PROFESSIONAL DEVELOPMENT TO TRANSFORM MIDDLE AND HIGH SCHOOL TEACHERS' UNDERSTANDINGS ABOUT DISTRIBUTION

Susan A. Peters
Department of Middle and Secondary Education
University of Louisville, Louisville, KY 40292
s.peters@louisville.edu

The concept of distribution is foundational to statistical reasoning. To facilitate students' development of statistical reasoning abilities, students, and thus their teachers, need deep understandings of distribution. Using data from a professional development program designed to deepen teachers' statistical understandings, we investigate how dilemma, critical reflection, and rational discourse affect middle and high school teachers' understanding of and reasoning about distribution. Framed by transformative learning theory, results reveal that critically reflecting on activities such as growing samples and engaging in rational discourse to consider multiple perspectives while working on these activities broadened teachers' perspectives about distribution. This study identifies factors associated with deepening teachers' statistical understandings and supporting their reasoning about distribution in increasingly mature and refined ways.

INTRODUCTION

Headlines such as "Statistician Projected as Top 10 Fastest-Growing Job" (thisisstatistics, 2016) underscore the importance of statistical study and development of statistical understandings and reasoning for today's students. Studies connecting student achievement with teacher knowledge (e.g., Mohr-Schroeder et al., 2017) suggest that students need knowledgeable teachers with robust understandings to develop appropriate understandings. Yet, anecdotal and empirical evidence suggests that many teachers may not have sufficient statistical experiences to develop the depth of understandings necessary to facilitate students' deep learning of fundamental content (e.g., Shaughnessy, 2007). As one step towards improving teachers' understandings, we designed, implemented, and studied a professional development (PD) program with middle and high school mathematics teachers focused on developing their reasoning about key statistical concepts and their understandings of those concepts. In this paper, we focus on teachers' work with distribution—a concept fundamental to statistics and statistical reasoning (e.g., Wild, 2006)—to address the research question: how do dilemma, critical reflection, and rational discourse affect middle and high school teachers' understanding of and reasoning about distribution?

BACKGROUND

Research suggests that students (and teachers) struggle to reason about data as an aggregate collection of values and thus struggle to reason about distribution—"the pattern of variation in a variable" (Wild, 2006, p. 11). Research also suggests instructional strategies that offer promise for helping students overcome their struggles with distribution and develop understandings of distribution that are foundational for understanding sampling variability and sampling distribution and hence for drawing informal inferences from data.

Developing an aggregate view of data—being able to view data in terms of the whole distribution—is important for reasoning about data and for understanding distribution (e.g., Hancock, Kaput, & Goldsmith, 1992). Views of data as an aggregate focus on patterns of variability, which include notions of shape, center, and spread, whereas pointwise views of data allow for calculation of summary measures such as mean, median, range, interquartile range, and standard deviation and consideration of individual deviations from patterns (Bakker & Gravemeijer, 2004). Statistical "experts" are capable of moving flexibly between pointwise and aggregate views of data, and this dual perspective aligns with views that understandings of distribution can be enhanced by viewing data as a distribution, or noise, around a signal, wherein the notion of central tendency embodies the idea of signal and variation embodies the idea of noise (Konold & Pollatsek, 2002).

Various instructional interventions offer promise for supporting individuals in deepening their understandings of distribution, including strategies to compare distributions (e.g., Leavy,

2006) and to represent and analyze multiple representations of data, often with the use of technology (e.g., Cobb, 1999). Arguably, the most implemented and researched instructional heuristic for distribution is *growing samples*, which was studied by Bakker (2004), Ben-Zvi (2006), and colleagues (e.g., Ben-Zvi, Bakker, & Makar, 2015). Growing samples involves cycles of first hypothesizing a sample drawn from an unknown population, examining actual samples from the population, and making inferences about larger samples as the sample size increases towards drawing inferences about a population. Data suggest that growing samples activities support students in changing their focus from individual datum to the aggregate of data (e.g., Bakker, 2004) and in isolating signals from the noise of data (e.g., Ben-Zvi, 2006) to reason from samples (e.g., Ben-Zvi et al., 2015) and make informal inferences about populations (Ben-Zvi, Aridor, Makar, & Bakker, 2012). Found to be effective for advancing students' understandings of distribution, these instructional strategies offer promise for deepening inservice teachers' understandings, particularly in light of evidence that suggests teachers' struggles with statistics are similar to those of their students (e.g., Batanero, Burrill, & Reading, 2011).

FRAMEWORK

Because pre-K–12 instruction focuses on developing new understandings or enhancing current understandings, instruction tends to focus on answering questions of what and how. Adult learning often results from focusing on the premises behind content and processes, suggesting that adult education should focus on answering questions of why. Transformative learning theory (Mezirow, 1991) acknowledges the unique nature of adult learning and takes into account learning from reflecting on premises. An overarching tenet of the theory is that powerful learning results from transforming meaning perspectives—the broad predispositions formed from experiences.

Perspective transformation can begin with a "disorienting dilemma" that triggers the examination of broad presuppositions or with a series of incremental dilemmas that prompt examination of particular knowledge or attitudes (Taylor, 2000). Dilemmas that induce questioning of assumptions for meaning schemes, which consist of specific expectations and knowledge used to interpret experiences implicitly and that are the building blocks for meaning perspectives, can be resolved by creating, enhancing, or transforming meaning schemes (Mezirow, 1991). Crucial for transforming meaning schemes or perspectives is critical reflection—reflecting on premises to question the importance, validity, or utility of knowledge. Critical reflection often is supported by rational discourse to examine alternative perspectives and to assess expectations and knowledge towards developing and acting on plans to resolve dilemmas.

We designed our PD program to facilitate adult learning and to build on the results of research related to students' reasoning and learning about distribution. The activities we developed include planned triggers for dilemmas in typical areas of struggle. To resolve dilemmas and enhance or transform skills and knowledge related to distribution, we provided considerable opportunities for teachers to explore content conceptually, to engage in rational discourse with each other and with PD facilitators, and to examine underlying premises to reflect critically on content. The program also incorporated characteristics of "high quality" PD such as sustained duration and focus on content (e.g., Goos et al., 2007). The program included a 40-hour intensive experience in which teachers actively explored statistics content as learners using the problem-solving process of formulating a question, collecting data, analyzing data, and interpreting results.

Prior to engaging with a growing samples activity, teachers explored formal measures of center and informal notions of distribution and variation. They explored mean as a balance point and took repeated blood pressure measures to isolate signal from noise. Two of the activities utilized a context of children's ages in U. S. households and included characteristics of growing samples activities. In particular, given a randomly selected sample of 15 children's ages from the American Time Use Survey (ATUS, 2011), teachers considered the composition of a hypothetical sample of size 30, compared their hypotheses with five actual random samples of 30 ages, considered the sample of size 150 formed by combining the five samples, considered the composition of the population, and compared their hypotheses and the samples with the population. The population is fairly uniform but contains lower frequencies for ages 18 and above.

Teachers then explored distribution further by engaging with an extended growing samples activity for which the population was 14,285 U. S. grade 9-12 students who completed the National

Youth Risk Behavior Survey (NYRBS; Centers for Disease Control and Prevention, 2011). After considering survey design features and how to investigate childhood obesity, the teachers hypothesized weights for a sample of 10 grade 9-12 students and compared their hypotheses with the hypotheses from other groups before comparing and contrasting their hypotheses to the distribution of weights from three actual samples of size 10. They then formulated hypotheses about the population of weights before engaging in two additional iterations of hypotheses and comparisons for samples of size 30 and samples of size 90. They concluded the growing samples activity by comparing the actual samples with the actual population of weights.

DATA SOURCES AND METHODS

Eleven middle- and nine high-school teachers participated in 40 hours of summer PD. They varied in their statistical learning and teaching experiences. All but one teacher completed a minimum of one introductory-level statistics course as part of a teacher preparation program, and one teacher completed one or more advanced or mathematical statistics courses. During a typical school year, two teachers taught as few as five statistics-related lessons and an additional seven teachers taught as few as ten statistics-related lessons. Seven teachers taught more than 25 statistic-related lessons, with some teaching the equivalent of an introductory college-level statistics course.

Data sources included audio- and video-recordings of large-group discussions and small-group activities from PD sessions and teachers' written work and reflections. For each recording, we created a content log and developed transcripts. We began analyses by considering aspects of the PD program designed to encourage dilemma, critical reflection, and rational discourse and identified relevant transcript passages. We searched teachers' written work and transcripts for evidence of dilemma, critical reflection, and rational discourse, paying particular attention to indications of insights, questions, or confusion; thoughts and reasoning beyond the immediately observable; content-related interactions with other teachers or consideration of multiple perspectives; and references to the preceding as potential evidence for dilemma, critical reflection, and rational discourse, respectively. Transcripts and teachers' work and reflections were coded using a combination of codes developed from the theoretical framework (e.g., dilemma, self-examination, assessment of assumptions, rational discourse, etc.) and codes that emerged from the data. We made comparisons across teachers to identify common themes.

RESULTS

In many respects, teachers' reasoning about distribution while growing samples followed the same general patterns observed with students. They tended to reason about individual values for small samples, such as each representing the size-10 sample distribution with dotplots, and reason about the aggregate of data for larger samples, such as each representing the size-90 sample distribution with a smooth curve. Unlike students, however, the teachers sometimes struggled when they reasoned from their prior experiences or when they considered teaching the statistics content.

Dilemma

Many teachers experienced multiple dilemmas throughout the PD program and when they explored distribution. With respect to hypothesizing and describing distributions, several middle school teachers struggled to describe the distributions' shapes. They seemingly wanted precise definitions for shapes and were bothered by questions for which "there's no [single] right answer. I hate that." Yvette, for example, told her group, "I'm really having issues with this" when they classified the distribution displayed in Figure 1 as skewed right. She noted, "So even though 90 percent of your data is right here [around 150] but because you have these two [weights near 260 and 300], that makes it skewed? ... Skewed means there has to be a lot more stuff in the tail." After hearing that some people might prefer to describe the distribution as somewhat symmetric with two outliers on the high end, whereas others might describe the distribution as skewed right because of the outliers, lack of multiple occurrences for weights on the high end, and context, Yvette's groupmate, Brooke, became perturbed. She wanted a definition for symmetric and a definition of skewed right so that she would know what her students needed to know about shape so her students could correctly answer standardized assessment questions about a distribution's shape.
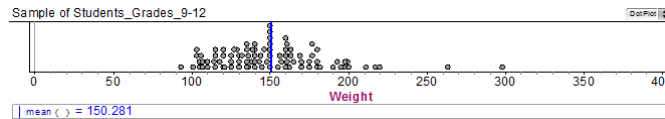
Figure 1. Weights from a random sample of 90 children who completed the 2011 NYRBS.

Several high school teachers experienced dilemmas when their observations conflicted with expectation and prior experiences. Ruthie, for example, expected the distribution of ages for children in U. S. households to be mound-shaped and approximately normal. Even after examining the five samples displayed in Figure 2, Ruthie stated, "I still think it is [bell-shaped]." Her partner, Valenta, questioned this premise, noting "I don't see how you can have both of these situations [Figure 2b and 2c]" with an approximately normal population. Ruthie countered, "I feel there's going to be more mounding in the middle," and another group member, Tim, agreed: "I feel that way, too." Tim, however, began to experience dilemma when he attempted to provide rationale for that shape: " [PD facilitator's] going to ask us why, so I was trying to think why." Ruthie and Tim also had expectations for the population distribution of child weights to be bimodal ("definitely two modes") to indicate differences in weights between males and females based on investigations they had done with adult heights. They only began to question their hypothesis after examining graphs for actual samples of size 10 from the population. Also related to expectations based on shape, several teachers had clear expectations for the relationship between mean and the median in skewed distributions. Nathan, for example, experienced some dilemma when he observed the value of the mean was smaller than the median for the distribution displayed in Figure 2c, noting "it's less than the median where I think it should be more" in a distribution he classified as skewed right.
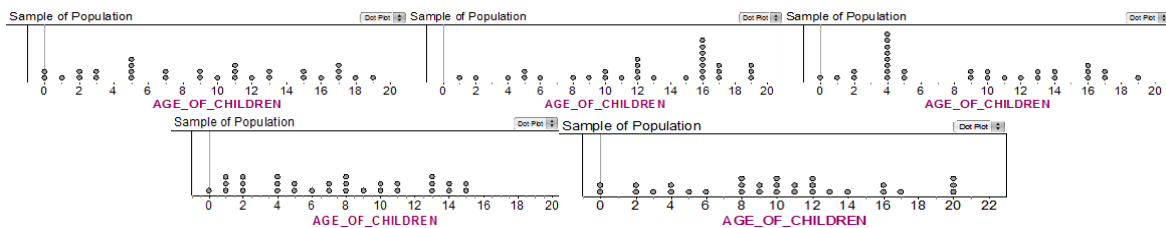

Figure 2 (a-e). Random samples of 30 children's ages from ATUS.

A last type of dilemma related to distribution arose early in the PD program and after teachers first considered appropriate sample sizes. Questions about sample size continued to surface with each subsequent design discussion. Some high school teachers proffered a sample of size 30 and above as sufficient for drawing inferences about a population based on the central limit theorem, but many of the remaining teachers struggled to see how such a small sample might be used to draw valid conclusions about large populations.

Rational Discourse

As teachers sought to resolve their dilemmas, they engaged in discourse with each other and with the PD facilitators and reasoned about data from multiple samples within context to examine alternative perspectives related to distribution. When examining the sample displayed in Figure 2b, for example, Ruthie experienced dilemma after she classified the distribution as skewed left with a median of 12 and mean greater than 12. She indicated, "This goes against everything that I've taught…The mean should be less than 12, right? If it's skewed left, then the mean has to be less than the median…but it's not." Tim and Valenta attempted to help Ruthie resolve her dilemma by presenting alternative perspectives. Tim indicated, "the skew's just probably not enough" and further noted that for a left-skewed distribution, he would expect "more values in here [upper end of the distribution]," which would increase the value of the median more than the mean. Valenta also pointed out the "big stack" of values at 12 that brought the left skew of the distribution into question. Ruthie seemingly accepted their reasoning as she agreed with Tim's and Valenta's points and stopped questioning the values of the mean and median for this distribution. As was true with teachers' discussions about shape, teachers' engagement in rational discourse enabled them to accept reasoning from the data over deterministic rules.

Somewhat unique to growing samples, the teachers also tended to interrogate data from actual samples within the given context to reconcile different perspectives of the population reflected in samples. For example, after Ruthie and Tim hypothesized that the population of weights for students in grades 9-12 would be bimodal, they observed that other teachers hypothesized that the population would either be symmetric or skewed right. Upon receiving the actual samples of size 10 from the population displayed in Figure 3, however, they began to question whether "we're still right about two clusters." They rationalized that even if the distribution of male weights and the distribution of female weights were mound shaped with peaks at different weights, the overlap between the two distributions would obscure the peaks and prevent the population from appearing to be bimodal, particularly in light of the approximate four-year difference in ages among students. Tim indicated: "I do think it's going to be mound shaped, just one mound, the more I think about it." To examine each actual sample of size 30, the group considered the shape, range and interval of weights, mean, and median for each sample. They also considered the distribution for the sample of size 90 that would have resulted from combining the three samples of size 30. Ultimately, they concluded that although one sample of size 30 was slightly bimodal in shape, when taken into consideration with other samples, the population was likely to be unimodal. They made no further changes to their hypothesis for the population after examining the samples of size 90. Even though the teachers initially believed strongly that the population would be bimodal, extensive examination and comparison of multiple samples within the context of student weights resulted in the teachers changing their opinions, albeit reluctantly.
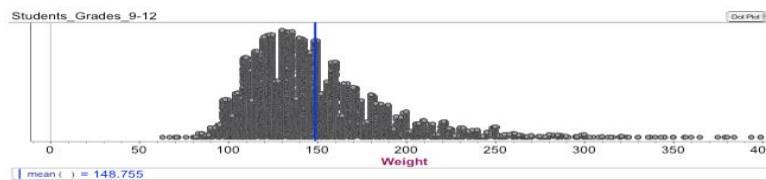

Figure 3. Weights of the 14,285 U. S. students in grades 9-12 who completed the 2011 NYRBS.

Critical Reflection

As indicated by Tim's observation that "[PD facilitator's] going to ask us why," activity-embedded and facilitator questions consistently focused teachers on examining the premises underlying statistical content and processes. As the examples presented above suggest, teachers often engaged in rational discourse to examine alternative perspectives and to reason about and justify their conclusions. For many of their explorations with distribution, the teachers resolved their dilemmas by reflecting on their rational discourse and interrogations of data. One exception was their consideration of sample size for drawing inferences about population distribution characteristics using a sample distribution. As teachers participated in the growing samples activity for student weights, they observed that their hypotheses for the population in terms of shape, peak locations, and ranges changed very little subsequent to examining actual samples of size 30. When asked why facilitators might have asked them to engage in a growing samples activity, Gabby noted, "We talked about sample size of 30…and was 30 actually an accurate sample size for a larger group?" She suggested that the activity confirmed that a sample of size 30 can provide useful information for a population as large as 14,285. Continued emphasis on design and teachers' observations and careful examination of samples from multiple populations allowed the teachers to resolve their dilemmas related to sample size for them to become more comfortable with drawing inferences about characteristics of large population distributions from relatively small samples.

CONCLUSION

Presumably due to teachers' familiarity with distribution, large transformations to teachers' meaning schemes for distribution are not evident. The evidence does, however, suggest enhanced meaning schemes based in teachers' reasoning about distributions using context and data in place of deterministic definitions and rules. Many dilemmas seemingly arose from working with real data whose distributions did not conform to conventional shapes and rules and with messy samples selected randomly from a population. These data, the PD emphasis on premises and conceptual understanding, and teachers' natural desires to understand the content for the benefit of their

students forced teachers to confront misconceptions they never realized they had. Although the struggles of middle and high school teachers sometimes differed, their working together to generate multiple perspectives focused on premises served to help them all resolve their dilemmas related to those struggles. This study's results suggest that PD activities designed with planned triggers for dilemmas, opportunities to engage in rational discourse, and consistent focus on premises can contribute to advancing teachers' understandings and development of content knowledge.

REFERENCES
Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, *3*(2), 64-83.

Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147-168). Dordrecht, The Netherlands: Kluwer.

Batanero, C., Burrill, G., & Reading, C. eds. (2011). *Teaching statistics in school mathematics: Challenges for teaching and teacher education.* Dordrecht, The Netherlands: Springer.

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. Rossman & B. Chance (eds.), *Proceedings of the seventh International Conference on Teaching Statistics*. Voorburg, The Netherlands: International Statistical Institute.

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM, 44*, 913-925.

Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics, 88,* 291-303.

Centers for Disease Control and Prevention. (2011). Youth Risk Behavior Survey data. Retrieved from www.cdc.gov/yrbs

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning, 1*, 5-43.

Goos, M., Dole, S., & Makar, K. (2007). Designing professional development to support teachers' learning in complex environments. *Mathematics Teacher Education and Development*, *8*, 23-47.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic enquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*, 337-364.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*, 259-289.

Leavy, A. (2006). Using data comparison to support a focus on distribution: Examining preservice teachers' understanding of distribution when engaged in statistical inquiry. *Statistics Education Research Journal, 5*(2), 89-114.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.

Mohr-Schroeder, M., Ronau, R. N., Peters, S. A., Lee, C. W., & Bush, W. S. (2017). Predicting student achievement using measures of teachers' knowledge for teaching geometry. *Journal for Research in Mathematics Education, 48*(5), 520-566.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester, Jr. (ed.), *Handbook of research on mathematics teaching and learning* (2nd ed., pp. 957-1009). Greenwich, CT: Information Age.

Taylor, E. W. (2000). Analyzing research on transformative learning theory. In J. Mezirow & Associates (eds.), *Learning as transformation: Critical perspectives on a theory in progress* (pp. 285-328). San Francisco, CA: Jossey-Bass.

Thisisstatistics.org. (2016, January 21). Statistician predicted as top 10 fastest-growing job. Retrieved from http://thisisstatistics.org/statistician-projected-as-top-10-fastest-growing-job/

United States. (2011). *American Time Use Survey*. Washington, DC: U.S. Bureau of Labor Statistics.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal, 5*(2), 10-26.