"I believe the most helpful thing was him skipping over the proof": Examining PCK in a senior secondary mathematics lesson

Nicole Maher
University of Tasmania
<Nicole.Maher@utas.edu.au>

Helen Chick
University of Tasmania
<Helen.Chick@utas.edu.au>

Tracey Muir
University of Tasmania
<TraceyMuir@utas.edu.au>

Pedagogical content knowledge is widely considered an essential and complex facet of mathematics teacher knowledge, but little research has focused on PCK at the senior secondary level. This study explores some of the complexities of PCK in a teacher's lesson for senior secondary students by analysing data from lesson observation, the teacher's own commentary on the lesson, and students' perceptions of the teacher's knowledge and actions. Findings suggest that classroom norms relating to the kinds of problems students are typically required to solve can affect priorities about what is taught, so that what is attended to and valued by students may impact on teachers' PCK demands.

Introduction

Effective mathematics teachers have knowledge of mathematical content and how to make it accessible to students (Hill, Ball, & Schilling, 2008). Pedagogical content knowledge (PCK) encompasses the way subject matter knowledge is transformed into the content of instruction, with Shulman describing PCK as knowledge that embodies those qualities of the content "most germane to its teachability" (1986, p. 9).

This study examines the PCK evident in a teacher's lesson for senior secondary students, an area in which there has been limited research. Data were collected from a lesson observation, the teacher's own commentary on the lesson, and students' perceptions of the teacher's knowledge and actions. The lesson focused on variance of discrete random probability distributions, one of several lessons on statistics from a wider PCK investigation. This paper will explore the following research questions: What aspects of PCK does a teacher of senior secondary mathematics demonstrate during a lesson on measures of variability of discrete random probability distributions? What do multiple sources of evidence of PCK reveal about the teaching and learning of measures of variability of discrete random probability distributions?

Review of Literature

PCK has received increasing attention in mathematics education following its introduction by Shulman in the 1980s (e.g., Hill et al., 2008). A number of frameworks have been developed to identify and unravel the complexities of teacher knowledge including PCK (e.g., Chick, Baker, Pham, & Cheng, 2006; Hill et al., 2008). The critical attribute of a useful framework of teacher knowledge is the extent to which it identifies "that knowledge needed for student learning and understanding" (Graeber & Tirosh, 2008, p. 124). Some scholars caution against the superficial use of such frameworks as simple checklists of teacher qualities or knowledge dimensions (Mason, 2008; Watson, 2008). Watson (2008) suggests that a typological approach to mathematics knowledge in teaching (e.g.,"knowledge of students", "knowledge of text-books") can obscure the complexities associated with the way the categories are connected and inform each other. Mason (2008) advocates for a richly conceived PCK that serves to remind and allow teachers to "be mathematical with and in

2016. In White, B., Chinnappan, M. & Trenholm, S. (Eds.). Opening up mathematics education research (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia), pp. 423–430. Adelaide: MERGA.

front of their learners" (p. 307). Therefore, an insightful and nuanced approach to applying teacher knowledge frameworks is needed.

The PCK elements in the Chick et al. framework for analysing PCK – used in this study – offer a set of filters through which to explore teaching in action. While space does not permit the inclusion of the entire framework, a full version may be found in Chick et al. (2006). This framework includes elements of PCK particularly relevant at the senior secondary mathematics level where content becomes abstract. One such PCK element is *Deconstructing mathematics into key components*, evident when the teacher identifies critical mathematical components within a concept that are fundamental for understanding that concept. *Choice of examples* is another salient PCK element, which is also found in other frameworks (e.g., Rowland, Huckstep, & Thwaites, 2005). The use of exemplification in mathematics teaching has been widely researched (e.g., Bills et al., 2006; Watson & Mason, 2006). An "instructional example", according to Zodik and Zaslavsky (2008), is any example considered within the context of learning a particular topic. Good instructional examples invite students to build appropriate generalisations about a mathematical idea by directing them towards key features that make it exemplary. Examples are frequently utilised in teaching senior secondary mathematics and featured prominently in the lesson discussed in this paper.

Methodology

This paper uses data from a wider investigation into PCK for senior secondary mathematics and explores aspects of PCK from the perspectives of a teacher, his students, and the researcher by examining one lesson. A Grade 11/12 Mathematics Methods class from an independent school in Tasmania participated. Mathematics Methods is the main pre-tertiary mathematics courses offered in Tasmanian schools for students who intend to pursue tertiary studies with significant mathematical demands. The course is assessed by internal unit tests, and a final external examination covering topics including function study, calculus, and statistics.

The participants were the teacher, Mr McLaren, and the 16-18-year-old students in his class. Mr McLaren has been a teacher of senior secondary mathematics for 12 years. Of the nine students enrolled in Mr McLaren's class, five (one female and four males) provided student data. Teacher and student names are pseudonyms in this paper.

Data were collected during an introductory lesson on variance of discrete random probability distributions. The lesson was observed, video-recorded, and transcribed in full. Following the lesson, participating students completed a short questionnaire about the explanations and strategies their teacher used that helped their learning. A 15-minute semi-structured focus group interview was also conducted with the five students in which they commented on aspects of the lesson that they perceived were particularly useful. Mr McLaren participated in a 20-minute interview after the lesson, discussing his actions in the lesson. Both interviews were recorded and transcribed in full.

Teacher actions depicted in the lesson transcript were analysed by the authors to see if they corresponded with any of the PCK elements of the Chick and colleagues framework (Chick et al., 2006). The teacher and student interviews and the questionnaire responses were aligned to events in the lesson transcript, and analysed for further insight into teacher PCK, based on what the teacher explained about his teaching choices, and what the students noticed and claimed was useful to them. The analysis thus focused on the PCK enacted in the lesson and any insights into teaching and learning variance of discrete probability distributions provided by the multiple sources of evidence for PCK: researcher lesson observation, teacher commentary, and students' perceptions. Space constraints mean that the analysis presented here is illustrative of some issues, rather than exhaustive.

Results and discussion

This section gives an overview of the lesson followed by three vignettes highlighting key incidents that give insight into PCK as illuminated by multiple sources of data.

Overview of the lesson

The lesson began with the teacher showing graphs of two different discrete probability distributions designed to illustrate the concept of variance (see Figure 1). Comparing the graphs visually led to discussion about the spread of the distributions and an introduction to variance as a measure of spread. Mr McLaren described variance as "a measure of each of the values of x and how far they are away from the mean, squared, and multiplied by their associated probabilities" and presented the formula Var(X) = $\sum(X - \mu)^2$Pr(X =x).

Figure 1. Graphs of two discrete random probability distributions chosen by Mr McLaren

The process of calculating variance was then modelled using data from each distribution. An alternative formula for variance was then shown: Var(X) = E(X^2) - $[\mathrm{E}(X)]^2$. Following some teacher demonstration of the use of the formula, students completed a text-book exercise involving problems such as the one shown in Figure 2.

5 A random variable has the following probability distribution.

x	2	4	6	8
Pr($X = x$)	0.15	0.3	0.42	0.13

Find (a) Var(X)
(b) Var(2X)
(c) Var(3X + 1)
(d) Var(-5X + 7)

Figure 2. A question from a class exercise from the prescribed text-book Hodgson (2013)

Developing the concept of variance by comparing two distributions

Mr McLaren produced two probability distributions designed to draw attention to the idea of measures of spread. The distributions were symmetrical and had the same mean (expected value), median, mode, and range, but different variances. This aspect of the lesson fits the PCK category *Choice of examples* from the Chick et al. framework, evident when the teacher uses an example that highlights a concept or procedure. In the post-lesson interview, Mr McLaren elaborated on the considerations he had taken into account when developing this pair of examples:

> I wanted to have enough values for *x* but not too many, umm like *x* from 1 to 30. I didn't want to have that many values but then I didn't want to have too few values because then the spread wouldn't be as obvious so I just wanted a bit in between. I wanted to have one [graph] with a wide spread and one with a narrow spread around the mean. Both with pretty much everything else the same so they had the same mean, the same umm *x* values so everything is the same but you can see that the probabilities were spread umm narrowly with one and spread more widely with the other. So then you can get an idea that spread might be important. [Mr McLaren; interview]

During the lesson students were asked first to identify the similarities between the two distributions. To a certain extent students noticed the intended features.

> **Kale:** Well it goes from one to ten for both of them [the x values for each distribution]
>
> **Toby:** They are both symmetrical
>
> **Kale:** They both have a median of 5.5.
>
> **David:** They both peak in the middle. [Mr McLaren acknowledges that all of the above are correct.]
>
> **Mr McLaren:** What can you tell me about the mean, the expected value? [Mumbling from class.]
>
> **Kale:** I can't really see the numbers, Oh yes I can, I don't reckon they would be the same because they've got three under 0.05 for the blue and two under for orange.
>
> **Mr McLaren:** Are you sure?
>
> **Kale:** Ummm. Oh wait I reckon they might be the same because the blue has three under 0.05 and there's two under the orange and the blue ones has a higher peak so I reckon so it should actually average out to be similar.

Some students appeared to have difficulty interpreting the mean of the distributions and did not seem to recognise that the mean is the balance point of the distribution and so for symmetrical distributions this is the point of symmetry. Mr McLaren responded as follows:

> It actually doesn't matter what the probabilities actually are, the fact that it's symmetrical will mean that the expected value or the mean is 5.5 in both cases. So it is the case that the mean or the expected value is the same in both these distributions. [lesson excerpt]

Although Mr McLaren's response partly explained why the mean is the same for both distributions, he did not link the symmetrical distributions with the idea of the mean as the balance point (centre of gravity) of a distribution. When asked to consider the differences between the two distributions the students gave valid responses based on what they saw; for example "there are different probabilities for x in each one" (Toby), and "the peaks are higher in the blue, but then it stays lower in the blue. It's kind of more exponential in the blue" (Kale). While these comments alluded to the blue distribution having a smaller spread as it is concentrated around the mean, the students did not appear to make this link readily. Also, the teacher did not probe their thinking further to enable them to make the connection themselves. Rather, Mr McLaren intervened with the following comment:

> All of those things you have said are correct. The thing I wanted you to see and perhaps you have thought about this, is that the probabilities in this one [points to the orange graph] […] are spread out more than in this one [spans hands around the blue distribution]. The probabilities in this particular distribution [blue] are really focused around the centre whereas the probabilities in this one [points to orange distribution] are been spread out more. Can you see that? [lesson excerpt].

While the link to the spread of the distributions was made explicit in Mr McLaren's explanation, more precise language in relation to the probabilities could have been used. Rather than saying the *probabilities* are focused around the mean, it is more appropriate to phrase the idea in terms of the graph having *values* clustered around the mean that have a higher probability of occurring. This linguistic ambiguity may be a faux pas in the heat of the moment, or may indicate limited content knowledge for describing distributions.

The focus group interview and questionnaires offered limited insight into the extent to which students perceived the pair of graphs as helpful for their learning of the concept of variance, and thus useful as exemplifiers of a concept. One student in the focus group interview commented that she found the graphs useful, but although she makes some connection to the idea of spread, it is not clear that she links this to the formula for expected value that they do later in the lesson.

> **Grace:** Well umm I liked how he did the graphs because it was sort of a more like a visual representation like you saw that – Oh – it's not just numbers.
>
> **Researcher (R):** How did that help you?
>
> **Grace:** You could sort of see the shape of the data, umm what it was doing.
>
> **R:** Could you explain further?
>
> **Grace:** It sort of like compared like more spread out numbers than all like close together ones or concentrated one.
>
> **R:** Did you find that helped with any other aspects of the lesson?
>
> **Grace:** Umm Oh sort of it gave a bit more motivation [giggles] it was like "Oh we have a shape; it's not just numbers".

In this case the teacher has the content knowledge and PCK to construct a pair of examples with the potential to illuminate the idea of variance for students, but the way in which it was used in the classroom – perhaps the lack of probing for students to articulate the variational differences and the possibly premature jump to a teacher-led explanation, though valid – may have reduced the impact of the examples.

Introducing the alternative rule for calculating variance

While the formula Var(X) = $\sum(X - \mu)^2$Pr(X =x) explicitly articulates how to obtain the variance of a discrete probability distribution, performing the calculation using it is a lengthy process. An alternative equivalent formula is commonly used for calculating variance: Var(X) = E(X^2) - $[\mathrm{E}(X)]^2$. This formula is, arguably, less intuitive for students, as it is not clear how it captures the idea of values varying from the mean, which is at the heart of the concept of variance. Mr McLaren introduced this formula to the students and chose not to address its derivation:

> There is a proof to show you how to get from here to here [points to Var(X) = $\sum(x - \mu)^2$Pr(X=x) and then to E(X^2) - $[\mathrm{E}(X)]^2$] but I'm not going to worry about that at the moment [as] you don't have to know it. Really what you need to remember is this here [points to $\mathrm{E}(X)^2$ - $[\mathrm{E}(X)]^2$]. [lesson excerpt]

The instructional decision to omit the proof of this form of the variance formula was particularly noticed by one student:

> I think the fact that he decided to skip over the proof of the $\mathrm{E}(X)^2$ - $[\mathrm{E}(X)]^2$ [as being the same as $\sum(x - \mu)^2$Pr(X=x)] was helpful because it didn't add too much information. He didn't overload us with too much stuff yeah. [Kale; focus group interview]

A similar view was expressed by Kale in his written response: "The most helpful thing I believe was him skipping over the proof of $\mathrm{E}(X)^2$ - $[\mathrm{E}(X)]^2$ so as not to confuse us further." This aspect of the lesson provides mixed evidence of the teacher's PCK. He knows both formulae are valid, so that there is content knowledge and also *Procedural knowledge* (Chick et al., 2006, p. 299), since he knows how to use the second formula and later demonstrates its use in class. However, his decision to omit the proof is, perhaps influenced by awareness of what students will need to know for their assessment (which does not include knowing the proof), and perhaps his beliefs about the students' capacity to follow a proof (which one could argue is *Knowledge of students' thinking* or knowledge of *Cognitive demand*). His choice certainly appears to be appreciated by Kale. However, the choice also indicates possible beliefs about whether or not proof is an important part of learning mathematics, suggesting that this aspect of being mathematical in front of learners (Mason, 2008) is not regarded as significant.

Making sense of the formulae.

The text-book exercises completed during the lesson required students to calculate the variance of a linear function of the discrete random variable, such as Var(2X) (e.g., see Figure 2). This involved the use of the relationship: Var($aX + b$) = a^2Var (X). The following lesson excerpt depicts Mr McLaren's response to a student's question about why the b "disappears" yielding a^2Var (X).

> **Mr McLaren:** If you're asked to do something like this [writes the rule Var(aX+b) = a^2Var(X) on the board] […] you do the following [writes Var (2X) is equivalent to 2^2Var (X)].
>
> **Grace**: What about the b up the top?
>
> **Mr McLaren:** It's just gone.
>
> **Grace:** Really! [She sounds amazed].
>
> **Mr McLaren:** Yeah. [He pauses for a couple of seconds.] We could look at why but I'm not too fussed. [As the students continue working, Mr McLaren seems to ponder this. He looks down at the open text-book, turns a couple of pages and appears to be thinking.]
>
> **Mr McLaren:** [Turns to the board again.] If you think in terms of what the variance actually is – I said I wasn't going to show you – but anyway [He clears the board leaving the expression Var(aX+b) and proceeds to express Var(aX+b) in the form Var(X) = E(X^2) – [E(X)]^2. He hesitates a little, checks the text-book, and writes the following: Var(aX+b) = E$(aX + b)^2$ - [E$(aX + b)$]^2 = a^2E(X^2) + b^2 - [E$(aX + b)$]^2]

Mr McLaren did not attempt to deal with the [E$(aX + b)$]^2 component but explained that "in the end what will happen is we end up with minus b squared on the end, so the b squared terms end up cancelling out. I don't think you're going to come across this at all". In the researchers' judgement, this explanation seems to lack some mathematical rigour or clarity for a learner. Later, however, as part of the post-lesson interview, Mr McLaren was shown the video footage of this part of the lesson and he reflected as follows:

> […] I'm figuring out a way of explaining it better. I think I was fumbling around a bit there but umm yeah but because 'a' is having a multiplying effect on all of the 'x' values and 'b' is having an additive effect. When it comes to variance you're not worried about the additive effect because the spread is all added by that […] 'b' so the variance doesn't get changed by the 'b'. So 'b' has no effect on the variance which is a measure of spread. The spread remains the same umm so 'b' umm that's why the 'b' disappears. But the 'a' does have a multiplying effect and because variance is about a square of the differences the 'a' has got to be squared. [Mr McLaren; interview]

This response suggests that, during the interview, Mr McLaren was able to draw upon his knowledge of both the concept of variance and the algebraic processes involved in calculating variance, by articulating rich connections between them. At this point it seems evident that he can *Deconstruct mathematics into key components*, as evidenced by the fact that he identifies critical mathematical components within a concept or procedure that are fundamental for understanding and applying that concept (Chick et al., 2006, p. 299). At this stage this seems to be mostly content knowledge, although the explanation is clear for the researcher; what is not yet evident is the extent to which this explanation would work for students. It is possible to imagine that the teacher could take one of the distribution graphs from the first vignette and then transform it by an additive shift, and a multiplicative stretch, and have a discussion about how this affects the variance.

The majority of the student responses (both interview and questionnaire) focused on specific details involved in calculating variance. For example, in the interview Carl made the following comment in relation to the example involving Var($2X$) = 2^2 Var(X): "The most helpful thing that I learnt was those cases when the variance is like ($2X$), it's not like take the 2 and times it by the X, it's like take the 2 and square it and umm just those small things." Carl also commented that he thought the examples from the text-book exercise assisted him to become more fluent with the formula Var(X) = E(X^2) - $[\mathrm{E}(X)]^2$.

> Umm when you have the two separate ones [i.e., E(X^2) and $[\mathrm{E}(X)]^2$] you have to minus them and I was getting a bit confused about which one I was taking away. But when I did them [the text-book exercises] I was able to actually understand that you take this one square it and then go minus just the normal expected value squared. It just helped me to remember it yeah. [Carl; interview]

Toby focused more broadly on the computation of variance and its associated notation:

> Umm when like it's much easier to see what you're doing when each different thing has a symbol to represent them. So like how the Var [hesitates a little] I've forgot what that's called like Var of X had like a formula and then you can just use that to find out the answers to the different questions, if that makes sense. [Toby; interview]

> The most helpful thing was going through the questions step by step and to follow the exact steps so that you don't make a mistake. [Toby; questionnaire]

The teacher's choice of specific text-book examples for students to solve seemed to be appreciated by them. The problems, however, did not require attention to the meaning of variance, but only its computation. The changes to a distribution that would have helped answer Grace's original question were not discussed. This may be symptomatic of a broader set of classroom or subject norms, where conceptual understanding is not probed, and computation is all that is required. Typical exam questions in this topic do not seem to examine understanding of how variance changes with a change of distribution. This is, presumably, known to the teacher, implying his PCK includes knowledge of assessment requirements. This, in turn, impacts on the emphases given to different aspects of the topic, perhaps at the expense of students' conceptual understanding.

Conclusions

On the surface, the graphs chosen by Mr McLaren as examples to highlight the concept of variance represented a good instructional example – the relevant feature was made visible by keeping others the same (e.g., mean, median, range). It was unclear, however, the extent to which the students were able to connect the idea of spread with the visible differences between the distributions. Furthermore, some student responses suggest a limited understanding of how to interpret a probability distribution, particularly in relation to the expected value (mean) and the relationship between the shape of the distributions and their spreads (variances).

Mr McLaren's procedural approach to the symbolic formula, Var(X) = E(X^2) - $[\mathrm{E}(X)]^2$ was highlighted as useful by one of the students. Even though this formula is conceptually removed from the idea of variance, its derivation was seen as unnecessary, by the teacher and thus by the class, particularly because applying the formula was all that was required by the text-book exercises.

The tension between showing the students how to calculate variance and explaining what it means was particularly evident when Mr McLaren discussed the rule for calculating the variance of a linear transformation of the variable, such as Var(2*X*). He called upon his own knowledge in the moment of teaching to address Grace's query and attempted to deconstruct the formula to show why the *b* "disappears". It was not until later in the interview that Mr McLaren reflected more deeply on the conceptual link between variance and how it is calculated. This kind of deconstruction, however, is not called upon either by the prescribed text-book exercises or the statistics component of the course content in general, yet students indicated they found them particularly helpful.

Although not analysed in detail these lesson vignettes and the perspectives from both teacher and students give insight into a teacher's PCK for this topic and what is valued by the teacher and students. There is a suggestion that classroom norms relating to the procedural demands of the "typical" exam and text-book problems students are part of a teacher's PCK and can affect priorities about what is taught and attended to by students. More detailed analysis of this lesson and others should reveal more about the complex ways in which PCK is brought to bear in the classroom and noticed by students.

References

Bills, L., Dreyfus, T., Mason, J., Tsamir, P., Watson, A., & Zaslavsky, O. (2006). *Exemplification in mathematics education.* Paper presented at the Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education.

Chick, H., Baker, M., Pham, T., & Cheng, H. (2006). *Aspects of teachers' pedagogical content knowledge for decimals*. Paper presented at the 30th Conference of the International Group for the Psychology of Mathematics Education.

Graeber, A., & Tirosh, D. (2008). Pedagogical content knowledge: Useful concept or elusive notion. In P. Sullivan & T. Wood (Eds.), *The international handbook of mathematics teacher education: Knowledge and beliefs in mathematics teachinig and teaching development* (Vol. 1, pp. 118-132). Rotterdam, The Netherlands: Sense.

Hill, H., Ball, D., & Schilling, S. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372-400. doi: 10.2307/40539304

Hodgson, B. (2013). *Maths quest 12 mathematical methods CAS* (2 ed.). Milton, QLD: John Wiley & Sons Australia.

Mason, J. (2008). PCK and beyond. In P. Sullivan & T. Wood (Eds.), *International handbook of mathematics teacher education: Knowledge and beliefs in mathematics teaching and teacher development* (Vol. 1, pp. 302-322). Rotterdam, The Netherlands: Sense.

Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education, 8*, 255-281.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Watson, A. (2008). *Developing and deepening mathematical knowledge in teaching: being and knowing.* Paper presented at the MKiT 6, Nuffield Seminar Series, 18th March, at University of Loughborough.

Watson, A., & Mason, J. (2006). Seeing an exercise as a single mathematical object: Using variation to structure sense-making. *Mathematical Thinking and Learning, 8*(2), 91-111.

Zodik, I., & Zaslavsky, O. (2008). Characteristics of teachers' choice of examples in and for the mathematics classroom. *Educational Studies in Mathematics, 69*(2), 165-182.