Statistical Thinking: One Statistician's Perspective

Maxine Pfannkuch
Department of Statistics
The University of Auckland
New Zealand

This paper discusses some characteristic ways of reasoning within the discipline of statistics from the perspective of someone who is both a practising statistician and teaching statistician. It is conjectured that recognition of variation and critically evaluating and distinguishing the types of variation are essential components in the statistical reasoning process. Statistical thinking appears to be the interaction between the real situation and the statistical model. The role of variation in statistical thinking and the implications for teaching are also discussed.

Background

The development of students' statistical thinking is considered by statisticians (Bailar, 1988; Barabba, 1991; Snee, 1993; Wild, 1994)) to be crucial for the enhancement of student learning in statistics courses. This challenge to the current statistics teaching practice of teaching procedures raises questions as to what are the characteristics of statistical thinking and what approaches should be used in teaching for the development of statistical thinking?

After interviewing some undergraduate students in two separate case studies (Pfannkuch & Brown, 1996; Pfannkuch, 1996) on their solutions and reactions to statistically based information and problems some hypotheses were formed on the development of their statistical thinking. In order to broaden and understand the findings from these studies it was decided to investigate the nature of statistical thinking from a practitioner's or 'expert's' perspective. These statisticians would be in the position of commentating either from reflection on their own statistical thinking and/or on their clients' or students' statistical thinking. Through this interaction with 'experts' it was hoped that further insights would be gained into the characteristics of statistical thinking and that this would be informative for conjecturing the type of statistical thinking that should be developed in students. It would also confirm or refute previous findings which had been solely derived from student data.

Method

Several statisticians were individually interviewed in depth for approximately ninety minutes. Broad questions were asked; the interview following a semi-structured protocol based on the statistical enquiry empirical cycle. The interviews were audio-taped and transcribed. Each statistician had his own perspective on statistics, dependent upon the type of statistical work he was involved in and his own beliefs and experiences. Each one focussed on different aspects in a statistical investigation to explain his perceptions.

For the purposes of this paper one statistician's view is presented and analysed. Ray (not his real name) is a private consultant, working mainly in quality management and is also a supervisor and lecturer for Stage 3 (final year) undergraduate project students. In the Stage 3 course he has specifically developed a module on statistical thinking. Therefore Ray was considered to be in a position to comment on statistical thinking from a practitioner's and teacher's perspective. An analysis and discussion of his responses, including implications for teaching, is presented in this paper. The analysis and interpretation of Ray's comments were presented to him for corroboration.

Analysis

Two interlinked components emerged from an analysis of Ray's interview on statistical thinking. The first component is that statistical thinking involves 'noticing', understanding, critically evaluating and distinguishing the types of variation. This is a very broad component but it is all pervasive throughout the statistical enquiry process from posing the question through to the conclusions and appears to be the core of statistical thinking.

> *Ray: "Statistics is the science of variation basically. Statistical thinking is about evaluating variation, I guess. Variation contains all the information about what's going on."*

The second component is the realisation that a sound judgement of a situation can only be made by gathering and analysing data.

> *Ray: "Statistical thinking I think is about knowing the only way we get any information about the world is by taking samples of data in one form or another . . . It's about saying how would I find out about that . . . I can gather some data and turn it into information that can help me understand this."*

The analysis of Ray's interview has been categorised from a modelling perspective of statistics. The three categories are: understanding the dynamics of the real world problem; moving towards a statistical model; and using statistical tools. Each of these is considered below in detail.

First Category: Understanding the dynamics of the real world problem

In order to understand the context of the problem, an enquiry process is set in motion, that depends upon an ability to ask questions and to 'notice'. The 'noticing' of variation, the wondering why and the thinking about how to collect data to answer the questions is perhaps a scientific way of viewing the world.

> *Ray: ". . . [questions].should be generated by you walking around and bumping into processes . . . you've got a dehumidifier you could start thinking gosh there's a lot of water in it . . . you're walking along the pavement and you're seeing bits of moss growing on the pavement . . . you're in a coffee bar and you're served coffee spilt into the saucer . . . most of the data isn't in numerical form. It's impressions and fleeting thoughts.*
> *Int: . . . in noticing the moss, you are noticing variation.*
> *Ray: Yes.*
> *Int: When you're served spilt coffee, you are noticing variation. This coffee has spilt, this one hasn't.*
> *Ray: Yes, that's right."*

When Ray's students started their projects they realised that defining and understanding the problem that was to be investigated required an investment of time, thought and 'noticing'.

> *Ray: "[my students] have a company who's got this vague interest in improving the quality of their operation somewhere along the line. They don't know how, they don't know what's important to their customers, they don't know who their customers are. The first six months they've got to spend learning about the business, understanding it."*

The development of a broad background knowledge about the business or situation from which data is to be collected is considered essential by Ray. Therefore context knowledge as well as subject knowledge becomes important in defining and understanding the problem. To determine, for example, what 'improve' means in a specific situation students need to adopt a view of the enterprise that 'notices' variation. With this 'noticing' of variation is the understanding and recognition that there are different types of variation, such as, common cause and special cause variation. Common cause or stable system variation is hard to link to any particular source and knowledge on how to improve such a system can only come from an in depth investigation. Special causes can be assigned to an identifiable source.

> *Ray: "They go in and talk to people who are operating the machines or whatever they are looking at. Finding out what the little nagging problems are because they are all things which are special cause variation. . . . there are special causes where we get something from a supplier which doesn't match the specifications . . . when they go through their projects they recognise all these different types of variation, that impacts on what's going on."*

Recognition of the type of variation is important as it determines what strategy to pursue in addressing it and the action to take.

Second Category: Moving towards a statistical model

Once the problem has started to crystallise, measurement issues arise. Students are faced with the conundrum of what can be measured that will capture the essence of the problem, that will reflect a partial truth or model of the actual situation.

> *Ray: "Probably one of the biggest problems they [students doing projects] have is taking what is a reasonably complex situation whether it's a management situation or a production situation or a customer service situation and finding appropriate measurements. Measurements which reflect what is important about the process . . . the things we are interested in are always fuzzy things. . . . the customer wants prompt service; what's prompt service?"*

The narrowing down of a problem was a difficult process for both the students and company executives. Invariably both tried to tackle a problem that was too big. The refining of the problem, to explanatory and response variables, to measures and to stratifications that inherently reflect the process under investigation required an ability to recognise variation and understand the real problem

In their interactions with clients, students were faced with the fact that their clients' perceptions didn't match with the reality.

> *Ray: " It causes a lot of problems for them. The really good ones manage to work their way around that and become a team with the client . . . the others say things like these people are too difficult to work with ...they have a different agenda . . . at some point we try to teach them that there's always more than one perspective for looking at something. Don't for goodness sake think you've got the main eye on the truth here. . . . Study their [everyone in the organisation] perspective as it is going to be informative. . . . so different to yours, maybe they know something you don't. You can always learn."*

The importance of being aware that one's own beliefs and interpretation of situations and statistical information are dependent upon a personal viewing lens and that one should actively seek alternative explanations is stressed to Ray's students. Multiple interpretations of information required an ability in students to distil and to analyse the differing points of view and in their final analysis to capture the essence of the problem through proposing what measurements to take.

> *Ray: "[the good students] are good at talking to people and finding out what really matters . . . and saying well it seems that* ***these are the key measures*** *we should be concentrating on we should train people on how to do them . . . so they reflect what's important to the customer, to the organisation. "*

Many of Ray's project students are presented with data that has already been collected. The first aspect some noticed was the variation in the data. Because the recorders of the information had no stake in improving the system or did not get feedback as to why the data was important, glaring errors were found. This facet, that measurement errors overwhelmed any signals to such an extent that the data was useless is an eye-opener to students. Many organisations collected data with little attention being paid to the data collection process or appeared to collect and store data for no reason at all. These organisations had not defined what questions they wanted answered and therefore what information they wanted to know. The 'poor' students tended to take the data at face value and use it whereas the 'good' students delved more deeply into finding out about the data and what was relevant and what was irrelevant. The distinction that Ray drew between the two types of students, those that treated statistics as an algorithm for data and those that did not, was critical thinking.

> *Ray: "I guess they're just not critical enough. They're not picking up on that thing John Tukey said, the more you know what's wrong with the figure the more useful it becomes. Mostly figures are wrong. Deming said, if you're going to use numbers boy you've got to understand them. You ought to know what's going on."*

Those students who understood what the data meant in context were noticing, understanding and critically evaluating the variation in the collected data and thus were able to propose measures that could model part of the real problem.

Third Category: Using statistical tools

In Ray's experience many students have an inadequate understanding about sampling when planning their investigation. Theoretically the students knew how to take a random sample from a population but when the reality was presented they wanted to measure and analyse everything, often in a specific time interval. Nor did they appreciate that a small sample, in proportional terms, was often adequate. *"It's only based on a very small, in proportional terms, sample of the population so they don't want to make any great statement about the reliability. So there is a lack of statistical faith" (Ray).* This 'lack of statistical faith' could suggest an inadequate understanding of variation and knowledge that statistical tools took into account the variation from sample to sample. That samples will say something about the population from which they are drawn and will help gauge whether the system is stable (common cause variation) or not.

Data representation also presented problems. Ray stated that students could respond with ease to a given graph and point out all its failings yet when they were in the position of having to produce and create their own graphs they resorted to bar graphs, pie charts and chart junk. *"They'll only do it in that sort of critical fashion so that it's an analytical tool rather than a synthetic tool. So that if they are creating something they'll go back to their old habits" (Ray).* Thus a critical attitude was not adopted for their own work. He commented that the students had difficulty in interpreting basic graphical tools such as histograms and time series plots. They did not realise that a histogram said something about what was going on inside the process, that there were clues to be read. He believed that this was because the key element of distinguishing between stable variation (common cause variation) and special cause variation was missing from their statistical education. Graphs should be inspected and fundamental questions asked, such as, what is going on here, is this common cause variation or is this special cause variation? The reading and interpreting of graphs required context knowledge so that reasons could be put forward as to why the data looked like this, so that data could be split in another way and so that data could be explored graphically.

The recognition of the difference between common cause variation and special cause variation was not easy. Ray reported that many people including statisticians got it wrong because they looked at individual data and said it was exhibiting special cause when it in fact it was common cause. There was a tendency to examine the extremes and call them special cause. For example, Ray said, they will say something went wrong there, but it wasn't that something went wrong it was the way the system was managed. As an illustration Ray gave the example of a bank account that was causing the owner problems as it had become overdrawn three times, which could, at first glance, be attributed to the special cause of a large dentist bill. However on thinking about it, the owner of the account was a certain age and knew the state of his teeth. Hence he should have been able to reduce the impact of these dental bills through financial planning. Thus what had seemed to be a special cause was now drawn into the system of common causes. However there were some things, such as accidents, that could not be planned for. *"This distinction between common cause and special cause is not taught well" (Ray).*

Another point that Ray raised was that it did not seem to occur to some students to use a significance test. Instead they considered a visual test of the graph sufficient. Underlying this aspect seemed to be a lack of understanding of variation in relationship to significance testing. The recognition that special causes can be revealed in what is thought to be common cause variation in the summary statistics he considered to be important. Context knowledge and subject knowledge were very powerful in making those interpretations, making the thrust beyond information to understanding.

> *Ray: "What always struck me when I became involved in quality management was that I had got this degree in statistics. I'd done a lot of consulting over the years of all sorts, mostly biomedical areas. And I hadn't really made the connection from the summary statistics in the information through to the understanding of what was going on. What must be going on to generate those patterns. There was more diagnostic information in that stuff than I'd ever realised."*

An interpretation of the summary statistics requires context knowledge and the desire to make the connections to achieve a deeper level of understanding of the problem situation. Subject knowledge and context knowledge seem to be inextricably linked if the findings are to be communicated in a meaningful way.

Discussion

Variation in its many forms appears to be a key component in statistical thinking. In Ray's view it starts with an ability to 'notice' variation, that is be aware of variation in everyday activities, in specific activities, in planning procedures for the gathering of data, in data collection, in data sets, in graphs, in statistical summaries and in interpretations and conclusions. Many statisticians concur with this perspective that variation is central to statistical thinking.

> *"Pupils in the future will bring away from their schooling a structure of thought that whispers 'variation' matters." (Moore, 1992, p. 426).*
>
> *"Variation is not a new concept. What is new is the awareness of variation and how it affects everyday activities is infiltrating the workplace. . . . Knowledge of the theory of variation alters people's view of the world forever. It influences practically every aspect of how companies are managed". (Joiner & Gaudard, 1990, p. 35).*
>
> *"I define statistical thinking as thought processes, which recognise that variation is all around us and present in everything we do . . . " (Snee, 1990, p. 117)*

Moore (1990) gave five core elements of statistical thinking: the presence of variation in all processes; the need for data; the design of data production with variation in mind; and that statistical analysis seeks to quantify and explain variation. The two interlinked components that emerged from the above analysis would agree with these tenets. However, according to Ray and these statisticians, to understand variation in its many forms requires a new awareness, and possibly a new conceptualisation of variation. Ray emphasised that an understanding of variation required the ability to evaluate the difference between common cause and special cause variation. Thus it would appear that an understanding of variation in data may be based on these factors: (1) the recognition of variation in all processes; (2) the distinguishing between the types of variation; (3) some variation may be explained on current knowledge; (4) other variation cannot be explained on current knowledge.

The question then arises on the place of random variation in this construct. Random variation is often defined as 'unexplained' variation (Konold et al, 1991), yet at an intuitive level, particularly with data not obtained from random devices, individual data can be explained. In quality control the view taken is: *"All variation is caused. Unexplained variation in a process is a measure of the level of ignorance about the process"*. (Pyzdek, 1990, p. 102). To overcome this conceptual barrier random variation may need to be defined as a multiplicity of causes that are chosen to be ignored in the data

(Falk & Konold, 1991), or that when a pattern cannot be discerned the phenomenon is modelled as random. From these various definitions it would appear that random variation is a model which can be superimposed as a means of coping with unexplained variation and that randomness is a convenient human construct which is used to deal with variation in which reliable patterns cannot be detected.

Statistical tools appear to be seen by some students as descriptive tools rather than analytical tools. For example, some of Ray's students preferred to make inferences from visual tests of the graphs. This may be because there is a confusion about the type of statistics they are undertaking. Much effort must be spent in dialogue with the data, looking for and extracting persistent patterns or signals and thinking of multiple explanations. One explanation that does not seem to naturally occur to some students is the possibility that the difference is due to chance and that some statistical tools will evaluate that possibility or if there is a difference, that statistical tools will take into account the size of the sample (Konold, 1993; Pfannkuch, 1995). Behind these analytical tools lies the concept of variation.

Measurement issues, such as how to capture data that are relevant to the problem and what is realistically measurable, yet still reflect a way of modelling the situation was one of the biggest problems facing the students. This is an area that does not seem to be addressed in some statistics courses (Pfannkuch, 1996).

Throughout the statistical enquiry process, interpretation is a key factor in statistical reasoning, from the defining of the problem through to the conclusion. Interpretation of information in all its forms will be largely based on the students' experiences, perspectives and context knowledge. Therefore Ray thought that it was essential that his students search out alternative explanations. Other education researchers (Hancock et al, 1992; Pfannkuch, 1996) and statisticians (Bartholomew, 1995; Barabba, 1991) concur with this idea. The development of students' metacognition as well as their cognition and disposition in the specific area of interpretation would seem to be advantageous.

Some Teaching Implications

Teaching statistics where variation is the central tenet implies that when students are using statistical tools to explore data, subject knowledge and context knowledge are essential for interpreting the variation. Steinbring(1991, p. 506) believes that *"[stochastic] knowledge is created as a relational form or linkage mechanism between formal calculatory aspects on the one hand, and interpretive contexts on the other."* This may mean then that to conceptualise variation a strong linkage has to be created between the statistical thinking tools such as graphs and statistical summaries and the context of the situation. The assumption behind this epistemological triangle (Fig. 1) is that the concept of variation would be subject to development over a long period of time with different tools and different contexts.

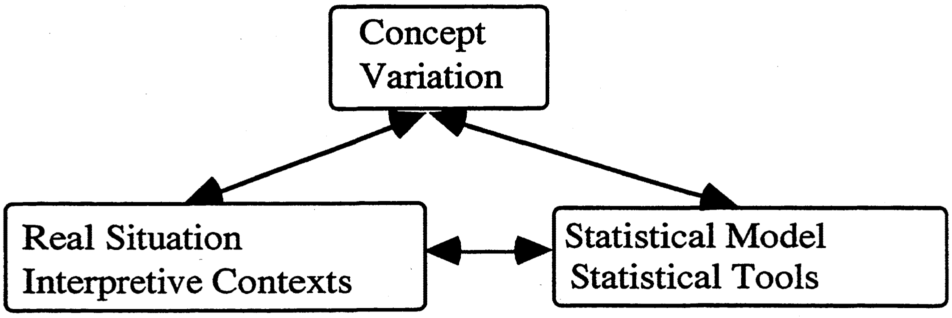

Figure 1. Epistemological Triangle

The development of statistical knowledge and concepts in this way would seem to suggest that it would also develop statistical thinking which could be regarded as the interactions between the real situation and its statistical model and between these and the resulting conceptual development. This interdependence between similar elements has been noted by statisticians such as Bartholomew (1995) who stated that statistical reasoning was based on the interplay of data and theory and educationists such as Pfannkuch (1996) who stated that context knowledge and subject knowledge were required to operationalise statistical reasoning.

The implication stemming from this is that teaching variation through games of chance alone will not give students sufficient understanding of variation. Understanding variation in data is quite different from understanding variation in random devices (Pfannkuch, 1996). The link between the two is very subtle. Historically probability has roots in two different lines of thought: the solution of gambling problems and the handling of data (Lightner, 1991). Today the gambling root of probability dominates teaching and textbooks. The other root in data needs to flourish alongside with the emphasis being on exploring variation rather than exploring probability as it is now in some curricula (Ministry of Education, 1992). Variation in all its contexts needs to be central to statistics teaching and thinking. And since context knowledge may be needed to operationalise statistical thinking then this implies that students should be taught from contexts that they have knowledge about.

Conclusion

Looking at the world from a statistical perspective means that there is a way of knowing, a way of thinking about the world that is uniquely statistical. Moore (1990) and Ullman (1995) state that statistical thinking, that is reasoning from uncertain empirical data, is a fundamental independent intellectual method. It would appear that this reasoning is based around the interplay between the real situation and the statistical model. Underpinning this, is the assumption that data is needed to understand the real situation. Statistical reasoning may not occur unless there is recognition of the underlying variation and an understanding of how to critically evaluate that variation from a real situation and a statistical model perspective. Distinguishing the types of variation and conceptualising variation may require the development of a new way of communicating these ideas in ordinary language and in intuitive models that are intellectually accessible (Fischbein, 1987). From this one statistician's perspective it would appear that an understanding of variation in the context of the original problem is a key element in statistical thinking.

References

Bailar, B. (1988). Statistical Practice and Research: The Essential Interactions. *Journal of the American Statistical Association.* 83(401), 1-8.

Barabba, V. (1991). Through a Glass Lens Darkly. *Journal of the American Statistical Association.* 86(413), 1-8.

Bartholomew, D. (1995). What is Statistics? *J.R. Statist Soc A 158*, Part 1 1-20.

Falk, R. & Konold, C. (1992). The Psychology of Learning Probability. In F. & S. Gordon (Eds.) *Statistics for the Twenty-First Century. MAA Notes*, Number 29 (pp. 151-164). USA: The Mathematical Association of America

Fischbein, E. (1987). *Intuition in Science and Mathematics.* Netherlands: Kluwer Academic Publishers

Hancock, C., Kaput, J. & Goldsmith, L. (1992). Authentic Enquiry With Data: Critical Barriers To Classroom Implementation. *Educational Psychologist.* 27(3), 337-364.

Joiner, B. & Gaudard, M. (1990). Variation, Management, and W. Edwards Deming. *Quality Progress*. December (pp. 29-37)

Konold, C. (1993). Understanding Probability and Statistics through Resampling. In L. Brunelli & G. Cicchitelli (Eds.) *Proceedings of the First Scientific Meeting of the International Association for Statistical Education.* (pp. 199-211) Perugia: University of Perugia

Konold, C. Lohmeier, J., Pollatsek, A., Well, A., Falk, R., & Lipson, A. (1991). Novice Views on Randomness. In *Proceedings of the Thirteenth Annual Meeting of the International Group for the Psychology of Mathematics Education - North American Chapter,* (pp. 167-173). Blacksburg: Virginia Polytechnic Institute and State University

Lightner, J. (1991). A Brief Look at the History of Probability and Statistics. *Mathematics Teacher,* 623-630

Ministry of Education (1992). *Mathematics in New Zealand Curriculum.* Wellington: Learning Media

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

Moore, D. (1992). Statistics for All: Why? What and How? In D. Vere-Jones (Ed.) *Proceedings of the Third International Conference on Teaching Statistics,* Vol 1 (pp. 423-428). Voorburg: International Statistical Institute.

Pfannkuch, M. & Brown, C. (1996). Building on and Challenging Students Intuitions about Probability: Can We Improve Undergraduate Learning? *Journal of Statistics Education* 4(1)

Pfannkuch, M. (1995). *Statistical Thinking Investigation: Report of Work in Progress.* Unpublished paper presented at NZARE Conference, Massey

Pfannkuch, M. (1996). Statistical Interpretation of Media Reports. In *New Zealand Statistical Association Research in the Learning of Statistics Conference Proceedings.* (pp. 67-76). Wellington: Victoria University

Pyzdek, T. (1990). There's no such thing as a common cause. *ASQC Quality Congress Transactions - San Francisco* (pp. 102-108).

Snee, R. (1990). Statistical Thinking and its Contribution to Quality. *The American Statistician,* 44(2), 116-121

Snee, R. (1993). What's Missing in Statistical Education? *The American Statistician* 47(2), 149-154.

Steinbring, H. (1991). The concept of chance in everyday teaching: aspects of a social epistemology of mathematical knowledge. *Educational Studies in Mathematics,* 22, 503-522.

Ullman, N. (1995). *Statistical or Quantitative Thinking as a Fundamental Intelligence.* Unpublished paper.

Wild, C. J. (1994). On Embracing the 'Wider View' of Statistics. *The American Statistician.* 48(2), 163-171.