

TEACHER KNOWLEDGE AND STATISTICS: WHAT TYPES OF KNOWLEDGE ARE USED IN THE PRIMARY CLASSROOM?

Tim Burgess¹
Massey University, New Zealand

Abstract: School curricula are increasingly advocating for statistics to be taught through investigations. Although the importance of teacher knowledge is acknowledged, little is known about what types of teacher knowledge are needed for teaching statistics at the primary school level. In this paper, a framework is described that can account for teacher knowledge in relation to statistical thinking. This framework was applied in a study that was conducted in the classrooms of four second-year teachers, and was used to explore the teacher knowledge used in teaching statistics through investigations. As a consequence, descriptions of teacher knowledge are provided and give further understanding of what teacher knowledge is used in the classroom.

Keywords: cKc; elementary schools; mathematics teacher education; statistical investigations; statistical thinking; teacher knowledge

INTRODUCTION

Statistics education literature in recent years has introduced the terms of statistical literacy, reasoning, and thinking, and they are being used with increasing frequency. Wild and Pfannkuch's (1999) description of what it means to think statistically has made a significant contribution to the statistics education research field, and has provided a springboard for research that further explores and contributes to an understanding of statistical thinking and its application. Increasingly, it is recognised that statistics consists of more than a set of procedures and skills to be learned. School curricula, including New Zealand's, advocate for investigations to be a major theme for teaching and learning statistics.

Debate about teacher knowledge and its connections to student learning has had a long history. An important question arises as to what knowledge is considered adequate and appropriate. Although much is known about teacher knowledge pertinent to particular aspects of mathematics, the situation for statistics is less clear. Arguably, the mathematical knowledge needed for teaching and the statistical knowledge needed for teaching do share some similarities. Yet, there are also differences (Groth, 2007), due in no small way to the more subjective and uncertain nature of statistics compared with mathematics (Moore, 1990). Pfannkuch (2006, personal communication) claims that, because of the relatively brief history of statistics education research in comparison with mathematics education research, there is still much that is unknown about the specifics of teacher knowledge needed for statistics.

¹ t.a.burgess@massey.ac.nz

This paper reports on a framework that was proposed and applied in a study that investigated teacher knowledge needed and used by teachers during a unit in which primary school students investigated various multivariate data sets. The focus here is on justifying the need for such a framework in relation to teaching statistics, and on providing descriptions of teacher knowledge as revealed in the classroom in relation to the framework for teacher knowledge that combines statistical thinking components with categories of teacher knowledge. Examples from the classroom are provided to support the knowledge descriptions in relation to some of the components from the teacher knowledge framework. Finally, the conclusions consider some of the implications of this research, particularly for teacher education, both preservice (or initial teacher education) and inservice (or professional development).

LITERATURE REVIEW

Research on teacher knowledge is diverse. The thread of research from that of Shulman (1986) who defined pedagogical content knowledge (as one category of the knowledge base needed for teaching) provides a useful way of examining teacher knowledge. Shulman claims that a teacher's pedagogical content knowledge goes beyond that of the subject specialist, such as the mathematician. Subsequent research has attempted to clarify the differences between categories of teacher knowledge, either using Shulman's categories, or others developed from Shulman's categorisation.

Much of this research, although conducted with teachers, has not been conducted in the classroom, the site in which teacher knowledge is used. Cobb and McClain (2001) advocate approaches for working with teachers that do not separate the pedagogical knowing from the activity of teaching. They argue that unless these two are considered simultaneously and as interdependent, knowledge becomes treated as a commodity that stands apart from practice. Their research focused on the moment-by-moment acts of knowing and judging. Similarly, Ball (1991) discusses how teachers' knowledge of mathematics and knowledge of students affect pedagogical decisions in the classroom. For instance, the subject matter knowledge of the teacher determines to a significant extent which questions from students should or should not be followed up. Similarly, subject matter knowledge enables the teacher to interpret and appraise students' ideas. Ball and Bass (2000) argue strongly that without adequate mathematical knowledge, teachers will not be in a position to deal with the day-to-day, recurrent tasks of mathematics teaching, and as such, will not cater for the learning needs of diverse students.

A focus on the knowledge of content that is required to deliver high-quality instruction to students has led to another model of teacher knowledge, which involves a refinement of the categories of subject matter knowledge and pedagogical content knowledge. Hill, Schilling, and Ball (2004) claim that teacher knowledge is organised in a content-specific way, rather than being organised for the 'generic tasks of teaching', such as evaluating curriculum materials or interpreting students' work. Two sub-categories of content knowledge are further clarified by Ball, Thames, and Phelps (2005): *common knowledge of content* includes the ability to recognise wrong answers, spot inaccurate definitions in textbooks, use mathematical notation correctly, and do the work assigned to students. In comparison, *specialised knowledge of content* needed by teachers (and likely to be beyond that of other well-educated adults) includes the ability to analyse students' errors and evaluate their alternative ideas, give mathematical explanations, and

use mathematical representations. Ball et al. (2005) also subdivide the category of pedagogical content knowledge into two components, namely *knowledge of content and students*, and *knowledge of content and teaching*. These two parts of teacher knowledge bring together aspects of content knowledge that are specifically linked to the work of the teacher, but are different from specialised content knowledge. *Knowledge of content and students* includes the ability to anticipate student errors and common misconceptions, interpret students' incomplete thinking, and predict what students are likely to do with specific tasks and what they will find interesting or challenging. *Knowledge of content and teaching* deals with the teacher's ability to sequence the content for instruction, recognise the instructional advantages and disadvantages of different representations, and weigh up the mathematical issues in responding to students' novel approaches.

Although statistics is considered to be part of school mathematics, there are some significant differences that have implications for the teaching and learning of statistics. In mathematics, students learn that mathematical reasoning provides a logical approach to solve problems, and that answers can be determined to be valid if the assumptions and reasoning are correct (Pereira-Mendoza, 2002), that the world can be viewed deterministically (Moore, 1990), and that mathematics uses numbers where context can obscure the structure of the subject (Cobb & Moore, 1997). In contrast, statistics involves reasoning under uncertainty; the conclusions that one draws, even if the assumptions and processes are correct, are 'uncertain' (Pereira-Mendoza, 2002); and statistics is reliant on context (delMas, 2004; Greer, 2000), where data are considered to be numbers with a context that is essential for providing a meaning to the analysis of the data. It becomes necessary when teaching statistics, to encourage students to not merely think of statistics as doing things with numbers but to come to understand that the data are being used to address a particular issue or question (Cobb, 1999; Gal & Garfield, 1997).

Statistical literacy, reasoning, and thinking have featured in the statistics education literature in recent years. Ben-Zvi and Garfield (2004) provide some clarity for these terms, although with regard to statistical thinking, Wild and Pfannkuch's (1999) paper provided a model for statistical thinking. Wild and Pfannkuch describe five fundamental types of statistical thinking: (1) a recognition of the need for data (rather than relying on anecdotal evidence); (2) transnumeration – being able to capture appropriate data that represents the real situation, and change representations of the data in order to gain further meaning from the data; (3) consideration of variation – this influences the making of judgments from data, and involves looking for and describing patterns in the variation and trying to understand these in relation to the context; (4) reasoning with models – from the simple (such as graphs or tables) to the complex, as they enable the finding of patterns, and the summarising of data in multiple ways; and (5) the integrating of the statistical and contextual – making the link between the two is an essential component of statistical thinking. Along with these fundamental types of thinking are more general types that could be considered part of problem solving (but not exclusively to statistical problem solving). Wild and Pfannkuch's dimension of 'types of thinking' is one of four dimensions that explain statistical thinking in empirical enquiry. The other three dimensions are: the investigative cycle (problem, plan, data, analysis, and conclusions – these are the "procedures that a statistician works through and what the statistician thinks about in order to learn more from the context sphere" (Pfannkuch & Wild, 2004, p. 41)); the interrogative cycle (generate, seek, interpret, criticise, and judge) – this "is a generic thinking process that is in constant use by

statisticians as they carry out a constant dialogue with the problem, the data, and themselves” (Pfannkuch & Wild, 2004, p. 41); and dispositions (including scepticism, imagination, curiosity and awareness, openness, a propensity to seek deeper meaning, being logical, engagement, and perseverance), which affect or propel the statistician into the other dimensions. All these dimensions constitute a model that encompasses the dynamic nature of thinking during statistical problem solving, and is non-hierarchical and non-linear.

This model for statistical thinking was developed through reference to the literature following interviews with statisticians and tertiary statistics students as they performed statistical tasks (Wild & Pfannkuch, 1999). Although it was developed as a model applicable to the statistical problem solving of statisticians and tertiary students, it has subsequently been used in a variety of other studies, such as an examination of the thinking of primary students (Pfannkuch & Rubick, 2002) and pre-service primary teacher education students (Burgess, 2001), through a professional development workshop with secondary teachers (Pfannkuch, Budgett, Parsonage, & Horring, 2004), and an investigation into how statistical thinking of learners can be encouraged through a teaching activity (Shaughnessy & Pfannkuch, 2002).

The Framework

Teacher knowledge frameworks from the mathematics education domain are inadequate for examining teacher knowledge for statistics because of the differences between statistics and mathematics, as discussed earlier. The development of a teacher knowledge framework that takes into account the particular needs of statistics teaching and learning is therefore required. Such a framework must be specific to statistics, since teacher knowledge is organised in content-specific ways (Hill et al., 2004). Consequently the framework on which this study is based draws heavily on the statistical thinking model of Wild and Pfannkuch (1999). The categories of teacher knowledge that are described by Hill, Schilling, and Ball (2004) and Ball, Thames, and Phelps (2005), namely mathematical content knowledge and pedagogical content knowledge, and each of these with two sub-categories, provide a good starting point for examining statistics content knowledge as enacted in classroom teaching.

A matrix for a conceptual framework, against which statistical knowledge for teaching can be examined, is shown in Table 1.

Table 1: The framework for teacher knowledge in relation to statistical thinking and investigating.

		Statistical knowledge for teaching			
		Content knowledge		Pedagogical content knowledge	
		Common knowledge of content (ckc)	Specialised knowledge of content (skc)	Knowledge of content and students (kcs)	Knowledge of content and teaching (kct)
Thinking	Need for data				
	Transnumeration				
	Variation				
	Reasoning with models				
	Integration of statistical and contextual				
Investigative cycle					
Interrogative cycle					
Dispositions					

The columns of the matrix refer to the types of knowledge that are important in teaching. These four types are: common knowledge of content (ckc); specialised knowledge of content (skc); knowledge of content and students (kcs); and knowledge of content and teaching (kct). Hill, Schilling and Ball (2004) and Ball, Thames, and Phelps (2005) describe the features of these four categories of teacher knowledge in relation to number and algebra. These descriptions arise from a consideration of the question, “What are the tasks that teachers engage in during their work in the classroom, and how does the teachers’ mathematical knowledge impact on these tasks?” From those researchers’ close examination of teachers’ work, it is apparent that much of what teachers do throughout their teaching is essentially mathematical.

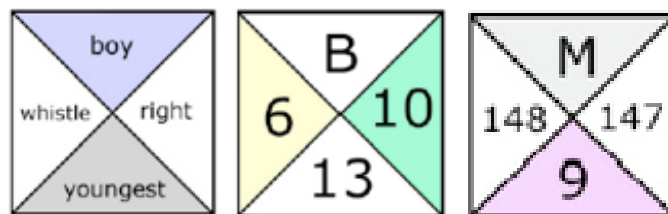
Just as Ball et al. (2001) claim that many of the everyday tasks of the teacher of mathematics are essentially mathematical, it is suggested that much of what a teacher engages in during the teaching of statistical investigations essentially involves statistical thinking and reasoning. Consequently, the four teacher knowledge categories are examined in relation to statistical thinking. The main feature that sets this framework apart from those offered for the mathematics

domain is the inclusion of the elements of statistical thinking and empirical enquiry (Wild & Pfannkuch, 1999), which are listed as the rows of the matrix.

THE STUDY

Since teacher knowledge is acknowledged to be important in relation to what and how students learn and is dependent on the context in which it is used (Ball & Bass, 2000; Barnett & Hodson, 2001; Borko, Peressini, Romagnano, Knuth, Willis-Yorker, Wooley et al., 2000; Cobb, 2000; Cobb & McClain, 2001; Fennema & Franke, 1992; Foss & Kleinsasser, 1996; Friel & Bright, 1998; Marks, 1990; Sorto, 2004; Vacc & Bright, 1999), it is argued that research should therefore take place in the classroom. Also, research on teacher knowledge must acknowledge and accommodate the dynamic aspects of teacher knowledge (Manouchehri, 1997), and be based on an understanding of how knowledge evolves. A post-positivist realist paradigm (Popper, 1979, 1985) was chosen because of the explanations about where knowledge comes from and how it grows in a dynamic fashion. Popper argued that knowledge develops through trial and elimination of error, and the logic of learning model (Burgess, 1977) was proposed as being appropriate for examining learning in classroom settings (Swann, 1999).

Using this post-positivist realist paradigm, case study research was undertaken with four inexperienced primary teachers (all in their second year of teaching), Linda, John, Rob, and Louise (all pseudonyms). The four classes were in the Year 5 (about 9-10 years old) to Year 8 (about 12-13 years old) level of primary school. The teachers were given a teaching unit that required students to investigate some multivariate data sets. The teachers developed their teaching based on this unit. The data sets generally consisted of 24 cases, each with four variables (or attributes). The first set used by each teacher included four category variables, while the other sets included at least two numeric variables along with the category variable(s). Each case was presented on a data card (see examples below from three different data sets), so that the students could easily manipulate and sort the cards in order to discover interesting things in the data.



Each lesson was videotaped, then edited by the researcher in order to focus on interesting episodes from the lesson. The edited videotape was shown to the teacher, and the discussion between the teacher and the researcher was audiotaped. The videotapes and the audiotapes from the post-lesson discussions were analysed in relation to the cells of the framework. Segments from the lessons or the discussions were identified in relation to the categories of teacher knowledge and the components of statistical thinking that were in evidence.

This paper reports on the results pertinent to the following research question:

What are the features of teacher knowledge in relation to aspects of statistical thinking that are used in the classroom?

DESCRIPTIONS OF THE FRAMEWORK

An understanding of the *need for data* on which to base sound statistical reasoning, instead of relying on and being satisfied with anecdotal evidence, is important in the development of statistical thinking. This corresponds to the first row of the framework. Classroom investigations can be conducted through two different approaches. First, an investigation can start with a question or problem to be solved and move onto data collection, which requires an understanding that data needs to be collected in order to solve the question or problem. The second approach is to start with a data set and generate questions for investigation from that data. By adopting this second approach for this study, teachers and students were not faced with the issues pertinent to establishing the need for data to help solve their questions. Consequently the *need for data* did not feature in this research. As such, the *need for data* is not described in relation to the four categories of teacher statistical knowledge for the framework.

Dispositions (corresponding to the final row of the framework), as another component of statistical thinking, did not emerge specifically in relation to the individual components of teacher knowledge but in a more general way. Teachers' statistical dispositions were apparent in the classroom. For example, inquisitiveness and readiness to think in relation to data along with an anticipation of what was to come was evident when Linda asked the students what they had started to notice when filling in their own data cards. She justified this question in the subsequent interview by saying that it was "to give them a hint of what was to come ... to see if the students had the inclination to start making their own conclusions already."

Common knowledge of content

As described by Ball, Thames, and Phelps (2005), *common knowledge of content* refers to what the educated person knows and can do; it is not specific to the teacher. They describe it as including the ability to recognise wrong answers, spot inaccurate definitions in textbooks, use mathematical notation correctly, and do the work assigned to students.

Wild and Pfannkuch (1999) describe *transnumeration* as the ability to: sort data appropriately; create tables or graphs of the data; and find measures to represent the data set (such as a mean, median, mode, and range). In general, transnumeration involves changing the representation of data in order to make more sense of it.

For teaching, *common knowledge of content: transnumeration* includes the knowledge and skills described above, along with the ability to recognise whether, for instance, a student gave the correct process or rule for finding a measure, had created a table correctly, or had sorted the data cards appropriately. Evidence of this category (as well as others involving common knowledge of content) was not often observed because the teachers generally used other types of teacher knowledge in relation to *transnumeration*. However if, for example, a teacher asked questions that led the students towards sorting the data in a particular way, it was assumed that the teacher

also had the *common knowledge of content* of how to do this for him or herself. There were instances where the researcher verified that this was indeed the case by asking the teacher during the interview to sort the cards, calculate a measure, or something similar. Consequently, *common knowledge of content: transnumeration* was subsumed within other categories of knowledge.

Consideration of variation in data is an important aspect of statistical thinking (Wild & Pfannkuch, 1999). It affects the making of judgments based on data, as without an understanding that data varies in spite of patterns and trends that may exist, people are likely to express generalisations based on a particular data set as certainties rather than possibilities.

The knowledge category of *common knowledge of content: variation* manifests itself in the classroom when the teacher gives examples of statements about data that acknowledge variation through the language used. Some of the more common situations that were observed related to inferential statements. Such statements were either about the actual data set and based on it, or generalisations about a larger group (population) from the smaller data set (sample). Such language included words and phrases such as “maybe ...”, “it is quite likely that ...”, and “there is a high probability that ...”. In addition, when the teacher talked about another sample being similar, but not identical, to the first sample, *common knowledge of content: variation* was evidenced.

For people to be able to make sense of data, statistical thinking requires the use of models. At the school level, appropriate models with which students could reason include graphs, tables, summary measures (such as median, mean, and range), and as used in this research, sorted data cards. If teachers demonstrated evidence of *common knowledge of content: reasoning with models*, it would be through making valid statements for the data, based on an appropriate use of a model.

Wild and Pfannkuch (1999) describe the importance of continually linking contextual knowledge of a situation under investigation with statistical knowledge related to the data of that situation. The interplay between these two enables a greater level of data sense and a deeper understanding of the data, and is therefore indicative of a higher level of statistical thinking.

The component of *common knowledge of content: integration of statistical and contextual* is characterised by the ability to make sense of graphs or measures, and by an acknowledgement of the relevance and interpretation of these statistical tools to the real world from which the data was derived. For example, John gave some possible reasons to support the finding that all the youngest students could whistle. He suggested that the older siblings could have taught the younger ones to whistle. This shows thinking of the real-life context in association with what the statistical investigation had revealed; such integration of the two aspects can sometimes enable the answering of ‘why might this be so’ that is being illustrated by the data.

One of the four dimensions of statistical thinking, as defined by Wild and Pfannkuch (1999), is the *investigative cycle*. This cycle, characterised by the phases of ‘problem, plan, data, analysis, and conclusions’, is what someone works through and thinks about when immersed in problem solving using data. If a teacher can fully undertake and engage with an investigation, then that teacher would be demonstrating *common knowledge of content: investigative cycle*. The teacher would be able to: pose an appropriate question or hypothesis, or set a problem to solve; plan for

and gather data; analyse that data; and use the analysis to answer the question, prove the hypothesis, or solve the problem.

For example, Linda discussed how data might be handled with an open-response type of question in a survey or census. Linda had considered, at the problem-posing phase of the investigation, how the responses from such an open-response type question would present a challenge at the analysis stage. This clearly indicated that Linda had some knowledge of the phases of the investigative cycle. She was able to maintain an awareness of a later stage of the cycle (analysis) while dealing with an early stage (planning data collection), and consider how decisions at that early stage could impact on the later stages.

A teacher would have *common knowledge of content: interrogative cycle* if it was evident that possibilities in relation to the data were considered and weighed up, with some possibilities being subsequently discarded but others accepted as useful. Engaging with data and being involved in 'debating' with it would be evidence of such knowledge. Likewise, developing questions that the data may potentially be able to answer is an aspect of *common knowledge of content: interrogative cycle*. Teachers who had immersed themselves with a data set prior to using it in teaching, so that they were aware of some of the things that might be found from the data, would be showing *common knowledge of content: interrogative cycle*. Such teachers would be prepared for knowing what their students might find in the data and what conclusions might be drawn from that data.

Specialised knowledge of content

A teacher requires *specialised knowledge of content: transnumeration* to analyse whether a student's sorting, measure, or representation was valid and correct for the data, particularly if the student has done something in a non-standard and unexpected way. It includes the ability to justify a choice of which measure is more appropriate for a given data set, or to explain when and why a particular measure, table, or graph would be more appropriate than another. Some of these skills, although considered part of statistical literacy (Ben-Zvi & Garfield, 2004), are still currently beyond what many educated adults can undertake. As such they are considered to be part of *specialised knowledge of content: transnumeration* rather than *common knowledge of content:transnumeration*.

Specialised knowledge of content: transnumeration was identified for all the teachers in the study. For example, Linda attempted to follow a student's description of how she had sorted the data and converted it into an unconventional table involving all four variables. The table consisted of: four columns labelled G, B, G, B; four rows with labels on the left to account for two more variables; labels on the right for three rows to account for the fourth variable; but no numbers or tally marks in the cells of the table to represent the sorted data. To determine the statistical appropriateness of that particular representation, Linda had to call on her *specialised knowledge of content: transnumeration* as she tried to make sense of the table. In another example in relation to some students deciding which measure or measures they should calculate for the data set (out of the mode, median and mean), Rob recognised that the mode would not be the most appropriate measure to use for the numerical data in question, and was able to give some justification regarding the inappropriateness of the mode.

Making sense of and evaluating students' explanations around whether it is possible to generalise from the data at hand to a larger group involves *specialised knowledge of content: variation*. For instance, when Linda asked whether there would be many boys who watched a particular programme on TV based on the class data that showed only a small proportion of such boys, a student answered, "Don't know; she hasn't asked all the classes yet." The teacher had to evaluate whether that was a reasonable response in relation to understanding of variation; Linda explained that there are factors that might affect the validity of this generalisation, but that the student's justification (about not having the data from the population so therefore it was not possible to make such a generalisation) was not a good reason for not generalising from the class data.

Specialised knowledge of content: reasoning with models is needed to interpret students' statements to determine the validity or otherwise of those statements. Students often struggled with making sensible and valid statements about the data based on a particular model they were using, and as a consequence it was not always straightforward for the teachers to make sense of the students' statements. Consequently, this category is seen as being quite distinct from *common knowledge of content: reasoning with models*.

Specialised knowledge of content: reasoning with models was a very commonly occurring component of teacher knowledge, especially as the focus of the unit was on finding interesting things in multivariate data sets, and making statements about these data sets. In many cases, students justified their statements through reference back to the model and as such, the teachers needed *specialised knowledge of content: reasoning with models* to help check the veracity of the students' statements. For example, the following interaction, initially between Linda and one student but later extended to the whole class, exemplifies the challenge for teachers to listen to and make sense of students' statements:

Student: That most girls can write with their right hand, ... most girls write with their right hand ... [inaudible].

Teacher: Sorry, I didn't catch what you said. Can you say that again for me? Slower this time.

Student: Most girls can write with their right hand are the youngest in ...

Teacher: Hang on. Most ... what are you saying? Most girls who produce their neatest handwriting with their right hand can whistle. ... [pause]. Okay ... [pause]. How many girls who produce their neatest handwriting with their right hand can whistle? ... [pause] Is that what you have got in front of you? [pointing at the cards on the desk] ... How many is that? [Student can be seen nodding as he counts cards] ... Is that these ones?

Teacher: So there are 5? ... These ones can whistle as well? But are they right handed? Okay. So what are you comparing that with? You said "most." So most compared with what? [No response from student.] In comparison with the right handed boys or in comparison with the left handed girls?

Student: Left handed girls.

Teacher: Okay... [pause] So R and J have taken that a step further and they have got ... [teacher moves to the whiteboard and starts drawing a type of two-way table – see Figure 1] ... here right-handed girls and right-handed boys and they have taken just this square [lower right] and sorted those people [the right handed girls] into different piles, into whistlers and non-whistlers. And they have found that there are more whistlers who are girls who are right handed than non-whistlers who are girls who are right handed. I think that is what they are trying to say.

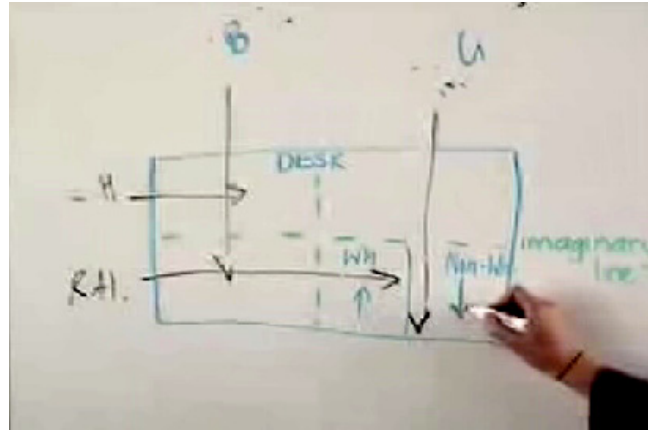


Figure 1: Diagram drawn by Linda to help students make sense of the statement from a student.

The interaction indicates the use of *specialised knowledge of content: reasoning with models* by the teacher, involving initially the model of sorted data cards on the student’s desk, followed by the model on the board that she created from transnumeration of the data cards.

Being able to evaluate a student’s explanation based on both statistical data and a knowledge of the context under investigation is one aspect of the category of *specialised knowledge of content: integration of statistical and contextual knowledge*. There were a number of situations in which the teacher prepared the students to gather data. Data collection questions had been suggested, such as, “What position are you in the family, youngest, middle or eldest?” When the students were considering the question prior to the actual data gathering, Linda was asked:

- Does it count if you have half brothers or sisters?
- What if your sister or brother has died?
- What if your brother or sister is not living at home?
- What would you put if you were an only child?

Each of these questions, and others involving the definition of family, were unexpected by Linda. She had to decide ‘on the spot’ how to respond to each question from students. She was required to weigh up the statistical issues related to answering such a data gathering question with the contextual issue of interpretation of ‘family’. Her answers indicated that she was able to do so satisfactorily and therefore were evidence of her having *specialised knowledge of content: integration of statistical and contextual*.

A teacher needs *specialised knowledge of content: investigative cycle* when dealing with students’ questions or answers in relation to phases of the investigative cycle, or when discussing or explaining various phases of the cycle and how they might interact. When thinking about suggestions for what could be investigated in a data set, the teacher needs to be able to evaluate the suitability of the problem/question, and whether it needs to be refined to be usable and suitable, in relation to the subsequent analysis.

So what does *specialised knowledge of content: interrogative cycle* look like, as distinguished from *common knowledge of content: interrogative cycle*? When a teacher has to consider

whether a suggestion from a student is viable for investigating within that data, the teacher requires *specialised knowledge of content: interrogative cycle*. Also, it involves determining whether a student's suggested way of handling and sorting the data would be useful to enable the later interpretation of results in relation to the question at hand.

Knowledge of content and students

The *knowledge of content and students: transnumeration* component includes: knowledge of the common errors and misconceptions that students develop in relation to the skills of transnumeration (including sorting data, changing data representations such as into tables or graphs, and finding measures to summarise the data); the ability to interpret students' incomplete or 'jumbled' descriptions of how they sorted, represented, and used measures to summarise the data; an understanding of how well students would handle the tasks of transnumeration; and an awareness of what students' views may be regarding the challenge, difficulty, or interest in the tasks of transnumeration.

There were situations in which students, when handling the data cards and sorting them, tried to consider too many variables at once and could not manage the complexity in the sorting of the cards and in making sense of what the cards showed. Linda was aware of this difficulty and guided the students to sort the cards 'more slowly'. She suggested sorting by one variable, and then splitting the groups by a second variable; she knew how many groups of data there would be from sorting by three variables and therefore that it needed to be simplified for the students. In general, the teachers did not realise how much the students would struggle with sorting the data cards, especially when the students were looking at numeric data such as arm spans, heights, and so forth. The teachers were surprised that the students did not naturally order the numeric data but simply grouped the data cards into piles. Furthermore, sorting data cards to check for and show relationships between two data sets was difficult for students, and most of the teachers underestimated the level of challenge that students would therefore face with sorting to show relationships in the data.

Knowledge of content and students: variation includes knowing what students may struggle with in relation to understanding variation, and to predict how students will handle tasks linked to variation. Whether students can appreciate and think about variation in data while looking for patterns and trends in the data is something that a teacher needs to listen for in students' explanations and generalisations. Although all the teachers posed questions as to whether it was possible to generalise from the class data to a wider group, there was no significant evidence of *knowledge of content and students: variation* being used by the teachers. It may be that for the investigations being conducted, such teacher knowledge of *variation* was not called on because the students were not ready for this inferential-type thinking. Since it was something new for the teachers to teach, they had not considered the statistical implications relevant to the students' readiness for thinking in relation to *variation*.

If a teacher can anticipate the difficulties that students might have with reasoning using models, or can make some sense of students' incomplete descriptions, then the teacher would be showing evidence of *knowledge of content and students: reasoning with models*. In one example of such knowledge, Rob described how he worked with a group of students who had made a statement

from the data cards comparing the number of boys with the number of girls who were right or left handed. Rob knew that the students were capable of proportional thinking so he encouraged them to consider proportions. He did so because the numbers of boys and girls in the data cards were different, and therefore using proportions for the comparison would be more appropriate than using frequencies. Rob knew these students sufficiently to encourage them to reason with a proportional model, which two of the students handled particularly well.

Can a teacher anticipate that students may have difficulty with linking contextual knowledge with statistical knowledge? Are students, through focusing on statistical knowledge and skills, likely to ignore knowledge of the real world, that is, contextual knowledge, or vice versa? Such aspects would give an indication of a teacher's *knowledge of content and students: integration of statistical and contextual*.

Whereas Linda's students' questions which related to the data question of position in the family (as discussed above) were unexpected, John anticipated such possible difficulties for his students and pre-empted their questions by asking the class how each child from a four-child family might answer the question, "Are you youngest, middle, or eldest in the family?" John's question encouraged the students to think about the data question (the statistical) in association with their knowledge of particular families (the contextual). This helped the students understand that statistics is not performed 'in a vacuum', removed from real issues, but deals with numbers that have a context (delMas, 2004).

Knowledge of where students might encounter problems or particular challenges in an investigation, and whether students will find an investigation interesting or difficult, are aspects of *knowledge of content and students: investigative cycle*.

One teacher predicted that students could have a problem with knowing how to interpret a data collection question so had to consider how he would deal with this potential problem within an early phase of the investigative cycle. The analysis phase of an investigation was predicted to present challenges for students in relation to them deciding on the form to present the data.

Some teachers were aware that students would be challenged within the investigative cycle with moving from the analysis stage to the drawing of conclusions or the answering of questions that had formed the basis of the investigation. Such awareness meant that those teachers had thought about how to address the students' difficulties.

Knowledge of how students would handle the development of appropriate questions for investigating the data, and the extent to which they might engage with the data and be prepared to question and consider various possibilities, are elements of *knowledge of content and students: interrogative cycle*.

There were a number of instances when teachers became aware that students, rather than fully engaging with the data and seeking possibilities, were focusing on a narrow aspect of the data, such as individual data points. The students then used this narrow focus to argue for or justify a particular position. Teachers who had *knowledge of content and students: interrogative cycle* were able to consider ways in which this tendency amongst students could be mitigated. Such considerations led to knowledge of content and teaching: interrogative cycle being utilised.

Knowledge of content and teaching

Knowledge of content and teaching, as far as mathematics is concerned, includes the ability to appropriately sequence the content for teaching, to recognise the instructional advantages and disadvantages of particular representations, and weigh up the mathematical issues in responding to students' unexpected approaches. So what are the features of knowledge of content and teaching with regard to statistics?

The ability to plan an appropriate teaching sequence related to transnumeration data, to understand which representations are likely to help or hinder students' development of the skills of transnumeration, and to decide from a statistical point of view how to respond to a student's answer, are all aspects of *knowledge of content and teaching: transnumeration*. All the teachers displayed this component of knowledge. Some examples of its use included suggestions: for how the data cards might be arranged on the desk when sorting; to spread the data cards within each group so that all the data cards could be seen, which helped with noticing patterns or irregularities within the data and then making statements about what had been found; and for creating a two-way table of frequencies as another useful representation of the sorted data cards.

How to structure teaching for understanding variation is the main component of *knowledge of content and teaching: variation*. Teachers intentionally modelled appropriate explanations and generalisations, through the use of language that acknowledged the existence of variation, and their questioning encouraged the students to consider whether various generalisations were appropriate. Students were challenged to consider the presence of variation in the data and therefore how it would affect statements that could be made about the data.

An example of a teacher using *knowledge of content and teaching: variation* arose when Linda challenged a student who claimed that, although all boys in the class could whistle, not all boys could whistle. She asked: "Why not? We have just found that all boys in this class can whistle. Why wouldn't it be the same everywhere else?" Linda justified this question as encouraging the students to think about "the bigger picture ... This was data for our class. It was just a sample of maybe everyone in our school". Another teacher, Rob, posed a question for the students to consider: "Will the things that we found out from the data squares yesterday be similar or different to our class?" This question was designed to encourage the students to consider variation; the challenge for students was to consider and account for similarities along with differences at the same time. Louise also posed a question that encouraged students to consider variation in data between samples; she asked how many boys in the school might have the same data square (i.e., respond identically to the four data questions), given that there were four boys in the class with that particular data square. When one student answered, "I don't know the right answer but there could be four in every class," Louise pushed the students' variation thinking further by asking whether there were other possible answers. By using her *knowledge of content and teaching: variation* in this way, she was encouraging the students to develop their conceptual understanding of variation.

Beyond asking questions such as in the examples above, the teachers did not know how to further develop the students' thinking about variation. Teaching the relatively sophisticated and

complex concept of variation and inference was new for these teachers. Therefore it is not surprising that evidence of *knowledge of content and teaching: variation* was relatively limited.

How should a teacher structure the teaching to encourage students' statistical thinking in relation to reasoning with models? This question is at the heart of the teacher knowledge category of *knowledge of content and teaching: reasoning with models*. A teacher with sound knowledge in this category would have considered various approaches to teaching this aspect, could justify a particular approach that was taken and maybe why other approaches were rejected, and could consider any statistical issues that might arise from students' statements or explanations.

John commented that because the students had tended to focus on only one variable at a time and make frequency-based statements for comparisons, he would structure the next lesson differently. He intended to encourage the students to consider two variables simultaneously, and would do this by posing some questions to focus the students, as well as suggest to them ways of sorting the data cards to enable the questions to be investigated. John's *knowledge of content and teaching: reasoning with models* and *transnumeration* developed as a result of becoming aware of a difficulty that the students had with reasoning with models, that is, as a result of a development of his *knowledge of content and students: reasoning with models*.

Knowing how to encourage students to consider the relevance of contextual knowledge in relation to the statistical investigation being undertaken is part of a teacher's *knowledge of content and teaching: integration of statistical and contextual*. The situations described above for *specialised knowledge of content: integration of statistical and contextual* (in relation to the definition of family and unusual cases) required the teacher to weigh up, prior to answering each student's query, the extent to which such interpretations of 'family' might affect the reliability of the data obtained. Linda commented:

Everyone has their own definition of what a family is ... so I decided that the children could, if they wanted to, include their half brothers and sisters.

Also John decided on an approach to teaching that involved asking students a question based on a 'what if ...' scenario, as he had anticipated a possible difficulty that students might have with interpretation of the question for a particular family. Louise encouraged her students to integrate the statistical and the contextual when she asked them to think of situations involving various aspects of statistics (such as graphs and summary measures of data), and what these are used for. These examples show that each teacher demonstrated some *knowledge of content and teaching: integration of statistical and contextual*.

Being able to encourage students to think about each phase of the investigation and to consider how these phases link to one another (i.e., to deal with the parts without losing sight of the whole) are components of the *knowledge of content and students: investigative cycle*.

Earlier, an example was given of a teacher predicting that students may have problems interpreting some data questions. John, based on this knowledge of students, considered how to approach his teaching so as to prevent the students from having such problems. He handled it in two ways: on one occasion, he discussed an example with the students about their experience of having students from another class gather data from them, and how they had found some of the data questions difficult to answer; on another occasion John asked the students about how they

would answer a particular data question, knowing that different interpretations were possible. By structuring the teaching in this way, based on *knowledge of content and students: investigative cycle*, the teacher successfully utilised *knowledge of content and teaching: investigative cycle*.

The strategies a teacher might use to address students' tendency to ignore a wide range of possibilities and, instead, be content with a narrow, restricted focus in their investigation of data, constitutes a part of *knowledge of content and teaching: interrogative cycle*. Being able to consider, from a statistical point of view, how such limited views of the data might impact on an investigation is another component of this category of teacher knowledge.

Linda decided that to assist the students to examine possible relationships in the data, it was important to spend some time discussing with the students what are relationships. Following this, Linda brainstormed with the class some possible relationships that might be investigated in the data. She considered that this was time well spent, as it enabled the students to focus quite quickly on the data and engage with it meaningfully from the outset. It was quite common for teachers to ask the students to think about what might be found in the data, once the students had an idea of what the data set contained (in terms of the variables), but prior to seeing the complete data set. Again, this teaching strategy helped the students to engage quickly with the data as they had already started to think about the data and had developed an interest in it. These examples are evidence of the teacher having *knowledge of content and teaching: interrogative cycle*.

SUMMARY AND CONCLUSIONS

The framework proved a useful tool for identifying aspects of teacher knowledge in relation to statistical thinking. These aspects were obtained from classroom episodes or interviews with the teachers that re-examined those episodes.

Generally, within one cell of the framework, it was found that there is a diversity of teacher knowledge pertinent to statistical thinking. Consequently, evidence of teacher knowledge as related to statistical thinking for one cell does not imply thorough and complete knowledge for those aspects in relation to the desirable knowledge associated with the lesson.

Some examples of each category of teacher knowledge are listed below. These examples have been derived from the study's data and discussion, and are by no means intended as a complete list of the knowledge that was observed in use or shown as needed in the teaching of investigations. Because so many statistical concepts were covered in the investigative process (from the posing of questions for investigation, consideration of data collection questions, analysis through sorting and other transnumerative processes, and concluding statements), the examples given are a small sample covering a wide variety of statistical concepts. Many classroom episodes resulted in multiple coding as more than one 'cell' of the framework was in evidence. Consequently, some of the examples given below show more than one type of knowledge and/or aspect of statistical thinking.

Examples of common knowledge of content (ckc)

Able to find the three measures of average (mode, median, mean)	ckc: transnumeration
Can explain why mode is not useful in certain instances	ckc: transnumeration
Considers the effect of sample size on generalising	ckc: reasoning with models
Knows that larger sample size leads to statement of greater confidence	ckc: reasoning with models
Changing the order of wording in conditional statement changes the group total and therefore the fraction (e.g., right handed whistlers or whistling right handers)	ckc: reasoning with models
Able to make a generalisation to a population	ckc: variation ckc: reasoning with models
Suggests reasons why the youngest child in a family is likely to be able to whistle	ckc: integration of statistical and contextual

Examples of specialised knowledge of content (skc)

Ability to make sense of students' data based statements, with reference to sorted data cards	skc: reasoning with models
Determines whether suggested data collection question is suitable	skc: investigative cycle
Recognition of inappropriate comparison of unequal sized groups	skc: reasoning with models
Ability of evaluate appropriateness of inferential statement	skc: reasoning with models and skc: variation
Explains why measures such as mean or median are used as appropriate summary of data	skc transnumeration:
Ability to link student's question about 'unusual cases' in relation to data collection question with contextual knowledge	skc: integration of statistical and contextual

Examples of knowledge of content and students (kcs)

Recognise the need for data collection questions to be closed, with only 2-3 possible responses otherwise students will	kcs: transnumeration
---	----------------------

struggle to sort and group data

Ability to anticipate students will struggle with making accurate inferential statements	kcs:reasoning with models and kcs:variation
Recognise that students will have difficulty with sorting to explore relationships between two variables	kcs: transnumeration
Recognise that students may find some data collection questions ambiguous	kcs: investigative cycle
Need to encourage students to examine the data, continually looking for patterns, interesting aspects	kcs: interrogative cycle
Recognise need for students to make links between what is found in the data with what they know about the real world	kcs: integration of statistical and contextual
Recognise that students find difficulty with making valid statements from data	kcs: reasoning with models

Examples of knowledge of content and teaching (kct)

Uses discussion with students to evaluate suitability of data collection question, and how to refine the questions to make them unambiguous	kct: investigative cycle
Can pose suitable questions to encourage inferential thinking	kct: reasoning with models and variation
Encourages students to predict what might be found in data, and revisits those predictions after sorting data and making data based statements	kct: investigative and interrogative cycles, and reasoning with models
Uses 2x2 table as suitable representation for helping make statements from data	kct: transnumeration and reasoning with models
Considers the statistical implications for data collection from student's questions about 'unusual' family situations (e.g., how would you answer the data collection question about your position in family if you have ½ brothers/sisters, if a brother/sister has died, ...)	kct: investigative cycle and integrating statistical and contextual
Gives examples of statements involving two variables that would be suitable for investigating to help encourage students with posing conjectures to investigate	kct: reasoning with models, interrogative cycle, and investigative cycle
Shows students a way to sort data by two variables, and suggests possible statements that can made from such a representation	kct: transnumeration and reasoning with models

This study, by conducting research on teacher knowledge in the classroom in which that knowledge is used, has provided a significant contribution to the research field. Literature searches have been unable to locate any other research in statistics education that both focuses on teacher knowledge at the primary school level and is classroom based. This study therefore provides important insights to what knowledge a teacher needs for teaching statistics, based on the reality of the classroom context.

With regard to the classification of knowledge through the framework, the category of *specialised knowledge of content* provided challenges in differentiating it from *common knowledge of content*. It is not possible, with what is known or not known about the common statistical knowledge of the 'typical' educated person, to be certain about the boundaries between *common knowledge of content* and a teacher's *specialised knowledge of content*. The research literature documents a considerable amount about statistical misconceptions, and the general need for a greater level of statistical literacy in today's world (e.g., Ben-Zvi & Garfield, 2004). This study's classification of and distinction between these two categories of teacher knowledge may need redefining. The suggested differences between these two categories are therefore tentative until proven inadequate.

This study focused on the knowledge for teaching statistics of teachers early in their teaching careers. As the teachers were all in their second year of teaching, some of their current knowledge could be attributed to development from the teachers' teaching experience or from knowledge that developed prior to their initial teacher education. This study's findings can provide guidance for what particular aspects of knowledge development should be the focus of initial teacher education programmes. As most initial teacher education students have not had the advantage of learning statistics through investigations, their *common knowledge of content* should be developed through immersing the students in investigations. As their *common knowledge of content* develops, their *specialised knowledge of content*, particularly for listening to and making sense of students' responses (such as through the use of video of students), will develop.

Knowledge of content and teaching (e.g., teaching sequences, advantages and disadvantages of various alternative data representations, and knowing how to respond from a statistical viewpoint to students' ideas, especially the unconventional ones) is dependent on *knowledge of content and students* (e.g., understanding the aspects of investigating data that present particular challenges for students, knowing the common misconceptions, or errors that students are liable to make). Consequently, these two categories of knowledge should also be a focus in initial teacher education programmes. Overall, all aspects of teacher knowledge must be targeted, as the connections between the categories of knowledge mean that individual categories of knowledge cannot operate in isolation.

Teaching statistics through investigations is a recent development in school statistics curricula. As most experienced teachers would have had little opportunity to teach statistics in this way, there are also implications for teacher professional development, irrespective of the length of teaching experience of the teachers. Targeting teachers' professional development in relation to

knowledge of content and students and teaching simultaneously with building their own *common knowledge of content* through investigations (and consequently also *specialised knowledge of content*) is considered an optimum approach.

REFERENCES

- Ball, D. L. (1991). Teaching mathematics for understanding: What do teachers need to know about subject matter? In M. M. Kennedy (Ed.), *Teaching academic subjects to diverse learners* (pp. 63-84). New York: Teachers' College Press.
- Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 83-104). Westport, CT: Ablex Publishing.
- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433-456). Washington, DC: American Educational Research Association.
- Ball, D. L., Thames, M. H., & Phelps, G. (2005). Articulating domains of mathematical knowledge for teaching. Retrieved May 13, 2005, from http://www-personal.umich.edu/~dball/Presentations/RecentPresentations/041405_MKT_AERA.pdf
- Barnett, J., & Hodson, D. (2001). Pedagogical content knowledge: Toward a fuller understanding of what good science teachers know. *Science Education*, 85(4), 426-453.
- Ben-Zvi, D., & Garfield, J. B. (2004). Statistical literacy, reasoning and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3-16). Dordrecht, The Netherlands: Kluwer.
- Borko, H., Peressini, D., Romagnano, L., Knuth, E., Willis-Yorker, C., Wooley, C., et al. (2000). Teacher education does matter: A situative view of learning to teach secondary mathematics. *Educational Psychologist*, 35(3), 193-206.
- Burgess, T. (1977). *Education after school*. London: Gollancz.
- Burgess, T. A. (2001). Assessing the statistics knowledge of pre-service teachers. In J. Bobis, B. Perry & M. Mitchelmore (Eds.), *Numeracy and Beyond (Proceedings of the 24th annual conference of the Mathematics Education Research Group of Australasia - Sydney)* (pp. 114-121). Sydney: MERGA.
- Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*, 104(9), 801-823.
- Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning*, 1(1), 5-43.

- Cobb, P. (2000). The importance of a situated view of learning to the design of research and instruction. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 45-82). Westport, CT: Ablex Publishing.
- Cobb, P., & McClain, K. (2001). An approach for supporting teachers' learning in social context. In F.-L. Lin & T. J. Cooney (Eds.), *Making sense of mathematics teacher education* (pp. 207-232). Dordrecht, The Netherlands: Kluwer.
- delMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 79-96). Dordrecht, The Netherlands: Kluwer.
- Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). New York: Macmillan.
- Foss, D. H., & Kleinsasser, R. C. (1996). Preservice elementary teachers' views of pedagogical and mathematical content knowledge. *Teaching and Teacher Education, 12*(4), 429-442.
- Friel, S. N., & Bright, G. W. (1998). Teach-stat: A model for professional development in data analysis and statistics for teachers K-6. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K - 12* (pp. 89-??). Mahwah, NJ: Lawrence Erlbaum.
- Gal, I., & Garfield, J. B. (1997). Curricular goals and assessment challenges in statistics education. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 1-38). Amsterdam: IOS Press.
- Greer, B. (2000). Statistical thinking and learning. *Mathematical Thinking and Learning, 2*(1), 1-9.
- Groth, R. (2007). Towards a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education, 38*(5), 427-437.
- Hill, H. C., Schilling, S., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal, 105*(1), 11-30.
- Manouchehri, A. (1997). School mathematics reform: Implications for mathematics teacher preparation. *Journal of Teacher Education, 48*(3), 197-209.
- Marks, R. (1990). Pedagogical content knowledge: From a mathematical case to a modified conception. *Journal of Teacher Education, 41*(3), 3-11.
- 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.

- Pereira-Mendoza, L. (2002). Would you allow your accountant to perform surgery? Implications for education of primary teachers. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS 6) - Cape Town* (pp. CD Rom). Voorburg, The Netherlands: International Association for Statistical Education.
- Pfannkuch, M., Budgett, S., Parsonage, R., & Horrying, J. (2004). Comparison of data plots: Building a pedagogical framework. Retrieved October 12, 2004, from <http://www.icme-organisers.dk/tsg11/>
- Pfannkuch, M., & Rubick, A. (2002). An exploration of students' statistical thinking with given data. *Statistics Education Research Journal*, 1(2), 4-21.
- Pfannkuch, M., & Wild, C. J. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 17-46). Dordrecht, The Netherlands: Kluwer.
- Popper, K. R. (1979). *Objective knowledge: An evolutionary approach*. Oxford: Clarendon Press.
- Popper, K. R. (1985). The growth of scientific knowledge (1960). In D. Miller (Ed.), *Popper selections* (pp. 171-180). Princeton, NJ: Princeton University Press.
- Shaughnessy, J. M., & Pfannkuch, M. (2002). How faithful is Old Faithful? Statistical thinking, a story of variation and prediction. *Mathematics Teacher*, 95(4), 252-259.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Sorto, M. A. (2004). *Prospective middle school teachers' knowledge about data analysis and its application to teaching*. Unpublished doctoral thesis, Michigan State University.
- Swann, J. (1999). The logic-of-learning approach to teaching: A testable theory. In J. Pratt & J. Swann (Eds.), *Improving education: Realist approaches to method and research* (pp. 109-120). London: Cassell Education.
- Vacc, N. N., & Bright, G. W. (1999). Elementary preservice teachers' changing beliefs and instructional use of children's mathematical thinking. *Journal for Research in Mathematics Education*, 30(1), 89-110.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-265.