson, J.M., & Kelly, B.A. (2003). Developing intuitions about variation: the weather. In C. Lee (Ed.), *Proceedings of the third International Research Forum on Statistical Reasoning, Thinking, and Literacy* [CD-ROM]. Mt Pleasant, MI: Central Michigan University.

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.

Graham A. Jones
Emeritus Professor, Griffith University
School of Education and the Arts
Gold Coast Campus
PMB 50 Gold Coast Mail Centre
Queensland 9726, Australia
g.jones@griffith.edu.au, graham.a.g.jones@gmail.com