APPROACHING THE BORDERLANDS OF STATISTICS AND MATHEMATICS IN THE CLASSROOM: QUALITATIVE ANALYSIS ENGENDERING AN UNEXPECTED JOURNEY

JANE M. WATSON
University of Tasmania
Jane.Watson@utas.edu.au

ERICA L. NATHAN
University of Tasmania
Erica.Nathan@utas.edu.au

ABSTRACT

To capture aspects of pedagogical content knowledge (PCK) not illuminated in an earlier written survey, an interview protocol was used with 40 middle school teachers. The scenarios were intended to elicit teachers' understanding of the big ideas, ability to anticipate students' answers, and intervention strategies for the classroom. This was expected to be a straight-forward journey based on teachers' responses to three context-based scenarios regarding students' answers to questions. Instead we were surprised by teachers' responses that revealed their perceptions that their experiences teaching mathematics and teaching statistics are very different. This led to further analysis of the PCK tasks and a suggestion that the mathematics embedded in the tasks was sometimes an impediment for the teachers, especially in relation to intervention strategies in the classroom.

Keywords: *Statistics education research; Teaching statistics; Teaching mathematics; Pedagogical content knowledge; Teacher interviews*

1. INTRODUCTION AND BACKGROUND

This paper reports on the evolution of a hybrid analysis of data from the three-year research project entitled StatSmart (Callingham & Watson, 2008). From the beginning, this project was designed to explore, refine, and improve pedagogical knowledge for teaching statistics to middle school students in Years 5 to 9 (aged 11 to 15). This remains the prime research focus as the analysis enters its final stage. Although the wider context of this research is important to keep in mind (e.g., Watson, Callingham, & Donne, 2008a, 2008b), it is the surprise intrusion of a subsidiary theme that developed from qualitative analysis that is the focus here. We offer a narrative of how this theme emerged in its own right and in a way that complemented our initial research focus. This theme relates to the differences that teachers perceive when teaching statistics and teaching mathematics. For the purpose of this paper, we highlight the illustrative value of this theme that emerged from our experience of applying both qualitative and quantitative analysis in educational research. We do not venture into the extensive debates that define the epistemological boundaries of these approaches.

Statistics Education Research Journal, 9(2), 69-87, http://www.stat.auckland.ac.nz/serj

StatSmart began in 2007 (Callingham & Watson, 2008), funded by the Australian Research Council with the support of project partners: the Australian Bureau of Statistics, Key Curriculum Press, and the Noel Baker Centre for School Mathematics. Initially forty teachers working in a range of Tasmanian, Victorian, and South Australian schools were invited to participate. The teachers and/or their schools had links to one of the research partners in their states. Most of these teachers, along with their middle school students, have remained in the project across the three years. Teacher surveys were devised to advance our understanding of pedagogical content knowledge (PCK) as first espoused by Shulman (1987). Attentive to the adaptations of Shulman's framework in mathematics education (e.g., Chick, 2007; Hill, Rowan, & Ball, 2005), we wanted to apply the concept of PCK to statistics teaching. The initial question was how could we best refine this concept, which has sustained its integrity for several decades, so that a measure of change in statistical PCK could be obtained?

After the first round of teacher surveys in 2007, we realised that some aspects of PCK were difficult to distil by relying on the tight format of the surveys alone. It was decided that interviews might offer greater opportunity to analyse the working components of PCK. These were completed in 2008. At that point some creative tension developed within the project. The interview material offered a further opportunity for a qualitative approach to the data that had the potential to dovetail with the teacher surveys. By opening the methodology to qualitative analysis, the interviews also offered the base material from which to consider a refinement to the conceptual framework of PCK. Furthermore, a qualitative approach encouraged new questions to emerge with the capacity to challenge the analysts' original research parameters. The conceptual challenges presented by the qualitative analysis of the interviews are considered elsewhere (Watson, Callingham, & Nathan, 2009; Watson & Nathan, 2010a, 2010b). In brief, we were encouraged to expand the conception of PCK suggested from the initial survey analysis (Watson et al., 2008a) by integrating a fourth component we called "Shift to General" (Watson et al., 2009). In this process, we discovered that a focus on compartmentalising different features for analysis and assigning quantitative value to them can mask the importance of the relational connections between the identified components that make up the "whole." In our study, the "whole" we aim to define more completely is statistical PCK, and it remains a central task that is being modified as the research progresses.

2. METHODOLOGY

In this paper we are particularly interested in one significant and unforeseen dimension that a qualitative approach exposed. It grew out of an interview design that was in two parts. The central part comprised pedagogical questions across three statistical problems in different contexts. Initially our selection of the three problems for the exploration of PCK was to represent a range of statistical ideas appropriate to the middle school curriculum. Teachers were asked to interpret a pictograph (see the example labeled "Tom" in Appendix A) that exemplified the idea of uncertainty, basic frequency, and proportional reasoning. The second problem involved the interpretation of a media report of a survey (see the example labeled "Marijuana" in Appendix A), which raised issues of sampling. Thirdly, a two-way table (see the example labeled "Lung Disease" in Appendix A) provided an opportunity for teachers to discuss the association of variables and more complex proportional reasoning. By asking teachers to respond to real answers that students had given for these problems we intended to glean the teachers' own understanding as well as their capacity to teach what they understood to be the ideas behind the content. The more general interview questions, that is, those not tied to the three specific statistical problems, were designed originally to supplement the core material as a means of allowing themes to emerge that were not so constricted by the interviewers' framework. A secondary function of these more open questions was to ease the teachers into and out of the interview. The complete interview protocol is provided in Appendix A.

Forty teachers, 14 from each of two states and 12 from the third, had been involved in the StatSmart professional learning project from 8 to 18 months. Although participating in the project because they taught Years 5 to 9, some taught to Year 12. They had teaching experience ranging from 2 to more than 25 years and a wide range of previous study in mathematics and statistics. They were interviewed in their schools by one of four members of the research team. The interviews were recorded and transcribed for our detailed analysis.

Initial analyses of the three central problem-based parts of the protocol were completed separately of each other, using the rubrics in Appendix B. The rubrics were developed to cover the initial questions and a combination of the responses to the student answers provided us with further prompts to explore the teachers' PCK. The rubric for each of the three tasks consists of four components. The first of these we called the "Big Ideas." Could teachers identify the essential idea/s behind the problem? Was this achieved immediately and expressed in specialist language or were the ideas gradually revealed as the interview segment progressed? The second component we titled "Anticipates Student Responses." Could teachers anticipate both appropriate and inappropriate answers that their students might give to the problem? Did the teachers discern a hierarchy of answers or reveal an insight into the thinking behind their students' answers? Our third component was "Content-specific strategies." Did the teachers move beyond "telling" the interviewer the appropriate range of responses to a problem and demonstrate their capacity to initiate or structure a discussion in a purposeful manner? Did they reference the problem by giving examples from other fields? And finally, our fourth component we called "Shift to General" (cf. Watson et al., 2009). This was a measure of a teacher's capacity to conceptualise, to link the specifics of a statistical problem to the underlying idea(s), to related concepts, and to the ambiguity of language. A teacher who was unable to make this "shift to general" remained encased in the details of the initial problem.

At the beginning of the interview teachers were asked, "What topics in statistics do your students have most difficulty with and why?" Later, two of the questions at the end of the interview were: "What's the best thing about teaching statistics? What do your students think is the best thing about learning statistics?" Although a rubric to assess anticipated answers could have been prepared for these questions as was done for the three problems central to the interview protocol, it would not have captured the significance of these questions beyond an initial listing of topics and "best things" about teaching and learning statistics. Repeated reading of the transcripts from the different perspectives of the two authors, taking note of our changing impressions with each reading (Bouma & Ling, 2004, p. 183), precipitated the emergence of the new theme concerning the differences in teaching mathematics and teaching statistics. Although initially this might be considered a grounded approach based on the questions in the protocol, and certainly there was developing sensitivity to the data as required by Bouma and Ling, as time progressed we used a clustering approach that is best described by Miles and Huberman (1994) to collect together evidence in the teachers' responses. The responses, both articulated and demonstrated, were not confined to the general interview questions. It was the depth of the responses that encouraged us to develop this theme even though it was unplanned. It is our experience that qualitative analysis allowed the interview material to speak to the researchers with greater force than if it had been a ready-made construct. Accepting that the language of the interviewees was an integral element of the analysis, we proceeded to categorise the differences between teaching mathematics and statistics that the teachers identified in their responses to the general questions and more inadvertently, throughout the interview. These differences were successively reread and eventually three themes emerged from the data: Teaching Practice, Curriculum Values, and Cognitive Experience. The extracts were then clustered around the three themes. Our new questions hence became the following: How do middle school teachers perceive the difference between statistics and mathematics in the classroom? What does this disparity tell us about teacher understanding of the discipline of mathematics for middle school students? What influence does this have on our understanding of the responses to the three problems in the core of the interview? A more qualitative approach allowed for unexpected responses to emerge and also allowed greater scope for the research to change direction or to keep direction. It was not just that this analysis added an unforeseen element to our research, it also enriched our capacity to refine statistical PCK.

Once this new theme was identified, the teachers' responses were categorized and entered in a table so that the distribution of this body of evidence could be more clearly discerned. Importantly, discovering this new theme also encouraged us to analyse the central problem-based material of the interviews differently in two respects. We

1. searched for evidence (quotations) that amplified the theme of a mathematics-statistics teaching spectrum, and
2. examined the rubric results for statistical PCK with the added perspective of profiling the three problems according to the inherent level of mathematics involved in the tasks.

The Results are hence presented in line with these two aspects of the new theme. Descriptive data on the supporting quotes and brief extracts are included in the Results section.

3. RESULTS

3.1 "SHADES OF GREY": MIDDLE SCHOOL TEACHERS' PERCEPTIONS OF THE DIFFERENCE BETWEEN STATISTICS AND MATHEMATICS IN THE CLASSROOM

As stated, the significance of this potential difference for the interviewed teachers became evident when their responses to the general questions, mainly about the "enjoyment" of teaching/learning statistics, evolved into a comparison of teaching the two subjects. More than that, on occasions their reflections on the difference leached into their responses to the statistical problems. A third indicator that this sense of difference was important to the teachers was in their language. Overall teachers spoke with authority, certainty, and sometimes with an energetic spark that was in contrast to other parts of the interview. This more assertive tone is present in the following statement volunteered by a teacher reflecting on likely student responses to a two-way table problem (italics represent stronger emphasis in the teacher's voice on the recording).

> I mean, they just, they get *zoned in* on number, if they *see two numbers the same*, and don't necessarily read the bits that go around it, cause it's two numbers the same, you know. So that's a very mathematical thing and *this is maths, this is about number*.… *Not about words.*… And I guess that's a struggle too, when you're teaching statistics and understanding is, that children don't want to do the reading and the thinking and the writing cause that's English or

> SOSE [Studies of Society and the Environment] or something. *That's not Maths. Maths is sums*.

We are not attempting to present this language aspect as evidence in itself, but we were sensitive to its inevitable influence and endorsed it as such within our qualitative analysis. It gave a context to the research that was both essential and yet difficult to acknowledge formally. A qualitative analysis allows the language of the interviewee to become a part of the teacher profile in a nuanced way. It indicates content knowledge, capacity to conceptualise, life experience, professional confidence, and relationship with students. And of course our role as interviewers and analysts is linked to a set of values depicted by the subtleties of language. To ignore its contextual reality is to deny our intrusion as analysts, and yet to abstract core meaning from a reliance on language was in this instance, difficult to justify. We also acknowledge that some teachers are naturally more expansive in their verbal expression than others.

All 40 of the teachers interviewed volunteered at least one comment that had relevance to these three themes of Teaching Practice, Curriculum Values, or Cognitive Experience, and more than half expressed views that related to more than one. The comments were largely oppositional, with the teaching of statistics being identified as a positive experience and the teaching of mathematics as negative. The themes themselves are now explained.

Teaching Practice Twenty teachers made comments that were classified as related to teaching practice. There were three different ways in which the teaching practice of statistics was distinguished from that of mathematics. First of all, a lively classroom dynamic was identified by 15 teachers. There was more "action" in the classroom, and more fun: "[You like best] Hands on, out of your seat, asking questions, moving, it's a bit of fun, it's interesting to them, it's relevant to them, particularly if you ask the right questions" (T33). This element of "fun" was particularly associated with the StatSmart school experience of the software *TinkerPlots*, developed by Key Curriculum Press (Konold & Miller, 2005): "[Students like best] They also, they are getting a lot of fun out of the *TinkerPlots*" (T18). Related in part to the use of this computer program, was the specific observation by three teachers that students found the content more accessible owing to an enhanced visual element: "When I teach statistics I like to have lots of visual stuff around the classroom, recording what they're doing. And it really is, it becomes a really wonderful way, that students can actually display their learning and then give them an opportunity to talk about it as well." (T6). Added to this picture of a more active classroom, was a more collaborative one suggested by four teachers: "Well I guess there's an element of socialising in it. I mean they usually work in groups … very rare to work independently in statistical things" (T29). Interestingly, it was observed by three teachers not only that more students overall were involved, but also that students not usually considered to perform well in mathematics had a greater role in the statistics discussions. In summary, the teaching practice of statistics was perceived as more active, more fun, and more collaborative than that of mathematics.

Curriculum Values Overall 23 teachers made comments that were judged to be related to curriculum values. Four teachers observed that in statistics students considered to be "low" achievers in mathematics generally had greater opportunity to contribute positively: "They [strong students] don't seem to dominate the group. And sometimes it's the weaker ones who actually have the more insight. It's quite interesting what they come up with" (T7). This notion of equity overlaps with the other themes that opposed a "mechanical" mathematics with a more contextual and contested statistics. The language of these comments suggested the persistence of a subject hierarchy in the curriculum, although whether this bias is held by the teachers or not is unclear. Issues of the context for presenting statistics and its relation to the rest of the school curriculum were uppermost for 19 of the teachers. There was some confusion expressed by eight teachers for where statistics should exist in the curriculum, of whether it should be aligned with "hard" mathematics or with "soft" English and social science. Teachers identified the curriculum challenge in statistics as the contextual interpretation of data that requires a level of conceptualisation identified with the humanities:

> I think about, it's a real world application and the fact that there's so much in their lives that they don't recognise as being statistics and it just, I know in the little bit of work we've done now, they're already sort of twigging to things they wouldn't have done before. (T36)

Statistics, unlike mathematics, was imagined across the curriculum. For some teachers this created a certain curricular discomfort as the boundary between mathematics and the humanities became blurred, but for others, the crossover was welcomed: "One of my very bright little ones, said but this isn't maths, when are we going back to ordinary maths, when I started; they don't see it as maths … it crosses [over] to me, maths and SOSE anyway. It's a very strong [link], it's not a pure maths thing" (T38). Also related to the cross curricular work was the implication of two teachers that statistics was practical, more concrete than abstract: "Well the only difference is it is not actually abstract, it is real. And so from that point of view it makes for more interesting maths because it is in fact something that they have actually—it is their stuff, like it is their statistics, so there is ownership there" (T19). This quality of social and personal relevance is also demonstrated by three teachers, including the following:

> And they really like to be able to see that it's got a real life application and they're much happier if they, if we can give them, you know, the majority of your, that kind of maths, so that there is a context to it that they can see it's important. (T35)

Cognitive Experience Twenty-two of the teachers interviewed explicitly mentioned the appeal of teaching statistics as promoting lively classroom discussion and the enhanced critical thinking by students. "Discussion" was identified as a novelty, and interestingly, in response to a later interview question, a few teachers considered their own "general" knowledge as being important in the teaching of statistics because there is such an emphasis on real world context: "There's more than one way of looking at that [statistical question] because of the way it's worded. So that sort of … the literacy of the maths is where I consider myself a lot stronger than, you know, the mathematical concept" (T4). It is this potential for situating statistics in a real world context that teachers believed works so well. The purpose of statistics could be demonstrated to students in a way that was contrasted with other mathematics: "In other words if you produce a concept and get the numbers and do the maths correctly then you can actually persuade people to move in a particular direction" (T22); "I think, like statistics is the easiest of all the maths to teach in terms of getting them to see where it comes from … they definitely understand statistics more than fractions, that's for sure" (T13). By experiencing statistics in the context of a wider argument there was greater opportunity for students' knowledge to be understood by them as integral to their values and purpose and not objectified as a box of content. Seven teachers expressed comments that touched on this theme of the students' relationship with knowledge being deepened. It was about students understanding how knowledge was constructed, and about their own power in that process: "And also being able to pose questions, they loved that. That sort of gave them control over it in a way that we have never really had before" (T26). These teachers expressed the students' sense of excitement, of real engagement. One teacher expressed this in terms of ownership: "It makes for more interesting maths because … it is their stuff, like it is their statistics, so there is ownership there" (T19). This was contrasted with algebra "which is not necessarily what they are into."

The following longer quote from a teacher in a boys' school sums up many teachers' views on the difference in the cognitive experience for students in relation to mathematics and statistics:

> … when maths is quite often taught in a stand-alone discrete way and quite often you have boys that are really, really able to do certain things in maths but have actually no understanding of what they do … they were expecting there to be a right or wrong [answer], they weren't expecting to actually be encouraged to see in between the lines … huge opportunities to borrow maths and apply it to something [that] was actually significant in the boys' learning … [Later] … generally the very notion that there are shades-of-grey between, and more than one way of actually looking at statistics is actually, completely new to most of them. Maybe one or two out of a class of 20, 22 boys were able to quickly see, oh hang on, this depends upon how you approach this question or looking at this set of data. Quite often they were looking for a yes or no, or what's the right answer and what's not. So that was a real challenge and continues to be that. (T4)

Overall 16 teachers made comments about a single theme used to cluster responses whereas 7 made comments related to all three. The Venn diagram in Figure 1 shows the number of teachers commenting in relation to the themes.

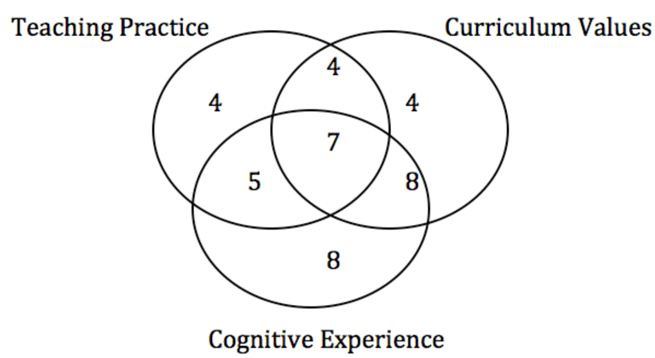


Figure 1. Numbers of teacher responses related to the three identified themes

3.2 REVIEWING THE STATISTICAL PCK EXPOSED IN THE THREE PROBLEMS IN THE INTERVIEW

After identifying the significance of the perceived difference between teaching mathematics and statistics we moved to reviewing the statistical PCK discerned in the teachers' responses to the three problems in the interview. We asked ourselves whether the differences in the teachers' statistical PCK could be related in part to the level of mathematics inherent in the three statistical problems.

Tom, Marijuana, and Lung Disease had encompassed the required range of statistical problems necessary for our prime research interest in PCK. When reviewed in the light of the supplementary theme of a perceived mathematics-statistics difference, the problems also held some unforeseen qualities. These qualities can perhaps best be described in terms of the dominant language of the problems. For Tom, the delivery was largely pictorial, for Marijuana it was textual, and for Lung Disease it was numeric. The degree of mathematical calculation required to think statistically escalated from marginal in Marijuana, to helpful in Tom, and to critical in Lung Disease. Our impression on reading the interviews was that most teachers struggled in responding to the Lung Disease questions. The hesitation of many with the mathematical calculation required was palpable. This in turn hampered their capacity to articulate the “big ideas,” to be convincing in their teaching strategy, and to abstract general principles. It seemed that as linguistic symbols, pictures and words were more conducive to statistical thinking than numbers. Pictures and words constituted a language that teachers and students shared with relative ease. The function of mathematics in enabling a higher order of statistical thinking was *not* well demonstrated by most of the interviewed teachers. Quite the opposite, mathematical calculations seemed to sabotage the statistical thinking, and by implication, the classroom teaching.

Having discerned an impressionist view of the teachers’ PCK using the four components already described, we were anxious to reconsider the rubric for these core questions in the interview. For the purposes of this paper, we shortcut to a series of graphs that do not in themselves constitute a complete measure of PCK, but do give relative weighting to the individual components across the three problems. Forty teachers responded to the Tom and Marijuana problems, 29 to Lung Disease. The other 11 teachers, primary teachers (Years 5 and 6), answered a different two-way table problem, considered originally to be easier than the Lung Disease version. It was felt that including these 11 teachers in the current analysis would bias the results as the teachers may not have had exposure to the mathematics involved in the problem. In the following graphs the results for the four parts of the three statistical contexts are presented as percentages of the number of teachers responding. Hence for each context (Tom, etc.) and each component the percentages sum to 100 percent. Figure 2 shows the distribution of response codes for each of the four rubrics: Big Ideas, Anticipates Student Answers, Content-specific Strategies, and Shift to General (details for the codes of the rubrics are in Appendix B).

More teachers found Marijuana, the problem that involved the least mathematics and was primarily delivered via text, the easiest to reduce to a learning idea. Next was Tom, then, by a small margin, followed Lung Disease. We expected a higher percentage of teachers to identify the Big Ideas in Tom. Perhaps owing to its placement as first of the three problems, teachers tended to jump into the detail of the response and to grasp the notion of uncertainty only as the interview progressed. This is evident in the second graph, which captures teachers’ capacity to anticipate student answers. A significantly greater percentage of responses scored level 2 for Tom and Marijuana, than Lung Disease. To anticipate a range of inappropriate and appropriate student answers does require a fair degree of confidence with the problem at hand. Contrary to this is the higher rating 3s for Lung Disease. When constructing the rubric it was decided to award a 3 if the response demonstrated the appropriate mathematical calculation required to access the interpretation of the two-way table. A small but notable number of teachers could complete the calculation, and therefore scored level 3. We wished to determine how many teachers were confident enough with the language of mathematics to enter the subtleties of the statistical problem. The almost 30 percent of teachers who attained a level 3 knew the mathematics of this type of problem well, although that same third did not necessarily appreciate the full potential of the problem. This point illustrates one of the many discussion points we negotiated as researchers, as well as the inevitable compromise that almost any rubric holds, especially those relying on qualitative analysis.

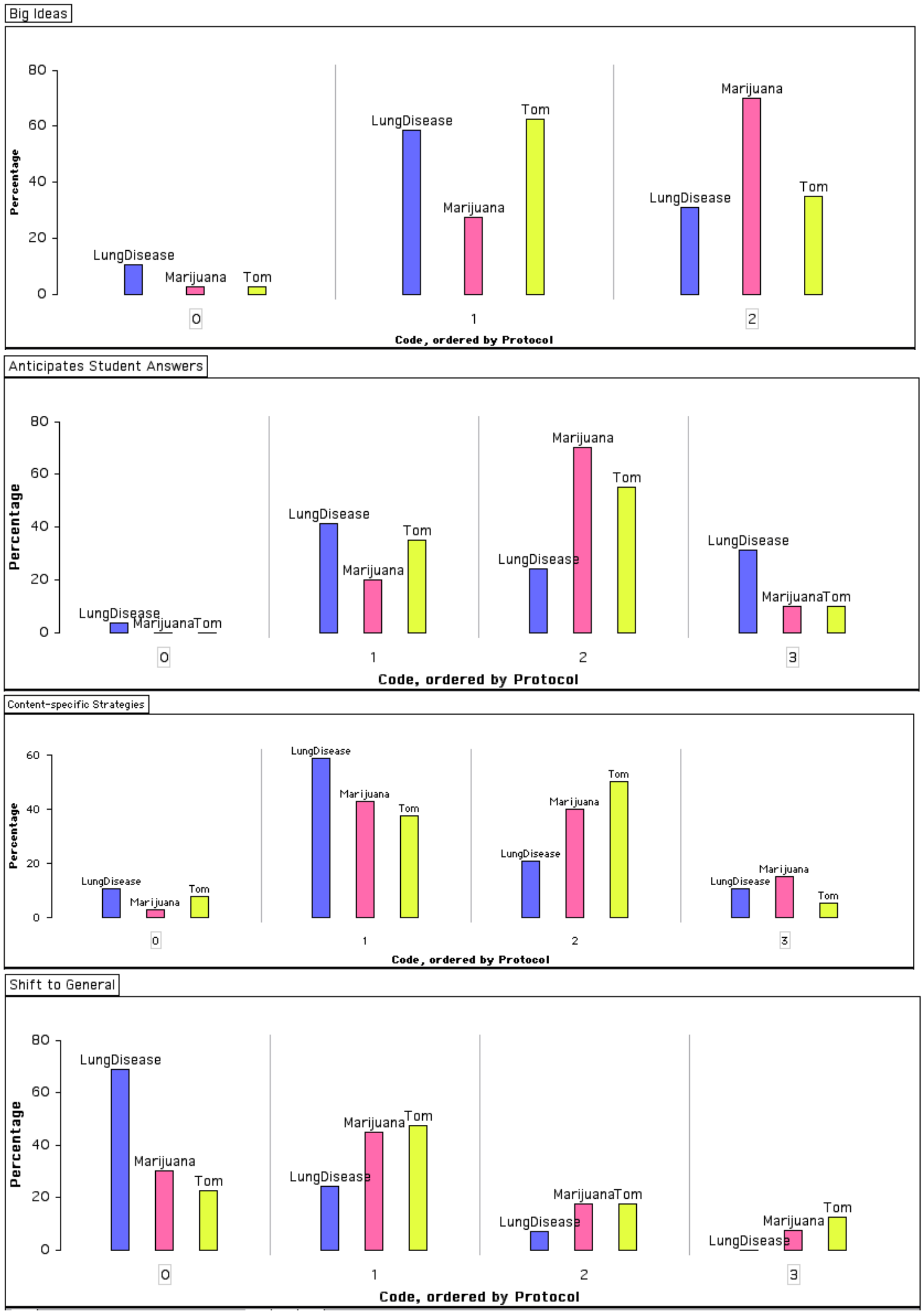


Figure 2. *Percentage of responses for each code for Big Ideas, Anticipates Student Answers, Content-specific Strategies, and Shift to General (the sum for each context—Tom, Marijuana and Lung Disease—is 100%)*

Although moving from surveys to interviews provided us with greater opportunities to dissect the teaching process, we remained frustrated at times with the brevity of responses by teachers. An "unevenness" of responses also emerged. If a teacher gave a full response to one problem, or to one of the sub-questions of the problem, then there was sometimes a more lightweight subsequent response, as if in compensation. We took measure of this pattern, erratic as it was, by rating these last components—Content-specific Strategies and Shift to General—across the two (Lung Disease) or three (Tom, Marijuana) sub-questions of the problem. A reading of the graph depicting Content-specific Strategies does support the view that teachers found it somewhat easier to communicate with students when the statistical problem was delivered more in text or pictures than numerals (especially codes 1 and 2). There does not seem to be any other factor besides the mathematics component that explains the relative difficulty teachers have in structuring a purposeful discussion about Lung Disease compared with the other problems. For the three contexts, however, only a slim percentage of teachers could demonstrate great creativity in their teaching approach (code 3). Likewise, few teachers exhibited the ability to shift from the specifics of the problem to its general conceptualisation. This is especially so for Lung Disease, where no response was coded as 3. Although the teachers also struggled to achieve a code of 3 on the other two contexts, the dramatic representation for Lung Disease in relation to the other two tasks (69% scoring 0) indicates that it was the most taxing of the three problems for the teachers to move beyond the mathematics in the classroom. Of the nine teachers who achieved code 3 for Anticipates Student Responses, five were coded 0, two were coded 1 and two coded 2 for Shift to General, which was only marginally better than the performance of the teachers who scored less than 3 for Anticipates Student Responses.

An overall perusal of Figure 2 shows the increasing struggle teachers had with PCK across the four components, with a trend downward in the codes. It is not possible to produce a measure of association between PCK and teachers' perceptions of teaching mathematics and statistics because the perceptions varied in ways that could not be numerically scaled. It is our belief, however, that overall the perception of difference in teaching mathematics and statistics is related to the observed differences in PCK across the tasks in contexts that reflected different levels of mathematics in their makeup. Of the four components of PCK, averaging the codes as an indicative measure only, Lung Disease was the most difficult on all four, whereas Tom was only easier than Marijuana on Shift to General. Overall, we might claim Marijuana with the least mathematical content was the easiest, then Tom, then Lung Disease. We speculate further on this relationship in the Discussion.

4. DISCUSSION

Qualitative analysis of the teacher interviews alerted us to the unexpected theme of perceived difference in the teaching of mathematics and statistics. This theme is of interest in its own right, but when considered as a new dimension in our primary research goal of formulating statistical PCK, it developed significance by association. It became a fresh filter for the researchers to view PCK as described in the four components summarized in the Results. The three interview contexts were recast as problems with a gradation of mathematics, increasing from Marijuana to Tom to Lung Disease, which was central to the teachers' display of PCK. This new theme of difference drew us to suggest that overall, the less mathematical content in the task the stronger the exposition of teachers' PCK in achieving statistical goals. This may not be a totally unexpected outcome given the results of Watson and Callingham's (2003) study of students' understanding of statistical literacy. That work, based on Rasch analysis (1960/1980), produced a six-level hierarchy of understanding from Idiosyncratic to Critical thinking, where the highest two levels of Critical thinking were distinguished mainly by the ability to employ proportional reasoning in the solutions to problems. We hypothesise that a similar difficulty arises for teachers, some of whom are themselves lacking in confidence in relation to their proportional reasoning content knowledge and how to transform it to enable their students to understand and apply it. Proportional reasoning was the mathematical component most difficult for students to incorporate into their critical analysis of questions and it appears that this same mathematical content may be the most difficult for teachers.

An appreciation of this theme impacted on our analysis of the teachers' PCK performance in another respect. It was clearly disconcerting for some teachers that the curricular status of middle school statistics was so ill-defined. The debate as to whether statistics holds credentials as a sub-discipline of mathematics or as a separate epistemological entity is a long-standing one (e.g., Cobb & Moore, 1997). Our research suggests that irrespective of the wider academic polemic, secondary school practitioners who most usually share the teaching responsibility of both curriculum domains do perceive a difference. The most common adjective applied to mathematics was *mechanical*. It was considered, by implication, to have qualities opposite to statistics. The term that perhaps best summed up the notion of difference was *critical thinking*. Statistics was seen to encourage students' critical thinking. This finding may reveal as much about the educational status of mathematics as it does statistics. Clearly the interviewed teachers did not appreciate the disciplinary problematic of mathematics, one demonstrated with great craft in the narrative of primary teaching experience by Magdalene Lampert (2001). This profile of school level mathematics as formulaic, uncontested knowledge is of importance here for the contrast it presents and is not a theme pursued in its own right.

The issue of perceived difference does have implications for curriculum policy in secondary schools where inter-disciplinary boundaries are more clearly set than in primary schools. One of the interviewed teachers expressed irritation that the level of discussion required to interpret statistics embedded in a social issue was not appropriate in a mathematics class. It was, he suggested, more appropriate fare for a teacher of English. By contrast, many of the teachers confessed to a kind of guilty pleasure in being able to enjoy discussions in a classroom normally devoid of them. They revelled in the idea of asking questions without definitive answers. The notion of *uncertainty*, so fundamental to statistics, was equated by teachers with *discussion* and *critical thinking*. It was this essence of controversy, the "shades of grey" according to one teacher, that offered a distinctive flavour to the teaching of statistics.

The provocative nature of statistics, from primary to university level, is firmly grounded in the context of real problems. Three curriculum positions regarding the "real-life" contextual significance of statistics became apparent from a reading of the interviews. Context can be of uppermost importance and statistics are presented as a central and often contested strand of evidence; context can be minimised or eliminated as a distraction to the mathematical acrobatics of abstraction that is statistics; and finally, context can dominate to such an extent that the mathematical validity of the statistics moves out of focus. Statistics, we believe, has the potential to manifest itself as a specialist subject that, although not abandoning "context," acknowledges mathematics as its essential language. This language of mathematics appears to us to emerge from the language of pictures and words, and becomes significant over the middle school years. Statistics also has a role in the general curriculum where a multiplicity of textual, visual, and mathematical representations can co-exist with broad social enquiry. For a teacher, the dedicated curriculum location of statistics at either the generalist core or at the specialist fringe, or ideally at both, would clarify what at the moment appears to be creating some confusion. If the disciplinary niche for school statistics could be asserted more forcefully then perhaps the teachers would have greater confidence to develop appropriate rationale and effective practice.

The larger issue of the place of statistics in national curricula around the world is more likely to reflect a historical perspective than the dilemma exposed by the teachers in the StatSmart project. In the United States, the National Council of Teachers of Mathematics places Probability and Statistics as a sub-domain of mathematics in its *Principles and Standards for School Mathematics* (2000), whereas the New Zealand curriculum grants the two areas roughly equal status in introducing the subject Mathematics and Statistics in its school curriculum (Ministry of Education, 2007). Advanced courses, such as AP Statistics in the US (College Board, 2008) and GCE A-level Statistics in England (Assessment Qualification Alliance, 2009), separate statistics from mathematics, but this is not done for example in Australian state or national curricula. The dilemma of how to situate statistics within an overall curriculum framework is far from resolved. In the meantime, the ambivalence of its curricular status would appear to be having very real repercussions in our middle school classrooms.

We would like to suggest a rather complex topic for further research arising from our analysis. We believe the uneasy relationship between mathematics and statistics that has evolved in recent years, as topics such as *chance* and *data* have filtered down to younger students who may not yet possess proportional reasoning skills, has resulted in a misconception about statistics as an entity separate to mathematics. When statistics was only taught at tertiary level as a theoretical subject, it was totally dependent on high level numerical skills from mathematics—this was never questioned. As statistical thinking has been introduced at the school level, the same numerical skills cannot be assumed and the underlying concepts related to sampling, data representation, data reduction, and informal decision-making have come to the fore. For teachers without the tertiary training in statistics (and the associated mathematics) these other aspects of statistics are appealing, especially as seen here, through the introduction of context and cross-curriculum opportunities. For the primary grades this content can hold its own without a high level dependence on numerical skills. This is seen in contexts like Marijuana and Tom, where teachers appear more at ease with teaching the statistical ideas through the language of words and pictures. Teachers who exhibit a robust statistical PCK can struggle when higher-level mathematical skills, such as proportional reasoning, are required in later years and in tasks such as Lung Disease. What we believe we have observed are teachers caught up in a curricular tug of war. Pulling with greatest strength is the new curriculum with its teaching demands of real world application and personal relevance for students who have fewer mathematical skills. It has been an exciting experience for teachers, as shown in their comments, but has provided challenges to their statistical PCK. It also provides challenges for curriculum developers.

As must often be the case when qualitative methods are used by researchers, unforeseen factors arise when detailed analysis takes place. This can be seen as both a benefit and a detriment. In the case of this study there was not a predetermined question in the interview, “Do you perceive there is a difference between teaching statistics and teaching mathematics?” Had there been such a question, some teachers might have emphatically said “no,” whereas others who might have been somewhat equivocal could have provided more “yes” examples than we actually observed. Although this issue is acknowledged as a weakness of the study in the sense of “missing data,” we also see it as a strength in the sense of being able to observe teachers’ unsolicited reactions to the teaching of statistics from their positions as teachers of mathematics. We feel that discovering these perceptions and consequent dilemmas for teachers without prompting provides evidence that warrants further research designed specifically to target the validity and significance of this "difference" theme.

ACKNOWLEDGEMENTS

This research was funded by Australian Research Council grant number LP0669106, in partnership with the Australian Bureau of Statistics, Key Curriculum Press, and the Baker Centre for School Mathematics. Thanks to our co-investigator, Rosemary Callingham, and to Julie Donne and Elena Zema, who conducted some of the teacher interviews.

REFERENCES

Assessment Qualification Alliance. (2009). *GSE A-level statistics.* UK: Author. [Online: http://web.aqa.org.uk/qual/gce/maths/stats_noticeboard.php?id=01&prev=01]

Bouma, G., & Ling, R. (2004). *The research process* (5th ed.). South Melbourne, Victoria: Oxford University Press.

Callingham, R., & Watson J. M. (2008). Overcoming research design issues using Rasch measurement: The StatSmart project. In P. Jeffery (Ed.), *Proceedings of the 2007 Australian Association for Research in Education annual conferenc*e. Victoria: Australian Association for Research in Education. [Online: http://www.aare.edu.au/07pap/cal07042.pdf]

Chick, H. L. (2007). Teaching and learning by example. In J. Watson & K. Beswick (Eds.), *Mathematics: Essential research, essential practice. Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 3–21). Adelaide, SA: MERGA

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly, 104*(9), 801–824.

College Board. (2008). *AP Statistics course description.* New York: Author. [Online: http://apcentral.collegeboard.com/apc/public/repository/ap08_statistics_coursedesc.pdf]

Hill, H. C., Rowan, R., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Konold, C., & Miller, C. D. (2005). TinkerPlots: Dynamic Data Exploration [Computer software]. Emeryville, CA: Key Curriculum Press.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Ministry of Education. (2007). *Draft mathematics and statistics curriculum.* Wellington, NZ: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press. (Original work published 1960)

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Watson, J. M., & Callingham, R. A. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3–46.

[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ2%282%29_Watson_Callingham.pdf]

Watson, J., Callingham, R., & Donne, J. (2008a). Establishing PCK for teaching statistics. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICME/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and the 2008 IASE Round Table Conference.* Voorburg, The Netherlands: International Association for Statistical Education.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/rt08/T2P10_Watson.pdf]

Watson, J., Callingham, R., & Donne, J. (2008b). Proportional reasoning: Student knowledge and teachers' pedagogical content knowledge. In M. Goos, R. Brown, & K. Makar (Eds.), *Navigating currents and charting directions. Proceedings of the 31st annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 555–562). Adelaide, SA: MERGA.

Watson, J., Callingham, R., & Nathan, E. (2009). Probing teachers' pedagogical content knowledge in statistics: "How will Tom get to school tomorrow?" In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides* (Vol. 2, pp. 563–570). Palmerston North, NZ: MERGA.

Watson, J., & Nathan, E. (2010a). Assessing the interpretation of two-way tables as part of statistical literacy. In C. Reading (Ed.), *ICOTS-8 Proceedings: Towards an evidence based society.* Voorburg, The Netherlands: International Association for Statistical Education, International Statistical Institute.
[Online: www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_5E1_WATSON.pdf]

Watson, J., & Nathan, E. (2010b). Biased sampling and PCK: The case of the marijuana problem. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education. Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 610–617). Fremantle, WA: MERGA.

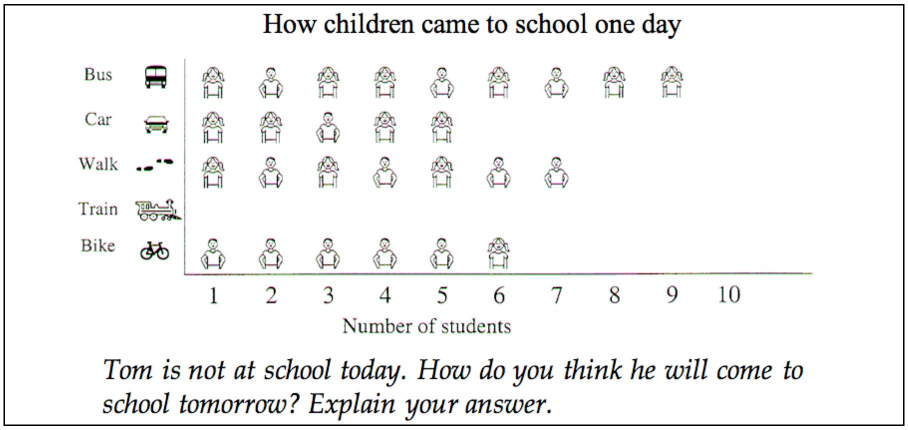
JANE M. WATSON
Faculty of Education, University of Tasmania
Private Bag 66
Hobart, Tasmania 7001
Australia

APPENDIX A: STATSMART TEACHER INTERVIEW PROTOCOL

Initial Question

Q1. What topics in statistics do your students have most difficulty with and why?

Use "Tom." Show the problem.



Q2T. What are the big statistical ideas in this problem? (Probe: What answer would you give?)

Q3T. Please can you give an example of an appropriate response and an inappropriate response that your students might give. (Probe: Can you explain why it is appropriate/ inappropriate?)

Q4T. What opportunities would this problem provide for you teaching? (Probe: Where would you place it in your lesson sequence? Or in your school's curriculum sequence?)

Show a student response: Bike, because the majority of boys ride to school.

Q5T. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Repeat the sequence. Show a student response: Tom will come to school by train because there is no-one next to the train so it must be him.

Q6T. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Repeat the sequence. Show a student response: Boy, because there is a pattern and the next one is a boy.

Q7T. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Use "Marijuana." Show the problem.

Decriminalise drug use: poll

SOME 96 percent of callers to youth radio station Triple J have said marijuana use should be decriminalised in Australia.

The phone-in listener poll, which closed yesterday, showed 9924 - out of the 10,000-plus callers - favoured decriminalisation, the station said.

Only 389 believed possession of the drug should remain a criminal offence.

Many callers stressed they did not smoke marijuana but still believed in decriminalising its use, a Triple J statement said.

Is the sample reported here a reliable way of finding out public support for the decriminalisation of marijuana? Why or why not?

Q2M. What are the big statistical ideas in this problem? (Probe: What answer would you give?)

Q3M. Please can you give an example of an appropriate response and an inappropriate response that your students might give? (Probe: Can you explain why it is appropriate/inappropriate?)

Q4M. What opportunities would this problem provide for you teaching? (Probe: Where would you place it in your lesson sequence? Or in your school's curriculum sequence?)

Show a student response: Yes, because 10,000 people is enough to get an accurate average of the view of the public.

Q5M. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Repeat the sequence. Show a student response: No, because it is not everyone in Australia voting.

Q6M. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Repeat the sequence. Show a student response: No, because some people could be lying.

Q7M. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Use "Lung disease/Smoking." Show the problem.

The following information is from a survey about smoking and lung disease among 250 people.

	Lung disease	No lung disease	Total
Smoking	90	60	150
No smoking	60	40	100
Total	150	100	250

Using this information, do you think that for this sample of people lung disease depended on smoking? Explain your answer.

Q2L. What are the big statistical ideas in this problem? (Probe: What answer would you give?)

Q3L. Please can you give an example of an appropriate response and an inappropriate response that your students might give? (Probe: Can you explain why it is appropriate/inappropriate?)

Q4L. What opportunities would this problem provide for you teaching? (Probe: Where would you place it in your lesson sequence? Or in your school's curriculum sequence?)

Show a student response: Yes, 90 who smoked got lung disease.

Q5L. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Repeat the sequence. Show a student response: 60 "no smoking lung disease" and 60 "smoking no lung disease: are the same.

Q6L. A student gave this answer. How would you move this student's understanding forward? (Probe: What would be the next step in learning?)

Final questions

What's the best thing about teaching statistics?

What do your students think is the best thing about learning statistics?

What contexts are they most interested in?

How do you stimulate student interest?

What knowledge do you have that you draw on when teaching statistics? You might consider statistical knowledge, teaching approaches, planning skills, understanding students, or anything else you think is relevant.

APPENDIX B: PCK RUBRIC FOR TOM, MARIJUANA AND LUNG DISEASE

PCK Feature 1: Recognises Big Ideas (for all three contexts)

Code/Level	Description
0	Response confused and /or incorrect
1	Response implied and/or understanding revealed beyond initial question
2	Immediate grasp of idea, language specific

PCK Feature 2: Anticipates Student Answers (for all three contexts)

Code/Level	Description
0	Response irrelevant
1	Appropriate or inappropriate but not both, or unclear
2	Distinguishes both appropriate and inappropriate
3	Demonstrates understanding of students' reasoning

PCK Feature 3: Employs Content-specific Strategies

Code/Level	Description		
	Tom	Marijuana	Lung Disease
0	Response absent or indicates misleading content, not highly relevant	Response absent or indicates misleading content	Response absent or indicates misleading content or not highly relevant
1	Content knowledge of probability and/or data representation requisite to initiate a discussion	Content knowledge of sampling requisite to initiate a discussion	Content knowledge of proportion requisite to initiate a discussion
2	Demonstrates questions or knowledge that might structure a discussion about probability and/or data representation	Demonstrates questions or knowledge that might *structure* a discussion about sampling	Demonstrates questions or knowledge that might structure a discussion about proportion and/or statistical variables (causation, independence)
3	Extends discussion by illustrating/referencing beyond the pictograph	Extends discussion by illustrating/referencing beyond the marijuana survey	Extends discussion by illustrating/ referencing beyond the table

PCK Feature 4: Constructs Shift to General

Code/Level	Description		
	Tom	Marijuana	Lung Disease
0	No shift to general evident	No shift to general evident	No shift to general evident
1	Considers elements of data representation in general terms (e.g., mechanics of data collection, limitations, clarity)	Considers elements of sampling design in general terms (e.g., size changeable with purpose; profiling of sample population; census vs. sampling; accounting for invalid responses; social sensitivity)	Awareness of context: popular beliefs of strong causation intruding on statistical interpretation; numbers versus real scenarios
2	Extrapolates from pictograph to consider one or more statistical concepts (e.g., the uncertainty of prediction, the majority, principles of proportional reasoning)	Extrapolates from survey to consider one or more statistical concepts (e.g., random, representation, average)	Discusses different representations of proportion (percent, ratio, numbers) and/or considers subtleties of application
3	Distinguishes between the pictograph as a statistical model and a representation of real data and/or introduces an awareness of language	Relates survey construction to wider context of argument and/or introduces an awareness of language (e.g., lying, public)	Explores the principles of association between variables; introduces an awareness of language (independence, variable in statistical context)