BEYOND CALCULATIONS: FOSTERING CONCEPTUAL UNDERSTANDING IN STATISTICS GRADUATE TEACHING ASSISTANTS

Jennifer L. Green[1] and Erin E. Blankenship[2]
[1]Montana State University, Bozeman MT, USA
[2]University of Nebraska-Lincoln, Lincoln NE, USA
jgreen@montana.edu

Incoming statistics graduate students who are expected to teach introductory statistics as part of their assistantships are very prepared to teach the procedural portions of the course. However, they are less comfortable with the statistical thinking/reasoning aspects of the course. To better prepare novice teaching assistants (TAs), we developed a course that focuses on how TAs can foster critical thinking and enhance learning in their classrooms. In this course, a team of faculty and exemplary TAs employs a variety of activities to promote conceptual understanding of key statistical ideas. The TAs leave the course with not only a cadre of examples, assessment questions and tips for how to teach introductory statistics, but also an enhanced understanding of statistical concepts. The course has helped to improve the level of teaching in the introductory course, not only by giving the TAs a set of classroom tools, but also by improving their confidence.

BACKGROUND

Like many large research universities, the introductory non-calculus based statistics course (STAT 218) at the University of Nebraska-Lincoln (UNL) is taught primarily by graduate students. The introductory course at UNL serves about 1200 students per year, and plays an important role in the university's general education program. Unlike many other graduate programs in statistics, UNL's graduate teaching assistants (TAs) serve as the primary instructors in the classroom, not as recitation leaders. They are not given a common syllabus, common homework or common exams. We allow them classroom autonomy because we believe that such independence has educational value and provides them with experiences that will be valued by future employers.

Because the TAs play an important role in undergraduate statistics education, their development as statistics educators is crucial (Moore, 2005). Most students admitted into the graduate program in statistics at UNL have an undergraduate major in mathematics and therefore may have had fewer opportunities to develop a conceptual understanding of elementary statistics than have statistics majors. While our incoming graduate students are very prepared to teach the nuts-and-bolts calculation portions of the introductory course, they are less comfortable with the statistical thinking/reasoning aspects of the course (Green, 2010). We are more aware of this dichotomy because statistical thinking/reasoning is now a major focus of the introductory course as a result of the reform movement precipitated by the publication of the Guidelines for Assessment and Instruction in Statistics Education (GAISE; Garfield et al., 2005). To help better prepare the TAs, we decided to develop a new course on how the TAs can use writing to foster critical thinking and enhance learning in their classrooms. The authors, with funding from an internal UNL grant, developed a TA preparation course along with Walt Stroup and Alison Friedow (a PhD candidate in English), and we taught the course as a team in the fall, 2008 semester. The course has been taught every fall semester since 2008 and is required for all new TAs. The TA preparation course is taught by an instructional team comprised of at least one faculty member and at least one PhD-candidate TA who has classroom experience and has excelled as an instructor. Course instructors of the one-credit hour course strive to not only develop the new TAs as instructors, but also to create a community of teachers and learners within the department of statistics.

As part of the course, the TAs construct classroom activities, lesson plans, assessment items and grading rubrics. We frequently use peer review to give the TAs immediate feedback, and considerable course time is devoted to discussion and modeling classroom scenarios. We visit each TA's classroom to observe his or her teaching, and the TAs also visit other STAT 218 classrooms to observe other instructors. Each TA keeps a teaching journal and submits a STAT 218 course portfolio at the end of the semester.

The TA course employs a schedule that allows us to teach the STAT 218 course material a few weeks before the same concepts are taught in the introductory course. By the time concepts

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

appear in the STAT 218 curriculum, the TAs have a cadre of examples and assessment questions as well as tips for how to teach the material and which writing-to-learn activities would be best-suited for the concepts. The TA preparation course does not address all of the STAT 218 content, but instead focuses on the areas with which students most often struggle. The major statistical concepts we discuss are variability, sampling distributions and statistical inference (interval estimation and hypothesis tests).

THE PROBLEM

The TA preparation course was developed to focus on low-stakes, writing-to-learn exercises, and how they can be used as a means of teaching statistical literacy (Friedow et al., 2012). Over the years, it has evolved into a course more generally focusing on pedagogical strategies useful for STAT 218. Because the TA course emphasizes GAISE (Garfield et al., 2005), we focus on statistical thinking and how to promote conceptual understanding of the STAT 218 content. However, to effectively develop student understanding, the TA instructors themselves must have good conceptual understanding of the material (Noll, 2011).

Like Woodard and McGowan (2012), we believe that the concept of sampling distribution is the most difficult idea in the introductory curriculum, both for the students and for the TAs. Similar to Noll (2011), we found that the TAs do not have a uniform, working understanding of sampling distributions. What knowledge they do possess is uneven, and seems to exist in isolation from related statistical concepts such as statistical inference. For example, in a class meeting at the beginning of the unit, the TAs are asked to define the term sampling distribution. It is not unusual to receive a definition that asserts that sampling distributions are used to estimate the population distribution. Similarly, TAs often believe that the effects of sample size can be explored with their STAT 218 students by adding additional points to the empirical sampling distribution built up through simulation. In particular, like Noll (2011), we find that the TAs have trouble distinguishing between theoretical and empirical distributions.

We also find that the TAs often equate sampling distributions with the normal distribution, and struggle when they are presented with a sampling distribution for something other than a mean or proportion (e.g., a range). Many of the TAs' definitions of sampling distribution refer only to the mean and reference the central limit theorem.

Due to the uneven knowledge of sampling distributions among TAs, we must ensure that they themselves possess a thorough working knowledge of the concept, in addition to pedagogical strategies for conveying that information to their students. To do this, we developed a unit in the TA preparation course that simultaneously exposes gaps in TAs' understanding and models effective discovery activities that can be used in the STAT 218 classroom. The remainder of this paper describes the activities we employ during the sampling distribution unit and some of our findings about TAs' overall perceptions of the course and its impact on their own learning and confidence to teach statistics.

SAMPLING DISTRIBUTION UNIT

The sampling distribution unit typically spans 4-5 classes of the 14 class meetings in a semester and encompasses multiple activities that allow TAs to reflect on their own understanding of sampling distributions, as well as on how they will teach sampling distribution concepts in STAT 218. We describe these activities below to help illustrate the progression of the unit and to provide a sense of what our course entails. The variety of methods we use to teach sampling distributions are representative of the methods we use to teach other content not discussed in this paper.

Activity 1: Sampling Distribution Definition Word Scramble

For this activity, each TA gets a blank strip of paper and is instructed to write a key word or phrase to describe a sampling distribution. For example, one cohort of TAs provided the following words: *sample, shape, general dispersement* [sic], *probability, behavior of a particular sample* and *difference between a value and an average value*. Then we collect these papers and project them on a document camera for the class to see. We also include additional strips of paper with other key phrases (*probability distribution, behavior, statistic, repeated, random samples,*

sample size, size n and *parent population*) and potential distractors (*random variable [RV], observed, distribution, parameter, mean, variance* and *normal*) that haven't already been provided by the TAs. After the words are projected, the TAs work in groups to put together as many of the listed words as they need to create a coherent definition of sampling distribution. Each group then writes its definition on the board and briefly describes the definition's major ideas to the rest of the class. Finally, as a class, we discuss and list the similarities and differences between their multiple definitions.

During this initial activity, the TAs are able to air their conceptions in a safe environment, and we are able to informally assess the robustness or limits of their understanding. In the past, groups of TAs have formed vastly different definitions that convey their misconceptions. For example, groups of TAs within the aforementioned cohort created the following three definitions:

- *The probability distribution of a population parameter of a RV is based on the observed behavior of a random sample.*
- *Probability distribution of repeated random samples of size n from the RV that describes the observed behavior of the statistic which is about the parameter(s) of the parent distribution with some mean and variance.*
- *Probability distribution of the sample statistic based on repeated samples of size n from the parent population with mean μ and variance σ^2.*

At this point in the unit, we don't correct any misunderstandings. Instead, we provide additional learning opportunities to help the TAs recognize their own gaps in understanding and build a stronger conceptual understanding of sampling distributions that they can then translate into their teaching.

Activity 2: Sampling Distributions Simulation

At the beginning of this activity, the TAs are instructed to write a response to the following questions: *What are the similarities and differences between the following distributions: parent distribution, data distribution and sampling distribution? Where does probability distribution fit into these?* After they are done writing, the TAs use the computer software, TinkerPlots (Konold & Miller, 2011), to complete a simulation activity that has them explore the sampling distributions of various statistics (sample minimum, sample maximum, median and range) when taking random samples of size n=30 from parent populations with various shapes (unimodal, skewed and multimodal). To maximize class time, each group is assigned one of the parent populations and asked to simulate the sampling distributions of the four statistics identified above. While they are conducting these simulations, the TAs are asked to identify whether each plot is a parent, data or sampling distribution and whether each is empirical or theoretical. After the TAs complete the simulation, they sketch their findings on the board and share their results with the rest of the class. This activity and its subsequent discussions allow us to distinguish between different kinds of distributions the TAs will encounter when teaching STAT 218 and to actively confront the misconception that all statistics are normally distributed. It also lets us demonstrate the power of software and simulation to explore various concepts, and we collectively discuss how the same activity could be used in a STAT 218 course or modified to teach other topics, such as the effect of sample size on sampling distributions. Finally, we have the TAs reflect on their beginning-of-activity writings and describe whether their answers have changed based on the simulation.

Activity 3: Sampling Distribution Definition Rewrite

Prior to class time, each TA writes a definition of sampling distribution that s/he would give to his or her students and submits it to the instructional team. These definitions are submitted early so that we can purposefully form groups based on the definitions. For example, sometimes these TA groups are formed based on similar definitions. However, if there is a TA with an incorrect definition, that TA is grouped with those who demonstrate a clearer understanding of sampling distributions. The TAs also bring these definitions to class and share them with other TAs in their group. Each group creates one definition that improves upon their individual definitions and

writes it on the board. For example, the same cohort of TAs described above created the following definitions within their assigned groups:

- *The probability distribution of a statistic based on all possible random samples of size n that can be drawn from a given population is called the sampling distribution. The computed statistic can be the mean, median, proportion, standard deviation, etc.*
- *A sampling distribution is the probability distribution of different values of the statistic, which is based off a random sample from a parent population.*
- *A sampling distribution of a statistic (descriptive measurements of a sample) is the probability distribution that specifies probabilities for all possible values of the statistic which result from repeated samples of size n.*

After the definitions are written on the board, the entire class uses these definitions to come up with one final overall definition. In this case, the class created the following definition of sampling distribution: *The probability distribution of a statistic based on repeated random samples of size n from a given population*. This particular activity is an example of 'writing to refocus a discussion' (Bean, 2001), and it lets the TAs use their own understandings of sampling distributions to collaboratively and progressively work towards one collective definition that they all helped to create. It also gives us the opportunity to discuss any misconceptions that might remain and demonstrates a similar process they can use with their own students to foster conceptual understanding.

Activity 4: Sampling Distribution Benchmarks

After the TAs complete the previous activities to support their own understanding of sampling distributions, we ask them to reflect on these concepts as instructors. Specifically, we ask the TAs to individually identify and write key learning objectives or benchmarks they expect their students to meet when learning about sampling distributions. After they have finished writing their thoughts, the class discusses the various concepts they hope their students will understand about sampling distributions, and we list those concepts on the board. For example, one semester a cohort of TAs came up with the following list:

- *Meaning of "statistic"*
- *Expecting variability in a calculated statistic*
- *Why do we need sampling distributions?*
  - *making inferences; likelihood of an observed statistic*
  - *because $\bar{x}$ is not equal to μ*
- *The relationship between sample size and variability of the sampling distribution*
- *The shape of the sampling distribution will not (necessarily) be the same as the parent population*
- *Empirical sampling distributions get "closer" to the theoretical sampling distribution with more samples*

This activity gives the TAs an opportunity to practice identifying major concepts and articulating them in the form of learning objectives and benchmarks. It also models how to use the purpose of a course and its goals to discriminate between "big ideas" and "minutiae" so that the focus of STAT 218 remains on developing statistical thinking, literacy and reasoning skills. In addition, the TAs have an example of how to use 'writing to probe a subject' (Bean, 2001), so that they can use a similar technique to encourage student interaction and discussion in their own courses.

Activity 5: Sampling Distribution Activity Peer Review

Before class time, each TA creates an activity to explore sampling distributions with his or her students and identifies which benchmarks or learning objectives the activity addresses. Each TA also writes an assessment to gauge student understanding and drafts an author's note, where

s/he writes down any questions or concerns s/he has about the materials s/he created. The TAs then bring their activities, assessments and authors' notes to class. During class, pairs of TAs exchange materials to review and provide feedback that addresses the comments on the author's note and that addresses the following instructor-provided questions:

1) *Would you be excited about carrying out the activity in your own class? If yes, why? What do you like about the activity? Is there anything you would want to change to make the activity work better in your classroom? If you would not be excited (or, in fact, would be unwilling) to carry out the activity in your own class, why? What about the activity would keep you from using it with your students? Are there any changes you would suggest that would make the activity more useful?*
2) *Is there enough detail provided in the activity that you could use it in your classroom? If not, what additional information would you need?*
3) *Does the assessment plan match the learning objectives? If so, in what way? If not, why not?*
4) *What knowledge will the assessment provide about student learning?*

After partners exchange their feedback and comments, each TA shares highlights of his or her partner's activity and the learning objective(s)/benchmark(s) it addresses with the rest of the class. By using peer review, we try to create a teaching community amongst the TAs that encourages them to reflect upon their own learning and teaching practices, and to support the teaching and learning of others.

TA PERCEPTIONS OF COURSE

We explore the TAs' development as instructors by asking them to reflect on their overall perceptions of the course and its impact on their own learning and confidence to teach introductory statistics. Each semester, we assign a "letter of learning" in which each TA reflects on and writes about his or her experiences as a student/learner in the course. Specifically, we have them write a narrative that makes a case for their learning and how their work in the course has supported the learning that is most important to them. We also have them explain how the intellectual community of the class has factored into their learning as both a student and a teacher of statistics.

We collected and read the TAs' letters of learning from the first five iterations of the TA preparation course (n=34 letters). In their letters, the TAs communicate that they developed a deeper conceptual understanding of statistics concepts, as well as a better sense of who they are as instructors as a result of the TA preparation course, as indicated by one TA in the reflection: "Several methods of engagement were [used]…all with dual goals. The first goal was to help us master the material at a more fundamental level and the second was to give us a tool to leverage in our own classes."

A few TAs specifically refer to their improved understanding of the concept of sampling distribution and how this contributed to their development as teachers. For example, one TA writes, "At the time sampling distributions were being discussed in this class…I needed help seeing the big picture idea of sampling distributions. Placing everyone's ideas on the whiteboard allowed me to see other interpretations of sampling distributions.

Another TA reflects, "[This course] has challenged me to really…make sure that I thoroughly understand the material I will be teaching my students. The section over sampling distributions was especially helpful as this is an area I haven't really taught in the past. By gaining a better understanding this semester, it will help me to be more effective an instructor in the future."

In general, a majority of TAs reflect on how the course not only gave them a set of classroom tools, but also improved their confidence to teach introductory statistics. One TA writes about this improved confidence from the perspective of teaching difficult concepts: "Overall,…the course has been quite helpful to me in preparing for the next semester. I still am a bit scared, but I do feel that there is a solid community for me to draw upon, and I feel considerably more prepared than I did at the start of the school year…The fear is waning as well—though still present. Confidence in explaining the difficult topics is likely what I gained most from this course, as well as an appreciation for active learning."

Yet, another TA discusses improved confidence beyond just knowing how to teach challenging topics, using a broader perspective of effectively communicating the overarching course goals: "The activities that we did throughout the semester really helped me to focus on big picture goals that I have for my class. I learned how to teach certain subjects that I was worried about, but more than that, I learned what I want my students to get out of the class. Also, I learned that all of us are in a similar situation. I gained comfort with my fellow TAs and know that if I am lost or need help, they can help me. I certainly gained more confidence in myself as an instructor and now know that I can be an effective teacher."

Overall, the TAs state appreciation for the opportunity to grow collectively as instructors, discussing how the TA preparation course fosters an intellectual community of teachers and learners of statistics for them within the UNL Department of Statistics.

CONCLUSION

The course has evolved since it was first offered in fall, 2008, as has the demography of the TAs enrolled in the course. For example, not all TAs are currently teaching STAT 218, and some are instead teaching other statistics courses. There is also more diversity in the TAs' background with respect to the stage of their graduate education. Nonetheless, we are still finding gaps in TAs' understanding of variability, sampling distributions and how these concepts connect to statistical inference. Through this course and, more specifically, the sampling distribution unit, we are able to explore fundamental concepts that are essential for their success not only as teachers, but also as statisticians.

The types of activities we describe and use in the course are not meant to be course-specific. We intentionally chose to describe the activities in the sampling distribution unit, because undergraduate (and graduate) students struggle to understand them conceptually (Noll, 2011; Woodard & McGowan, 2012). By describing the activities in depth, we hope that others can adapt them to best fit their context and needs. For example, we think these activities can be adapted and used in undergraduate and graduate-level statistics content courses, not just in a course like the one described. We also think this type of course would be beneficial to a variety of other educators, including faculty and TAs with varying course assignments and levels of instructional responsibility. We do not focus on the logistics of the course per se, because we think course goals can be accomplished in a variety of different course structures and institutions. We also think the variety of methods we use (e.g., writing to probe a subject, definition word scrambles, etc.) are not content specific and can be utilized in anyone's classroom to promote conceptual understanding.

REFERENCES

Bean, J. C. (2001). *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom.* San Francisco, CA: Jossey-Bass.

Friedow, A. J., Blankenship, E. E., Green, J. L., & Stroup, W. W. (2012). Learning Interdisciplinary Pedagogies. *Pedagogy, 12*(3), 405-424.

Garfield, J., Aliaga, M., Cobb, G., Cuff, C., Gould, R., Lock, R., … Witmer, J. (2005). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report.* Alexandria, VA: American Statistical Association.

Green, J. L. (2010). Teaching Highs and Lows: Exploring University Teaching Assistants' Experiences. *Statistics Education Research Journal, 9*(2), 102-122.

Konold, C., & Miller, C. D. (2011). TinkerPlots: Dynamic Data Exploration (Version 2.0) [Computer software]. Emeryville, CA: Key Curriculum Press.

Moore, D. S. (2005). Preparing Graduate Students to Teach Statistics: Introduction. *The American Statistician, 59*(1), 1-3.

Noll, J. (2011) Graduate Teaching Assistants' Statistical Content Knowledge of Sampling. *Statistics Education Research Journal, 10*(2), 48-74.

Woodard, R., & McGowan, H. (2012). Redesigning a Large Introductory Course to Incorporate the GAISE Guidelines. *Journal of Statistics Education, 20*(3). Retrieved from http://www.amstat.org/publications/jse/v20n3/woodard.pdf