INVESTIGATING THE NATURE OF STATISTICAL THINKING

Maxine Pfannkuch and Chris Wild, Department of Statistics,
The University of Auckland

In this exploratory study six professional statisticians were interviewed on their perspective of the nature of statistical thinking and on how they approached statistically based projects. Building on this and other studies some characteristics of statistical thinking are proposed and briefly described with reference to some of the statisticians' stories.

BACKGROUND

Amongst statisticians there is an increasing clamour for statistical courses to develop students' statistical thinking (e.g. Snee, 1993). Their argument is that the traditional teaching approach has focussed on the techniques and procedures of statistics and has failed to produce an understanding of statistical thinking. They suggest that there is a need to focus on authentic activity with students involved in problem solving, project work and working with real data sets. It would appear from their debate that they believe statistical thinking is developed through experiencing statistics in the broadest sense from the problem formulation stage through to the interpretation stage. Traditionally statistics has been predominantly associated with the analysis stage, but a group of statisticians (e.g. Bartholomew, 1995) believes that statisticians bring a unique perspective to all areas of the empirical enquiry cycle. This call for reform in the teaching of statistics raises the question of what is statistical thinking and how can it be characterised in the broad statistical domain.

In quality management there is a growing awareness of the need to develop statistical thinking at all levels in business and hence it is in this area where one finds a definition of statistical thinking. Britz, Emerling, Hare, Hoerl and Shade (1997, p. 67) state that "statistical thinking is a philosophy of learning and action based on the following fundamental principles: all work occurs in a system of interconnected processes; variation exists in all processes; understanding and reducing variation are keys to success". Statisticians such as Moore and Cobb (1997, p. 801) believe that the core of statistical thinking is the "omnipresence of variability" and that "statistics requires a different kind of thinking, because data are not just numbers, they are numbers with a context". Amongst statistics educators Shaughnessy, Garfield and Greer (1996, p. 206) state that “[statistics] is mathematical detective work within a context, and neither the context nor the principal players should be disassociated from the data”. From these ideas can be obtained the notions of variation, explanation, the importance of context and the need to dialogue with data. However there does not seem to be a full articulation of statistical thinking in general. The quality management definition of statistical thinking is from a business perspective. Therefore it is appropriate to try and define the characteristics of statistical thinking from a wider perspective for education.

Thus the research domain is the *broad* thinking skills that are invoked during the carrying out of a statistical enquiry. This exploratory research is aimed at uncovering some of these skills. It is part of a larger study using many sources to better understand statistical thinking. In order to gain access to some of these broad thinking skills six professional statisticians were interviewed for approximately ninety minutes. The interviews followed a semi-structured protocol based on the empirical enquiry cycle, PPDAC, as depicted by MacKay and Oldford (1994). (PPDAC is an acronym for problem, plan, data, analysis, conclusion.) The statisticians chosen worked in different fields of application and work environments: science (agriculture and forensics); brain mapping; biology; medical; market research; and quality management. The interviewees comprised two women and four men ranging in age from their mid twenties to late forties. The interviews, in the form of a conversation, naturally drew on their stories and insights into aspects of their projects that had surprised and impressed them. The interviews were conducted by the first author, were audio-taped, transcribed and initially analysed independently by each of the authors who then reached a consensus on the interpretation of the transcripts. These understandings were checked with the subjects. The data set was analysed using NUD*IST software.

EMERGING ELEMENTS

From the analysis of these interviews the following emerging characteristics of statistical thinking were proposed as follows.

Statistical thinking is the *integration of statistical and real problem understanding*. Certain elements underpin and/or facilitate it. For example

- interconnecting processes
- understanding and dealing with variation
- seeking explanations

- transnumerating (a coined word)
- interrogating constantly (including imagining)
- encapsulating complexity
- acknowledging and dealing with limitations

For this thinking to occur the statistician must interact with the problem situation. This means that the statistician will bring, dispositions, and the environment of the problem situation will bring, constraints, that will impinge on the statistical thinking.

Transnumeration is a coined word meaning a ‘numeracy transformation for facilitating understanding’. It occurs: when there is a quantitative description of the real system; when data are transformed in the statistical system; and when statistical summaries are changed to forms that relate more directly to the real system problem. The problem with the above categorisation and labels is that the meaning of them may not be interpreted as the authors intended. The following brief illustrations will give an indication of the authors’ meaning for some of these elements. (References to individual statisticians are by their field. For example, the market researcher is referred to as Market.)

DISPOSITIONS

It appeared that personal disposition provided the driving force behind the instigation, implementation and completion of a project. Curiosity, scepticism, openness and perseverance were mentioned in relation to disposition.

CONSTRAINTS

Constraints emerged such as the psychology of measurement and the primacy effect. There were also constraints related to the statistician, to the environment in which the problem was set and to resources. A number of these constraints resulted from the realities of working in a consulting environment.

The psychology of measurement was a surprising theme and was in relation to the statisticians’ clients. This psychology of measurement pervaded every phase of the empirical cycle and resulted in the statisticians talking about “managing their clients’ expectations” and “building up trust”. Science was adamant that for many of his clients the “original measurement has sanctity”, a phrase that describes a large part of this attitude towards measurement. For example, Market said that some of her clients did not want the data cleaned as “numbers were seen as immutable objects that can be trusted”. At the analysis stage Science stated some clients wanted him to stick to the same measurement and classification scale as the original measurement and were unhappy with transformations of data sets. Brain discussed the story of how a person he worked with felt that he had “stolen or deprived him of his measurements” when he reduced his forty measurements to ten in a paired comparison analysis on ten subjects. In the conclusion stage this “sanctity of the original measurement” resulted in clients wanting a picture that related to what they had originally measured. As a consequence all the statisticians mentioned using simple generic graphs or visuals that were meaningful to the particular field or context specific visuals that allowed direct interpretation to the real situation and related directly to the measurement process. An example of context specific visuals was given by Brain who related how the people he worked with had “developed an awful lot of tools for visualisation . . . that they can directly relate to and they superimpose the statistical images [e.g. p-values] onto a picture of the anatomy [of the brain] so that they can in fact see where these red blobs [e.g. extremely small p-values] lie relative to known structures of the brain.”

The primacy effect category (Fischbein, 1987) was distinguished in the interviews when an assumption at the beginning influenced the way a problem was subsequently viewed. Examples of this occurred when it was assumed that the ‘common practice’ use of statistics was correct or a commonly held community assumption was correct.

A STATISTICIAN’S STORY

In order to amplify further some of these elements emerging from the interviews, an abbreviated transcript of one Market’s stories is given. Phrases are numbered for future reference.

> “We did another job, for the City Council and that was measuring people’s attitudes towards the Bluebin recycling scheme (1), both attitudes and participation and volumes and what we were trying to do is get an understanding why (2), maybe, certain suburbs . . . actually had lower levels of participation in the bluebin recycling (3) . . . we did that by conducting a telephone survey (4) asking them what they recycled, the volumes they recycled, how often they recycled and then a lot of questions relating to their attitudes towards recycling and then demographics about the household (5), . . . got the data in and we analysed it and it’s mostly sensible (6) and it actually agrees with what the council thought about what suburbs are good at recycling and what suburbs are bad at recycling (7) and this showed sort of the... you

> know, the yuppie suburbs like [a named high socio-economic area] are really the best at recycling and the suburbs like, you know, [a named low socio-economic area] aren't really so good at recycling (8) . . . And that was the assumption (9) . . . our clients [assumption] and we had their understanding of what was happening (10). But, it turned out to be one that was completely wrong (11) and it took us to look at the data, look at the numbers and to think about them and realise that we had this assumption completely wrong (12), . . .we looked at what sorts of things were being recycled in the different suburbs (13) and the way they were being measured (14) . . . Our client was measuring recyclings based on weight (15). It turned out that people in [the high socio-economic area] looked like they were better recyclers, simply because they were recycling heavier things, drinking wine [from glass bottles], . . . whereas people in [the low socio-economic] were recycling plastic drink bottles (16) . . . and when you realise that you see that there's very little difference between the suburbs in terms of taking part in the recycling scheme and that really is a sensible interpretation of the data (17). . . . very near the end of the process . . . we were sitting down just talking and thinking (18) about how we would write the report and put the presentation together (19) and it sort of just dawned on us, we've been doing this thing all wrong . . . but yet at the time I had the sense that we might have missed it".

DISCUSSION ON A STATISTICIAN'S STORY

The above story of Market will now be used as a basis for describing some elements of statistical thinking that have emerged from the interviews. Note that (not incl.) means (discussed in the full story but not included in the excerpted parts of this story).

At the beginning of a project understanding the dynamics of the system in which the problem is set is necessary for the statistician to build up a mental model or picture of the *interconnections* in a system. This includes the understanding of how people operate in the system and the clients' perception of the problem and system. To obtain this understanding, *interrogation* of people in the system and *interrogation* of the client occurs, as well as the observing and noticing *variation* in the system and the seeking of alternative *explanations* for the phenomenon under study. Before Market could produce a survey on recycling, she needed to build up a model of how the recycling scheme worked, and to build up knowledge about recycling and the psychology of the people in the system (not incl.). At (1) she had made these *interconnections*, while at (2) she was seeking an *explanation* for the noticed *variation* (3). Once the system was understood and the problem *encapsulated*, measurement issues arise as to how to capture data from the system. *Transnumeration* occurs when measurements that reflect and capture the important elements from the real system are used at (5). The data collection process (4) by telephone was influenced by a cost *constraint* (not incl.). *Variation* was acknowledged and was dealt with in the design of the questions in the survey - that is, variation with regard to recycling and to the psychology of people being surveyed (not incl.). Despite this, a particular group's data still had to be discarded as a result of *interrogating* the data for reasonableness and of recognising the *limitations* of what can be captured by measurement (not incl.).

At (6), as the data are being analysed, there is an *interrogation* process whereby with an internal dialogue Market evaluates and judges with reference to what she knows about the real situation. She also checks externally (7) with the clients as to whether the analysis is reasonable and sensible in terms of their understanding of the real situation. At (8), *transnumeration* occurs when the data are converted from summary statistics to a word interpretation form in order to facilitate communication. However a *primacy effect* blinkered Market to other possible interpretations of the data. At (9) and (10) the *constraints* of beliefs and expectations from the self and from the environment in which the problem is set resulted in a wrong assumption. At (11), *encapsulation* of an aspect of the analysis is possible through *interrogating* the data, seeking another *explanation* and noticing *variation* (conveyed implicitly in phrase 12) through *transnumeration* (13). Here, transnumeration is used in the sense, that data are looked at from many perspectives through reclassification, multiple graphs and transformations of the data. The *interrogation* at (14) is an example of how, in each phase of the empirical cycle, a mapping to the real system and statistical system is engaged in, as well as looking forwards and backwards in the cycle in order to check the integrity of each phase. According to the statisticians interviewed it is crucial to check that measurements in the statistical system adequately capture crucial elements in the real system (*transnumeration*) which is demonstrated at (15). At (16) there is now an *explanation* for the *variation* and finally at (17), an *encapsulation* about how to interpret the data is determined. The importance of *interrogation* (18) is underlined at the final *transnumeration* phase (19) of communicating the data in a form that is understandable to the client (not incl.).

CONCLUSION

From these interviews it would appear that statistical thinking is differentiated by some key features: the role of variation; the integration of statistical knowledge with context knowledge; and transnumeration between and within the real system and statistical system. Even though the other elements such as interrogation could be considered generic, they are data based and therefore appear to be linked uniquely to a statistical perspective. It is to be hoped that these proposed categorisations will be useful in characterising and capturing some aspects of statistical thinking. And, if these emerging elements do help in characterising the nature of statistical thinking, then the question is raised as to how these elements should be developed in the teaching process so that statistical literacy is promoted.

SOFTWARE

*NUD*IST*. (1995). T. and L. Richards, Qualitative Solutions and Research, Box 171, La Trobe University Post Office, Vic 3083, Australia.

REFERENCES

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

Britz, G., Emerling, D., Hare, L., Hoerl, R., and Shade, J. (1997). How to Teach Others to Apply Statistical Thinking, *Quality Progress*, June 1997, 67-79.

Cobb, G. and Moore, D. (1997). Mathematics, Statistics and Teaching, *The American Mathematical Monthly*, *104(9)*, 801-823.

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

MacKay, J. and Oldford, W. (1994). *Stat 231 Course Notes Fall 1994*. Canada: University of Waterloo.

Shaughnessy, M., Garfield, J. and Greer, B. (1996). Data Handling. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick and C. Laborde (Eds.) *International Handbook of Mathematics Education*. (pp. 205-238). Netherlands: Kluwer Academic Publishers.

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