Variation and Expectation as Foundations for the Chance and Data Curriculum

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

This paper considers the evolution of research in statistics education since the introduction of chance and data into the Australian mathematics curriculum in 1991 and presents selected outcomes of research into students' understanding of the content in the chance and data curriculum, using them to argue for a change in emphasis in the classroom in the teaching of chance and data. These suggestions might also influence current curriculum revisions taking place within Australia and New Zealand. Building on the history of the discipline of statistics and its introduction into the school curriculum, it is argued that topics in the curriculum associated with expectation, such as the mean, generally have preceded those associated with variation, such as the standard deviation. Research however, suggests that children develop an appreciation of variation before expectation, and this knowledge should influence the order of the introduction of associated topics and their juxtaposition in the curriculum and the classroom.

The history of the discipline of statistics is shorter than that of most of the other mathematical sciences, particularly those included in the school curriculum. Emerging from concerns with chance mechanisms in the 18th century (Hacking, 1990) and the social concerns of the 19th century (e.g., Bernstein, 1996), the 20th century saw the maturing of the theoretical and practical areas of the field to produce a sophisticated discipline that contributed greatly to many other fields of endeavour (Salsburg, 2001). As with many other areas of study, especially science, which came to prominence in the 20th century, the question arose as to how much foundational understanding of statistics was required for non-specialist citizens and where this should be placed in the school curriculum.

By the time the field had matured, the basics could be expressed in relation to measures of central tendency and spread, mainly the mean and standard deviation. In considering these two measures in terms of the associated number skills being taught in the school curriculum, the mean required addition and division whereas the standard deviation required the mean itself, as well as differences, squares, addition, and the square root. In terms of the complexity of these operations the mean could be taught in Grade 5 or 6, whereas the standard deviation was left until much later, at least Grade 10, if not Grade 11 or 12. Unfortunately the result of such a focus in the curriculum was that the idea of central tendency – *expectation* – was introduced but spread – *variation* – was not emphasized. The idea of spread in a distribution of data values was implicit in the need to calculate an average or central value but this was rarely made explicit.

As Shaughnessy (1997) pointed out, the neglect of ideas of spread in school mathematics curricula is likely to be associated with the mathematical complexity of calculating the formula for the standard deviation. One suspects that a mentality has long existed in the mathematics curriculum that if there is not a calculation to perform in relation to a concept, then the concept is not worth including in the curriculum. Considering the structure of the curriculum suggested by Holmes (1980) – data collection, data tabulation and representation, data reduction, probability, and interpretation and inference – it was the areas of data reduction and probability that received initial attention in the mathematics curriculum. The arithmetic mean appeared as early as Capel's (1885)

Catch Questions in Arithmetic & Mensuration and the calculation of probabilities was common in algebra books from the mid-twentieth century (e.g., Hart, 1953). The research of Mokros and Russell (1995) on school students' understanding of average reflected these observations, indicating that students appreciated various ideas of middle but rarely displayed an understanding of an average value in some sense representing the set of data values from which it was derived. When the curriculum was expanded to take in Holmes' ideas (National Council of Teachers of Mathematics, 1989; Australian Education Council [AEC], 1991; Ministry of Education [MoE], 1992), it was natural that teachers continued to focus on the aspects that were familiar to them from their own previous experiences in mathematics. Concepts associated with data collection and sampling, data tabulation and representation, and interpretation and inference, were much more descriptive and nebulous, often not leading to a single correct answer.

The Rise of Variation

From a theoretical viewpoint the averages of data reduction and the elementary probabilities of chance are examples of the mathematical idea of expectation. Focusing as they do on specific values they draw attention away from the data sets or distributions that create them. As David Moore (1990) pointed out, however, in his seminal work entitled "Uncertainty," it is the omnipresence of variability in the world that creates the need for statistics, otherwise there would be no discipline. Wild and Pfannkuch (1999) supported this view based on their interviews with practising statisticians, where they found that variation was consistently singled out as the significant element of all investigations. Moore also acknowledged the importance of expectation in his definition of random when he juxtaposed the expected pattern of outcomes with uncertain individual outcomes. The recent contribution of Konold and Pollatsek (2002) in using the metaphor of "data analysis as the search for signals in noisy processes" (p. 259) is significant in providing another angle on the issue and stressing the pre-eminence of variation. It is this tension between variation and expectation that makes the study of probability and statistics and their applications so interesting.

In 1993, Green, in considering the new data handling curricula around the world, began to ask specific questions about students' understanding of variation.

- What do students understand of variability and how does this originate?
- What are the implications for not introducing the concept of spread in the education of young children?
- What are the essential experiences needed to develop a full appreciation of variability? (pp. 227-228)

With the impetus of the calls from Green and from Shaughnessy (1997), several projects (e.g., Shaughnessy, 2002; Watson, 2000) began to focus more directly on school students' understanding of variation in relation to the overall chance and data curriculum, in an attempt to answer these questions. Although some findings from Australian research have been reported elsewhere, this paper focuses on the beginning understanding of students and development into the middle years in order to make suggestions for the classroom and curriculum in terms of both variation and expectation.

Method

The data for this paper are extracts from student interviews conducted with seven six-year-old children and with many older children, mainly from Grades 3 to 9 but a few in the senior grades. The six-year-olds had participated in a rich mathematics program but had had no specific work related to chance and data (Watson & Kelly, 2002). The older students had experienced the programs in chance and data their schools provided but no instruction related to the research project at the time the interviews took place. The protocols were developed for children in Grade 3 and above. The topics of interest were related to drawing 10 lollies from a container of 100, of which 50 were red, 30 green, and 20 yellow, and predicting how many would be red (Kelly & Watson, 2002), to explaining the meaning of an average daily maximum temperature for their city for a year of 17^{o}C (Watson & Kelly, in press), and to creating a pictograph with cards depicting books and children to show how many books each child had read and to make predictions based on the display (Watson & Moritz, 2001). Details of data collection are in the aforementioned papers.

Young Children's Appreciation of Variation and Expectation

Evidence from six-year-old students suggests that they develop an intuitive appreciation of variation in the world about them before they develop intuitions about expectation. With respect to drawing 10 lollies from the container with 50 red, 30 green, and 20 yellow, the six-year-olds interviewed made reasonable estimates but did not have the experience to provide reasons based on the proportions in the container. This is illustrated in the following extract.

> (S1) [I: How many reds do you think you will get if you pull 10 out?] Umm, probably umm something like 5 or more. [I: Why do you think that?] Because when I put my hand in umm, some reds are on the top and some reds are on the bottom and I would get the top reds and there's sort of, something like 10 on the top and I'd be doing it in a corner or something and there'd be like 5 reds in one corner. [I: Now if we did this several times do you think you would get the same number of reds every time?] No. [I: Why don't you think you'd get the same number every time?] Because every time I mix them up some would be on the bottom and then others would be on the top because there could be less on the top. [I: Now I want to ask if we drew the 10, what number of reds would surprise you?] Something like 8. [I: Why?] Because – they are so little that they can slip out of your hands and I'd be surprised that I've got so many. [I: Suppose six of you do this experiment. What do you think is likely to occur for the numbers of red lollies that are written down?] 4, 5, 1, 3, 6, 8 [I: Any particular reason you chose those?] Umm, because they are little numbers and that they can – they could probably fit into my hand.

The comment "5 or more" acknowledged possible variation, as did not getting the same number every time and the six different suggested values; the supporting reasoning however was idiosyncratic. The closest justification to proportional reasoning was shown by a boy who thought a long time at several places, responding as follows.

> (S2) [I: How many red ones do you think you might get?] I think I would get … about 5. [I: Why do you think you might get 5?] … Because there's 50 so I think I might get 5, because there's 5 pl.. 10, so … [I: … Would you get 5 again?] [shakes head, no] [I: You might get something different?] [Nods head, yes] [I: Why?] Because every time you do something it's a different way. [I: How many would surprise you?] Umm I think 6, … because 6 is my favouritest number. [I: Suppose six of you do this experiment. What do you think is likely to occur for the numbers of red lollies that are written down?] 4, 6, 5, 3, 2, 8 I thought 4 and 6 because 4 and 6 would be 10 and … 5 and 3 … were 8, and so I thought I would put the 8 at the end, and the 2, I just thought that is my second best number.

At first this child showed potential for appreciation of both expectation ("there's 50 … so 5") and variation ("it's a different way") but then reverted to idiosyncratic reasoning with respect to the task. For the protocol on drawing lollies from a container, all children appreciated that different outcomes were likely on multiple draws but there was only one slight intuition about proportion (noted above).

For the protocol on the maximum daily temperatures in their city over a year, it was realised that six-year-old children would have little appreciation of the meaning of an average yearly maximum but it was of interest to explore their understanding of how temperatures vary and whether there would be any type of expectation displayed. When asked what the average maximum of 17^{o}C meant, most indicated that it was “hot” or “cold.” All were aware of temperature as a measurement and most had watched the nightly weather forecast on television. When asked if all days had a maximum of 17^{o}C, almost all disagreed, acknowledging variation. When asked to explain, several could distinguish summer and winter. The most articulate of these explanations was the following.

> (S3) [I: What does this tell us about the temperature?] That is quite hot if it was 17. [I: Do you think all of the days of the year had a temperature of 17^{o}?] No, because you get summer, winter – summer, spring, winter, autumn, then summer again. [I: What does that mean?] You get, it’s like hot … mild or cool, cold, mild or cool, and then hot again.

Although not all children were certain about whether all days had a maximum temperature of 17^{o}C, one suggested the other extreme, “Every single thing [day] is different.” Asked to predict temperatures for six different days of the year, the children suggested different distinct values, some of which were not reasonable (like 88). They did not return to seasons in explaining the six chosen values and were inconsistent in predicting monthly maxima for July and January in relation to each other and the entire year.

In the protocol on representing books children had read, ideas about variation were expected in the pictographs produced and subsequent descriptions of them, whereas ideas about expectation would be related to the ability to make predictions in association with the pictographs. All six-year-olds could count correctly but some produced piles of cards that could not be distinguished from each other, whereas others produced visual displays laid out in rows or columns. After the students had produced their representations of the books each child had read, they were asked what someone could tell who came into the room and looked at their displays. Other than telling imaginative stories, expressing differences, i.e., variation, in the numbers observed was the extent to which students progressed in describing their displays. When asked “Who likes reading the most?” they all replied the student with the greatest number of books beside her name. They hence could compare and contrast the numbers of books assigned, although some had to recount their piles of cards. The idea of expectation for this protocol was related to predicting the number of books read by Paul, a new boy introduced into the scenario, and by Helen, a new girl. The data provided to this stage of the interview had gaps for some numbers and the girls had read more books than the boys. The six-year-olds found the prediction task quite difficult. One boy refused to predict altogether – he did not like guessing. Other predictions included that Paul had read three books, “because one of my sisters is three [years old],” and that Paul had read zero “because it was his first time in the library and he doesn’t know how to choose books.” None of these students used the information available in their representations to express an expectation for Paul but one boy suggested Helen had read six books, “because no one else has got six.”

What is learned from six-year-olds? They are aware that variation occurs in the world around them: in numbers of red lollies drawn from a container, in the maximum daily temperature, and in the number of books different children have read. What of course they do not appreciate is appropriate variation around some expected value: the proportion of red lollies in the container, the yearly average maximum temperature, or the pattern shown in a representation they have produced. In fact in nearly all cases, they do not appear to be aware of the existence of an expected value.

The Development of Appreciation of Variation and Expectation

This section focuses mainly on student development through the middle school years. In relation to the predictions related to drawing lollies from a container, all children suggested variation in the six trials they were asked to predict. Improvement occurred across the grades in focusing the variation more appropriately both in terms of an expected value of five red and the total range of values. Often primary children suggested outcomes ranging from 0 to 10 red, that is all possible values, rather than a range of likely values, say 3 to 7, or 2 to 8. Justifications for the choices moved from mainly idiosyncratic ideas, to arguments based on there being "more reds" or "more reds than yellow or green" in the container, to arguments based on "half are red." These last responses initially had an intuitive focus on centre, such as "mostly around 5 and mostly reds," or "some might go higher, some might be lower, but half of them is red." Few students by Grade 7 or 9 had an appreciation of the appropriate distribution in terms of the degree of variation that would be reasonable about the middle value of 5. The following extract from a Grade 7 student illustrates the highest category of response for middle school students.

> (S4) [I: … and pull out a handful, how many red do you think you might get?] Five. [I: Why do you think you might get 5?] Because half of the contents of the container is red and so you should expect to get half the amount in what you pull out. [I: Suppose you did this a few times… would you expect to get the same number of reds every time?] No. Because it's just the luck of the draw most of the time. You'll get around the same amount but not exactly the same amount. [I: How many reds would surprise you?] I reckon about 8 or 9. [I: So why do you think 8 or 9?] …cause again there's only half the container filled. So you'd still expect to get some yellow and green in there, so you wouldn't expect just to pull out this huge handful of red ones, cause they'd all be mixed up. [I: Suppose 6 of you've come along and done this experiment… Can you write down for each of the people the number of red that would be likely?] 5, 3, 6, 4, 5, 4 [I: So why have you chosen these numbers?] I've chosen them because they're around the middle number that I chose of 5 and so there's a bit of give and take for different mixtures… cause obviously they'd mix them up after each go and you never know they might bring all of the other ones up to the top.

This student produced a graph for 40 imagined outcomes for drawing 10 lollies that reflected the appropriate shape of the distribution about 5 reds but the spread was somewhat greater than would be expected.

In considering the protocol on the average maximum daily temperature for a year, the sources of variation that were suggested reflected many different aspects: differences during the day itself, differences from values predicted on the weather forecast, differences between their city and other parts of the country or world, differences in the seasons, and short-term fluctuations over a few days. Not all of these responses were relevant to the questions asked but they illustrated the breadth of student intuitive appreciation of variation in the context.

> (S5) Well sometimes you can't always rely on the weather… because I can remember one day when I was down in Hobart, that it was freezing cold and it was supposed to be 17˚ … and well sometimes it's hard when you're sort of thinking about what the weather's going to be, knowing what to put on, when it can change later in the day. (Grade 3)

The most common explanation for variation in the temperature was seasonal change. In predicting six temperatures throughout the year, however, some younger students based their judgements solely on personal experience: "Well the maximum temperature which we have ever had was 32 … probably like at least around the twenties, 25 … 18, 21, 16, 26 [Why?] Basically because we had like temperatures around them" (Grade 5). Appreciation of appropriate spread grows over the middle years, as does the appreciation of clustering around the yearly average. The subtle development in understanding is shown in the following two Grade 7 responses.

(S6) [I: What does this tell us about the temperature in Hobart?] It's not really high, like up in Darwin but it's not absolutely freezing like in Antarctica or somewhere. [I: So do you think all days might have a maximum of 17˚C?] No. Because some days you would get like a day that might go to 30˚C, if it is really hot, and a lot of days could get much colder. […] Six Temperatures: 12, 23, 17, 19, 14, 20 [I: Why?] Because like, it could be anything basically – it depends, but the average is 17, so it would be more likely to be within a certain range, but up like 40 or down to zero. [I: … So what have you done there (see Figure 1)?] It's the highest in January, February, and December cause that's the middle of summer... The coldest would be around here in winter. In around these sections, it's around middling. [I: It's interesting you've got May a little bit higher here…] Yeah, it could change. There'd be a lucky day sometimes. It could just go up over. [I: So are these temperatures, are they what, maximums, or averages or…?] Yeah, maximum averages.

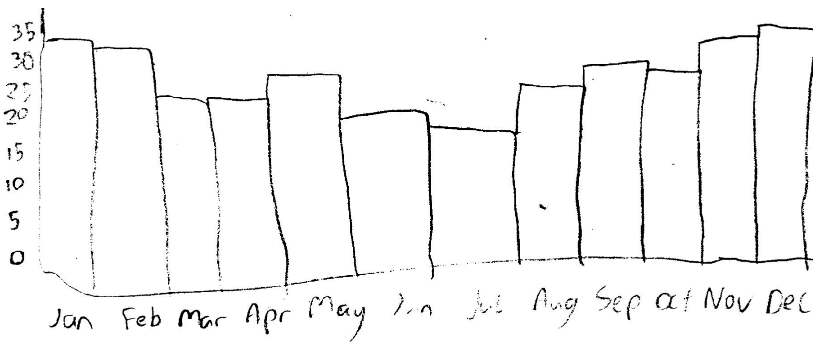

Figure 1. Grade 7 graph of average maximum daily temperature over a year.

In contrast to the previous student, the following one is more focussed on the distribution of temperatures directly related to the questions and indicates an appreciation of two types of variation, daily fluctuation and seasonal trends.

(S4) It is cold. Either you get – in Hobart obviously it means that there's a lot of cold days but then there's a few hot days in there but the number of cold days is outnumbering the number of hot days bringing the total down. [I: So do you think all days have a maximum of 17?] No, I think they could be hotter – some of them – most of them – a fair few of them might have been higher but then you have got all these ones that are really low. Like dismal. [I: What would the maximum be for 6 different days of the year?] So is this like for a set period of time, like during a week, or can is be like one during winter, one during summer…? [I: Do what you want, just go for it.] Six Temperatures: 19, 29, 15, 11, 35, 31. [I: Why these?] I just, I made the choices … to give a wide range of the possibilities because quite often you have a very cold day but then of course you have very hot days and so the rest are just spread out through the middle to show that they are through the middle and all different, you can get all different temperatures no matter what.

In responding to the protocol to represent the numbers of books children had read, primary students had no difficulty in arranging layouts of the book to show the variation among children. A few older students went further to rearrange their pictographs when given extra information.

(S7) I would probably also change it around so it went from the person who read the least books … and then slowly getting bigger. It would be a bit more easier to look at and you could follow it down logically … it's quite clear that Jane and Anne both read 4 books. (Grade 12)

The older students were more likely to be aware of the usefulness of the distribution of the values to tell the story of the number of books read. They were also more likely to be speculative in responding to questions such as who likes reading the most: "Well it could be Lisa [who had the most] but it doesn't really mean that she *likes* it. She might be forced to" (Grade 7). Considering the expectation of the number of books read by Paul or Helen, the new children in the protocol, responses ranged from "I couldn't tell because I haven't been given the information" (Grade 5) to "3, because everyone has read 1 to 6 and 3 is in the middle" (Grade 7). The latter response considered the range of values in the pictograph but some students appeared to ignore the pictograph and jump straight to an algorithm: "[No value given] Just the average of these. [I: How?] … You just add them up, then divide it by the total number of people" (Grade 10). Another student who suggested using averages also considered the variation shown for boys and girls in the pictograph: "…with Paul you could take an average … because boys don't seem to read as many as girls. So I reckon Helen would probably read more than Paul" (Grade 10). It could be argued that with respect to this protocol, students' first intuitions involve variation, later they focus on "most," and after being taught the mean algorithm this overrides intuitions and they perform calculations. Finally, however, some put the ideas of variation and expectation together to provide balanced conclusions.

Returning to Green's (1993) first question about the origins of understanding of variability by students, what is seen in the development of student understanding of variation and expectation documented in interviews from early childhood to the middle years is the following progression:

- intuitive appreciation of variation as change or non-uniformity of outcome;
- intuitive appreciation of expectation without the ability to associate it with mathematical theory (proportional reasoning or averages);
- developing appreciation of appropriate variation in straightforward contexts (random generators and common physical phenomena such as height or weather);
- ability to use qualitative terms to express expectation, such as "more" or "most";
- slow development of ability to apply proportional reasoning to quantify expectation;
- eventually the ability to tie together expectation and variation as integrated notions.

These observations from interviews support the hypothesis on the development of the concept of variation put forward by Watson, Kelly, Callingham, and Shaughnessy (2003). They proposed four levels of development: prerequisites for variation, partial recognition of variation, applications of variation, and critical aspects of variation. Throughout their model, increasing ability to handle mathematical expectation was observed, with proportional reasoning only appearing at the highest level.

Practical Implications for the Curriculum and Classroom

These outcomes and Green's (1993) second question about the implications of not introducing variation to young children demand a closer look at current curriculum documents and at the issue of whether variation and expectation are deeply embedded in them. The terms variation and expectation are not found in the Chance and Data section of *A National Statement on Mathematics for Australian Schools* (AEC, 1991). Band B (around Grades 3 to 6) Data Handling experiences in Australia, for example, focus on questions, representations, and sampling, with little hint of variation involved in any of them. The term variation appears once in the organisers in the Australian *Profile* (AEC, 1994), in relation to Chance: Understanding, estimating and measuring chance variation. This is striking because *nowhere* at any level in the detailed description of this organiser is the word used again. In *Mathematics in the New Zealand Curriculum* (MoE, 1992) the Statistics section also avoids the word expectation but variation and associated ideas are used meaningfully from Level 5 (around Grades 9 and 10). Other words associated with mathematical expectation and variation in terms of chance and data handling occur sporadically from the early bands or levels in both documents. A few examples are given in Table 1. These could be linked more directly to the underlying concepts.

Table 1.

Examples from Australian (AEC, 1991) and New Zealand (MoE, 1992) Curricula

Experiences with chance should be provided which enable children to:	**Experiences with data handling should be provided which enable children to:**
Describe possible outcomes for familiar random events and one-stage experiments (AEC, p. 166)	Recognise that repetitions of the same experiment may produce different results (AEC, p. 176)
For random events, systematically list possible outcomes, deduce the order of probability outcomes and test predictions experimentally (AEC, p. 170)	Use and interpret interquartile range and compare with range (AEC, p. 178)
Determine the theoretical probabilities of the outcomes of an event such as the rolling of a die or drawing a card from a deck (MoE, p. 188)	Use their own language to talk about the distinctive features, such as outliers and clusters, in their own and others' data displays (MoE, p. 178)

The terms expectation and variation have sophisticated mathematical definitions in relation to distributions of variables (James & James, 1959, p. 151). This may be one reason why they rarely appear in curriculum documents in relation to chance and data at the school level. Both words however have colloquial meanings that can encompass ideas associated with pattern and prediction or with slight change and spread. These ideas can be built in informal fashions that can lay a foundation for later more mathematical understanding, in particular starting with an awareness of variation. In answer to Green's (1993) third question about the essential experiences needed by students, these should be made more explicit in both curricula and classrooms.

- Put Variation and Expectation at the top of the curriculum as organisers for all of the detailed descriptions that follow.
- Begin with young students' innate intuitions about the variation they see around them in the world.
- Use discussion and description to develop ways of representing variation.
- See graphs as a way of showing variation in data sets – stating this explicitly as often as possible in the story-telling process.
- Use variation as a contrast to expectation and in straightforward non-mathematical contexts begin to ask questions such as "What do we think will happen? Why might it not happen? What might happen instead? Would we be surprised? Why?"
- Move to expectation with simple random generators where repeated experimentation can take place.
- Constantly reinforce what is observed and recorded in terms of varying outcomes.
- Record and graph both chance outcomes and data values, providing reinforcement for both what is expected to happen and how variation from this occurs.
- Provide many repetitions of random experiments and other data collections to reinforce the juxtaposition of variation and expectation.

In some settings the expectation is for a single value, such as the number of red lollies in a handful of 10, and the variation from that value produces a distribution of outcomes that reflects the theoretical distribution expected from a random process. Similarly if the heights of a number of students were measured and graphed, a distribution would be produced. In this case the expectation might be discussed in terms of middle of the distribution and the average calculated to reflect this representative value. In the first case the expectation might be discussed first and then the distribution graphed to display the variation from expectation. In the second, variation is likely to be the main interest as heights are graphed and then the mean (expected value) calculated to summarise the display in a single value. It is the juxtaposition of the two ideas that is important.

In more complex settings the expectation may be based on the distribution itself, questioning for example if variation in an observed distribution of outcomes or measurements is sufficiently different from the expected distribution to arouse suspicion about whether the observed process is consistent with the expected process. This type of question is considered by Watson and Kelly (2004) in terms of many repeated trials of 50 spins of a 50-50 black-and-white spinner. A scenario is set of classes carrying out the trials of 50 spins many times with some classes cheating and others actually completing the random process. Expectation in this case is related to a symmetric distribution around an expected value of 25 “successes” and variation is observed in relation to the idea of difference from a reasonable spread of values around 25. A distribution centred around 35 would not be consistent with the expected distribution, nor would a uniform distribution spread across 0 to 50. Discussion of such “meta-variation” and “meta-expectation” can take place in concrete contexts in the middle school long before statistical tests of goodness-of-fit are introduced. The intuitions built upon graphing experiences and discussion of variation should help make sense when interpreting *p*-values in later years.

If variation is the foundation of statistics, without which statistics would not exist (Moore, 1990), then it is expectation that draws information *from* variation to lead to an inference. Both statisticians (e.g., Wild & Pfannkuch, 1999) and statistics education researchers (e.g., Konold & Pollatsek, 2002) now indicate that variation is the pre-eminent idea, and this is supported by the research presented in this paper. It is *then* the balance between variation and expectation that provides confidence in statistical decisions made. This is expressed in tests of significance or confidence intervals in senior courses but learning to appreciate variation, and its relationship to expectation, can and should begin very early with suitable concrete experiences and questioning.

Acknowledgements

This research was funded by Australian Research Council projects, A79231392, A79800950, A00000716, and DP0208607. Thanks to Helen Chick for comments on an earlier draft of this paper.

References

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

Australian Education Council. (1994). *Mathematics – A curriculum profile for Australian schools.* Carlton, VIC: Author.

Bernstein, P. L. (1996). *Against the gods: The remarkable story of risk*. New York: Wiley.

Capel, A. D. (1885). *Catch questions in arithmetic & mensuration and how to solve them*. London: Joseph Hughes.

Green, D. (1993). Data analysis: What research do we need? In L. Pereira-Mendoza (Ed.), *Introducing data analysis in the schools: Who should teach it?* (pp. 219-239). Voorburg, The Netherlands: International Statistical Institute.

Hacking, I. (1990). *The taming of chance*. Cambridge: University Press.

Hart, W. L. (1953). *College algebra* (4th ed.). Boston: D. C. Heath.

Holmes, P. (1980). *Teaching statistics 11-16.* Slough, UK: Schools Council and Foulsham Educational.

James, G., & James, R. C. (Eds.). (1959). *Mathematics dictionary*. Princeton, NJ: D. Van Nostrand Company, Inc.

Kelly, B. A., & Watson, J. M. (2002). Variation in a chance sampling setting: The lollies task. In B. Barton, K. C. Irwin, M. Pfannkuch, & M. O. J. Thomas (Eds.), *Mathematics education in the South Pacific* (Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia, Vol. 2, pp. 366-373). Sydney, NSW: MERGA.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33***,** 259-289.

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

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26*, 20-39.

Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.

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

Salsburg, D. (2001). *The lady tasting tea: How statistics revolutionized science in the twentieth century*. New York: Henry Holt.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.), *People in mathematics education* (Proceedings of the 20th annual conference of the Mathematics Education Research Group of Australasia, Vol. 1, pp. 6-22), Waikato, NZ: MERGA.

Shaughnessy, J. M. (2002). *An investigation of secondary students' and teachers' conceptions of variability*. (National Science Foundation project, No. REC 02078420). Portland, OR: Portland State University.

Watson, J. M. (2000). *The development of school students' understanding of statistical variation*. (Australian Research Council grant No. A00000716). Hobart: University of Tasmania.

Watson, J. M., & Kelly, B. A. (2002). Emerging concepts in chance and data. *Australian Journal of Early Childhood*, *27*(4), 24-28.

Watson, J.M., & Kelly, B.A. (2004). Statistical variation in a chance setting: A two-year study. *Educational Studies in Mathematics, 57*, 121-144.

Watson, J. M., & Kelly, B. A. (in press). The winds are variable: Student intuitions about variation. *School Science and Mathematics*.

Watson, J.M., Kelly, B.A., Callingham, R.A., & Shaughnessy, J.M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology*, *34*, 1-29.

Watson, J. M., & Moritz, J. B. (2001). Development of reasoning associated with pictographs: Representing, interpreting, and predicting. *Educational Studies in Mathematics, 48*, 47-81.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*, 223-265.