A Day in the Life of a Statistical Knowledge for Teaching Course

Randall E. Groth

Salisbury University, Salisbury, Maryland, USA

Email: regroth@salisbury.edu

Summary: Teachers' statistical learning needs differ from those of individuals in other professions. Along with learning statistical content, they must develop the ability to teach statistics to others. This paper illustrates how both needs can be addressed simultaneously in an undergraduate course grounded in a statistical knowledge for teaching framework. A typical day in the course is presented to demonstrate application of the framework. It is also suggested that practicing teachers can benefit from using the framework to set their own professional development goals.

Keywords: Statistical knowledge for teaching; Teacher preparation; Undergraduate statistics

STATISTICAL KNOWLEDGE FOR TEACHING

The knowledge needed by teachers differs sharply from the knowledge needed by other professionals. Like other professionals, teachers need to know statistical content, but they also need to know pedagogical strategies for helping others learn statistics. Adding to the complexity of the issue, content knowledge and pedagogical knowledge are not always neatly separable. Shulman (1987) coined the phrase "pedagogical content knowledge" to describe a "special amalgam of content and pedagogy that is uniquely the province of teachers, their own form of professional understanding" (p. 8). If teachers are to realize their full professional potential, they must attend to the development of pedagogical content knowledge along with subject matter knowledge.

This paper illustrates how statistical subject matter knowledge and pedagogical content knowledge can be addressed simultaneously within the context of an introductory statistics course for prospective teachers. It offers a glimpse of a typical day in a course I designed and taught that was guided by a framework of teacher knowledge being developed by the Learning Mathematics for Teaching (LMT) Project (Hill, Ball, & Schilling, 2008). Within the framework, there are three primary components of subject matter knowledge and another three of pedagogical content knowledge. Subject matter knowledge is hypothesized to consist of common content knowledge, specialized content knowledge, and horizon knowledge. Pedagogical content knowledge is hypothesized to consist of knowledge of content and students, knowledge of content and teaching, and curriculum knowledge. Although this paper describes an undergraduate course, practicing teachers can also benefit from analysing their own knowledge structures in terms of the LMT framework. Becoming familiar with the framework can help teachers reflect on their own knowledge and identify suitable professional development goals.

Subject matter knowledge is closely related to the content one would expect to encounter in a conventional introductory statistics course, yet some points of departure exist. Common knowledge consists of concepts and procedures needed by teachers as well as other professionals. For example, teachers and others have a need to understand descriptive and inferential statistics and their potential uses. Horizon knowledge is closely related to common knowledge, since consists of content one might encounter in a conventional introductory statistics course. Horizon knowledge allows teachers to understand how the content they teach fits within the broader discipline of statistics. Specialized content knowledge marks a departure from the content of conventional introductory statistics courses, since it pertains to knowing representations likely to help children learn and being able to appraise strategies invented by young students. Watson, Fitzallen, Wilson, and Creed (2008), for instance, described a statistical representation called a "hat plot," a simplified version of a box plot that prepares children to understand the structure of condensed statistical representations. Knowing of such a statistical representation is of interest to teachers, but not necessarily to other professionals.

Pedagogical content knowledge further distinguishes the knowledge needed for teaching from the knowledge needed in other professions. Knowledge of content and students involves being able to diagnose students' thinking and recognize common errors. Teachers with knowledge of student errors and misconceptions are in position to address them in their instructional plans. Knowledge of content and teaching involves having a wide range of strategies to draw upon to facilitate students' learning. Knowledge of curriculum allows teachers to identify curricular sequences likely to enhance student learning.

A summary of how the activities for the class session to be described in this paper align with the LMT framework is provided in Table 1. It should be noted that the LMT framework is a theory of *mathematics* teacher knowledge, and hence does not explicitly deal with important non-mathematical elements of the discipline of statistics (Groth, 2007). Some of the activities to be described helped develop primarily non-mathematical knowledge elements required for teaching statistics (e.g., issues in survey design).

<INSERT TABLE 1 HERE>

HOMEWORK LEADING INTO THE CLASS SESSION

On the day of the class session, a homework assignment was due that required reading an article and responding to an accompanying set of questions. The article described the common error of attempting to compute a median for nominal categorical data and offered teaching strategies designed to address it (Leavy, Friel, & Mamer, 2009). The article pointed out that when children work with nominal categorical data, they may attempt to arrange nominal categories in alphabetical order to determine a median or may compute the median of category frequencies. To steer children away from such strategies, the article suggested prompting children to analyse their own thinking. For instance, in one case in the article, when an even number of nominal categories existed and “fish” and “bird” were the two middle categories alphabetically, children were prompted to notice that the median would have to be a nonsensical category such as “fird” (a combination of the two category names). The assigned article and accompanying questions were intended to build knowledge of content and students by drawing attention to children’s thinking patterns and knowledge of content and teaching by suggesting a potential strategy for dislodging misconceptions.

My students were required to submit responses to the assignment online before class, so I read their work while planning the lesson. In reading their responses, I noticed that most were able to write original examples of how children might errantly alphabetize category *labels* to determine a median. However, when asked to provide their own examples of how children might order category *frequencies* to determine a median, many instead described the mistake of ordering numerals that serve as nominal labels (e.g., zip codes, uniform numbers). Since we had previously discussed the latter type of error in class, it was not surprising they had it in mind. However, it was surprising that they were not able to recognize the "new" error of ordering category frequencies described in the article and explain how it might surface in classrooms. Additionally, when asked to describe how they would go about teaching children that the median is not suitable for categorical data, many relied on very directive approaches, opting to simply tell children it is not suitable rather than leading them to recognize errors in their own thinking.

To provide another opportunity to study the student misconceptions and teaching strategies described in the assigned article, I had my students discuss a case of classroom practice from the *Developing Mathematical Ideas* (DMI) series (Russell, Schifter, & Bastable, 2002) to begin class. The DMI case described a situation where children analysed data from a classroom survey. The data consisted of responses to the question, "What is your favourite place to swim?" The teacher in the case encouraged children to find useful ways to represent and organize the responses. The children produced rudimentary bar graphs by stacking data cards into categories. A great deal of discussion occurred about how condensed categories should be (e.g., "the high school pool" and "Kellerman Pool" were originally in separate categories, but later merged). Considering the case gave my students another opportunity to think about teaching strategies that guide children to analyse their own thinking. I also prompted them to discuss whether it would be useful to compute the mean or median of the category frequencies for the data set described in the case. Discussing these ideas allowed those who initially missed them in the assigned homework reading to revisit them in the context of a class discussion.

CONSTRUCTING A CLASS SURVEY

After they read and discussed the DMI case, I led my students in constructing and administering their own class survey. Curriculum documents such as the *Guidelines for Assessment and Instruction in Statistics Education PreK-12 Report* (Franklin et al., 2007) recommend that children conduct classroom surveys when first learning statistics. Hence, the activity was intended to help develop knowledge of content and teaching by providing a chance to experience a teaching strategy similar to the one described in the DMI case. My students worked in small groups to design questions to ask the entire class. I specified that their questions should generate both categorical and quantitative data (providing another opportunity to develop common knowledge of the distinction between the two data types). Students were then asked to contribute their favourite questions to a class survey that everyone would take.

As we compiled questions, a number of survey design issues emerged. The first issue was about inclusiveness of language. One group wanted to ask, "How warm do you keep your dorm room at night?" Others objected, since some students lived in apartments rather than dormitories. Still others, who lived in houses, pointed out the need to make the question apply to an even broader group. Eventually the class agreed on asking "How warm do you keep your bedroom at night?" Another issue was the manageability of data generated by survey questions. One group wanted to ask "What do you do on a Friday night?" Although others agreed it would be an interesting question, concerns were raised about how to summarize responses to it. Ultimately, they agreed to leave the question on the survey but to give short, one-word responses when answering it.

Another issue emerged as we began to gather responses to the survey questions. I compiled students' responses in a spread sheet as they said them aloud. When we arrived at the question, “Which grade level do you prefer to teach?,” some responded with numerical grade levels such as “1” and “3,” but others responded with “Pre-Kindergarten.” I asked the class how these responses should be coded. Some wanted to be consistent in giving each grade level a number, so they proposed “-1” for Pre-Kindergarten and “0” for Kindergarten. Others objected, since these numerical designations are not commonly used for Pre-Kindergarten and Kindergarten. Ultimately, this discussion drove us back to distinguishing between different types of categorical data. Since the numerals “1,” “3,” etc. can be viewed as ordinal category labels for grade levels, we decided to use the familiar ordinal category labels, “Pre-K” and “K,” to code “Pre-Kindergarten” and “Kindergarten” responses.

Issues that arose while constructing and administering the survey questions also helped to develop the prospective teachers’ horizon knowledge. As children construct surveys, they may pose questions requiring knowledge beyond the scope of the statistics they have studied in the school curriculum. For instance, a question that requires a grade level in response is simple to pose, but children may need assistance understanding how a numeral can represent a number in some cases and a category label in others. The nature of the objects represented by numerals in a data set dictate the types of data displays and summary statistics that are suitable for studying it. Horizon knowledge enables teachers to judge which types of displays and summary statistics are appropriate and lead children in productive directions.

ANALYSING SURVEY DATA

Dynamic statistics software packages allow students to quickly produce, compare, and contrast multiple representations of data. They also allow students to explore how changing a value in a data set affects the accompanying statistics and graphical representations. The representational capabilities of dynamic statistics software are of special interest to teachers because of their educational value. To help familiarize my students with dynamic statistical representations, which resonate with the notion of specialized knowledge, I had them use *TinkerPlots* (Konold & Miller, 2005) and *Fathom* (Key Curriculum Press, 2006) on various occasions. On the day of the class described in this article, they used both programs to look for patterns in the data generated from the class survey.

After looking for patterns in the class survey data, I asked my students to share their findings. The ensuing class discussion helped us delve into the capabilities of dynamic representations. Some had produced displays of the grade levels survey respondents preferred to teach (Figure 1). For these data, *Fathom* provided the option to produce a bar graph, but not a histogram. This led to the observation that the "K" and "Pre-K" codes we used during data collection signalled that the data were categorical. Others used *Fathom* to produce dot plots of the data from the survey question about night time bedroom temperature. When this graphical representation was shared, I used *Fathom* to quickly produce a histogram for the same display, and asked the class how the histogram was related to the dot plot. We also observed how the shape of the histogram could be changed by varying bin widths (Figure 2). Data explorations done with *TinkerPlots* helped bring out the idea that informal displays can be helpful to explore before using more formal conventional ones. The *TinkerPlots* display shown in Figure 3 provided a beginning comparison of credit hours and number of courses taken in a semester that eventually could be organized into a scatterplot by ordering values within each quadrant.

<INSERT FIGURE 1 HERE>

<INSERT FIGURE 2 HERE>

<INSERT FIGURE 3 HERE>

HOMEWORK DUE BEFORE THE NEXT CLASS SESSION

The homework assignment for the next class session was designed to help develop curriculum knowledge about the roles of graphical displays within instructional sequences for children. My students were to read an article describing an extended teaching experiment designed to promote statistical discourse (McClain, 1999). In the teaching experiment, children were encouraged to invent graphical representations for data. Two of the curriculum-related questions to be answered as part of the assignment were:

(1) On p. 374, the author asked, “Do students first need to know how to construct various types of graphs before they can engage in an analysis of data, or can they learn how to construct various types of graphs by engaging in data analysis?” Write a response to the author’s question. Explain how your response compares to the position taken by the author of the article.

(2) On p. 375, the author stated, “My assessment of their (the students’) performance would not be based solely on whether they made a histogram and made it correctly but would focus more on how they reasoned about organizing and representing the data.” Do you agree with this decision? Why or why not?

My students’ responses to the items revealed their uneasiness about setting children loose to analyse data before teaching them conventional graphical representations. Those who had experiences in classrooms embracing more traditional teaching methods had trouble envisioning an alternative curricular structure. Hence, responses to the homework items helped me identify another worthwhile objective for the course: familiarizing prospective teachers with children’s potential learning trajectories within innovative curriculum sequences.

CONCLUSION

Several tools currently exist to help teachers develop subject matter knowledge and pedagogical content knowledge in tandem. In the class described above, I used several of them, including a case study of children's learning, dynamic statistics software, and professional journal articles written for teachers. Although these tools were employed in the context of an undergraduate course for prospective teachers, they are appropriate for practicing teachers as well. I hope the experiences I have described will catalyse the thinking of teachers and teacher educators about how the LMT framework can inform the selection and use of tools for enhancing statistical knowledge for teaching.

References

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education* (GAISE) *report*. Alexandria, VA: American Statistical Association.

Groth, R.E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, *38*, 427-437.

Hill, H.C., Ball, D.L., & Schilling, S.G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*, 372-200.

Key Curriculum Press. (2006). *Fathom dynamic data software* (Version 2.03) [Computer software]. Emeryville, CA: Key Curriculum Press.

Konold, C., & Miller, C.D. (2005). *TinkerPlots*: *Dynamic data exploration*. (Version 1.0) [Computer software]. Emeryville, CA: Key Curriculum Press.

Leavy, A.M., Friel, S.N., & Mamer, J.D. (2009). It's a fird! Can you compute a median of categorical data? *Mathematics Teaching in the Middle School*, *14*, 344-351.

McClain, K. (1999). Reflecting on students' understanding of data. *Mathematics Teaching in the Middle School, 4*, 374-380.

Russell, S.J., Schifter, D., & Bastable, V. (2002). *Developing mathematical ideas*: *Working with data*. Parsippany, NJ: Dale Seymour Publications.

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

Watson, J.M., Fitzallen, N.E., Wilson, K.G., & Creed, J.F. (2008). The representational value of hats. *Mathematics Teaching in the Middle School, 14*, 4-10.

Table 1. Relationship between the class activities for the day and a statistical knowledge for teaching (SKT) framework

SKT Component	**Corresponding Classroom Activities**
Common Content Knowledge	• Considering the feasibility of computing the median of frequencies for nominal categorical data as part of *Developing Mathematical Ideas* case discussion and homework article responses. • Distinguishing between categorical and quantitative data in constructing survey questions.
Specialized Content Knowledge	• Using dynamic statistics software packages to explore data generated from a class survey and to become familiar with the features and capabilities of dynamic statistical representations.
Horizon Knowledge	• Deciding how to code and analyse data generated by student-constructed survey questions.
Knowledge of Content and Students	• Reading article describing student misconceptions about the median and categorical data. • Studying *Developing Mathematical Ideas* case with a focus on student strategies for organizing nominal categorical data.
Knowledge of Content and Teaching	• Analysing *Developing Mathematical Ideas* case in which a teacher used a class survey as a pedagogical device. • Constructing a class survey like the one in the case.
Curriculum Knowledge	• Forming an opinion on the optimal time to introduce conventional statistical representations in a curricular sequence by responding to a homework article about an inquiry-oriented curriculum.

Figure 1. Bar graph representation of grade levels the prospective teachers preferred to teach

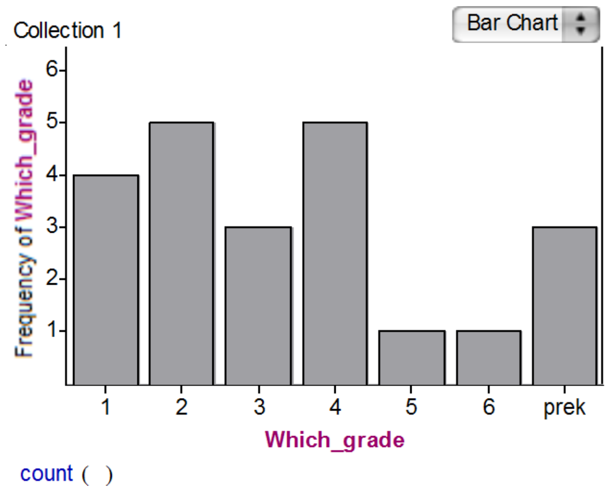

Figure 2. Data set on bedroom temperature represented as a dot plot and as histograms with various bin widths

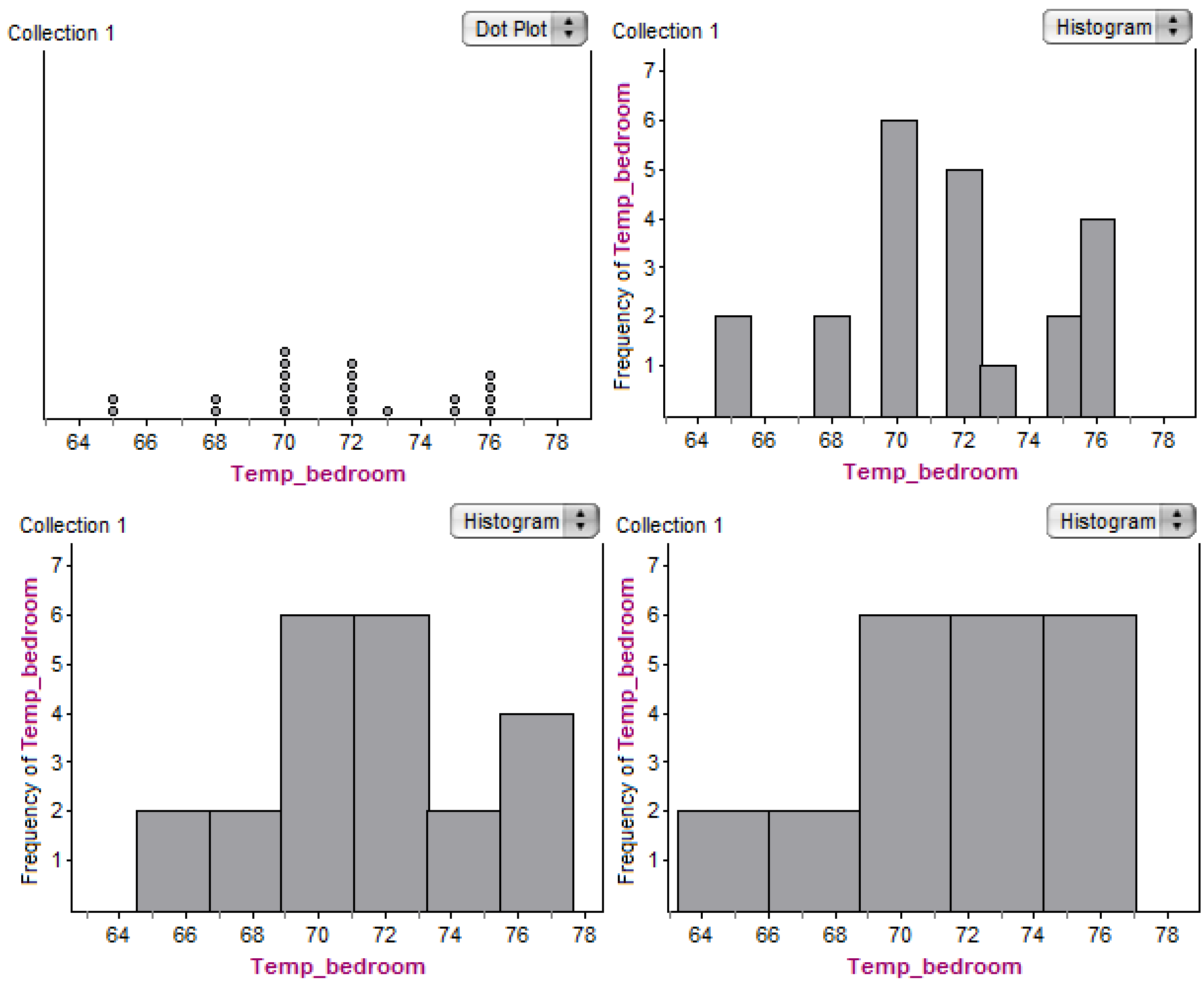

Figure 3. *TinkerPlots* graph comparing the number of credit hours to the number of classes taken