

## Using the Five Practices Model to Promote Statistical Discourse

**Summary:** Statistical tasks that can be solved in a variety of ways provide rich sites for classroom discourse. Orchestrating such discourse requires careful planning and execution. Stein, Engle, Smith, and Hughes (2008) suggested five practices to help teachers do so. The five practices can be used to structure conversations so that coherent classroom narratives about solutions to tasks may be formed. In this manuscript, two classroom examples that illustrate the five practices are offered. It is argued that employing the five practices can lead to higher quality classroom discussion than some commonly used arrangements.

**Keywords:** Teaching statistics; Classroom discussion; Sampling; Variability; Informal solution strategies; Comparing groups

## DISCOURSE IN STATISTICS CLASSROOMS

Classroom discussions of rich statistical tasks are of great value. Well-structured discussions allow students to make their thinking public, challenge one another's ideas, and engage in collaborative problem-solving (Franklin et al., 2007). Such discussions help establish classroom discourse communities that value risk-taking and diversity of thought (Morrone, Harkness, D'Ambrosio, & Caulfield, 2004). As students share their statistical thinking, teachers can gain valuable insight about how best to design instruction to address students' learning needs (Cobb, 1999). Hence, discussions have great potential as sites for both teaching and assessment.

Successfully orchestrating classroom discourse is a non-trivial matter. It can be challenging to get students to participate and to keep the conversation focused on worthwhile statistical ideas (Silver & Smith, 1996). Teachers must decide how to organize discussions so that important aspects of the task at hand are discussed. They must also support students without taking over the problem-solving process (National Council of Teachers of Mathematics, 2000). Simply having students volunteer to share their thinking in an unstructured manner is generally not optimal, as it may result in little more than a "show-and-tell" session that leaves important ideas unaddressed (Ball, 2001). Teachers must strike a balance between providing too much and too little structure.

### THE FIVE PRACTICES MODEL

Recognizing the challenges inherent in orchestrating productive classroom discussions, Stein, Engle, Smith, and Hughes (2008) offered a "five practices" model for facilitating discourse. The five practices they recommended were: (i) anticipating students' responses, (ii) monitoring students' responses, (iii) purposefully selecting responses, (iv) purposefully sequencing students' responses, and (v) connecting students' responses. These five practices are meant to provide structure for discussions while still giving students the primary responsibility for thinking and problem-solving.

The five practices model is relevant to facilitating discussions about tasks that lend themselves to multiple solution strategies. Once such a task has been selected, the teacher *anticipates* how students may interpret it and the array of strategies (correct and incorrect) they might use. After posing the task to students, teachers *monitor* their responses. This is done with an eye toward deciding which strategies it would be productive to include in whole-class discussion. Teachers then *purposefully select* strategies to be shared. Purposeful selection helps teachers ensure that an appropriate array of strategies is represented in whole-class discussion. If a key strategy or important idea is not used by any of the students in their responses, the teacher may elect to introduce it. Next, the teacher *purposefully sequences* the selected responses. Teachers might choose to sequence in a variety of ways, such as having the most commonly used strategy presented first, starting with an easy-to-understand strategy, starting with a common misconception, or putting contrasting strategies next to one another to draw out important differences. The overall objective in sequencing, regardless of the particular sequence chosen, is to form a coherent narrative. Finally, as students share their responses, teachers play an important role in *connecting* the responses. At this stage, teachers might compare and contrast the strategies being presented, ask students to do so, emphasize important concepts in the responses, and discuss the efficiency and generalizability of the different strategies.

I employed the five practices model in an introductory statistics course for undergraduate education majors. To do so, I posed homework tasks for which I anticipated various solution strategies. Students submitted their responses to the homework tasks online a few hours before the beginning of the class in which the homework would be discussed. This allowed me to monitor the solution strategies they used and purposefully select and sequence strategies for our in-class discussion of the task. As students presented their strategies in class, I facilitated by connecting the important ideas they contained. In the following sections, I present two examples of how using the five practices model helped promote discourse about statistical ideas.

### EXAMPLE 1: CRITIQUING A SAMPLING STRATEGY

One of the first topics in my course was statistical study design. For homework, I asked students to critique study design ideas similar to those they might hear from children. One such study design scenario was:

A group of elementary school children wanted to determine the percentage of students in school who would like to have pizza in the cafeteria on Fridays. Johnny wants to set up a booth outside of the cafeteria. In the booth, he will have a ballot on which students can vote for whether or not they want cafeteria pizza on Fridays. Students who wish to vote can come up to the booth and cast a ballot.

In posing the task, I *anticipated* a variety of responses. I expected some to struggle identifying flaws in the study design, since the method of voting described in the scenario is very commonly used. I expected that others would reject the study design because it did not use any methods we had discussed in class for drawing a random sample. I thought that some would be able to give specific difficulties that would emerge from having a non-random sample. I also thought that some would recognize it as a self-selected sample and automatically distrust it on those grounds, since we had discussed limitations of this type of sampling in class.

As I *monitored* the responses my students submitted, I noticed some of the expected thinking strategies as well as some I had not originally anticipated. Most responses noted that Johnny's sampling design would be likely to attract only certain groups of children. For example, some wrote that the sampling design might only catch students who eat in the cafeteria, those who had strong opinions about pizza, or those who were Johnny's friends. Some expressed concern that Johnny's plan would exclude students not in school that day. All of these ideas represented legitimate concerns about self-selected samples. In other cases, the primary concern about Johnny's plan was that the sample size would not be large enough. For example, one student wrote, "Although Johnny does have a good strategy by setting up outside the cafeteria, maybe instead he should make posters and have a teacher announce it over the intercom so it's much more publicized." In this response, the only critique was the number of respondents the survey would be likely to draw. Additionally, although virtually everyone recognized that Johnny's sample was non-random, the sampling strategy was referred to by a variety of names, including "self-directed," "self-motivated," and "convenience sampling."

In *selecting* responses to be shared during class discussion, I had a number of considerations. I wanted students to voice a broad array of critiques. Several legitimate flaws associated with self-selected sampling were represented among the responses, but no single response contained all of them. I believed a collective discussion of the flaws would be valuable in prompting students to analyze such situations more thoroughly. I purposefully avoided selecting responses stating that Johnny's sample size was not large enough. I conjectured that

putting the focus instead on flaws in the overall design would help those who believed that Johnny just needed a larger sample re-think their position. I also felt that selecting responses that were off-target in this manner might discourage those with less-developed responses from participating in class discussions in the future. This concern was especially salient for this particular class discussion, as it occurred near the outset of the course.

To *sequence* the class discussion, I chose to start by calling on a student who correctly identified Johnny's sampling strategy as "self-selected" and pointed out that he may get mainly the opinions of his friends. I did this to help establish the language for describing Johnny's strategy and to provide a start on identifying its potential flaws. Next, I called on two students who each had identified additional flaws in the sampling design. The first of the two had noted that only those with strong opinions about pizza were likely to participate, and the second added the idea that children buying lunch were more likely to respond. The final two students I selected to present during class discussion had unique ideas that were not represented in any other responses. The first of these two pointed out that Johnny presumably had no way to determine if someone voted more than once. The second of the two said that a mandatory school-wide survey would yield more reliable results.

As students shared their responses in the sequence I arranged, I facilitated by *connecting* ideas in the responses to one another. At times, this required prompting students to voice all of the important elements in the written responses they had submitted. For example, the student who noted, in writing, that Johnny did not seem to have a way to tell if someone had voted more than once did not share that element of his response without my prompting. I also introduced elements to the conversation that I believed were necessary in order to provide a thorough analysis of the given task. One of the key elements I felt the need to add to the conversation was that Johnny may have been able to improve his design by using random sampling. Although several of my students identified flaws in Johnny's strategy, and they had previously studied random sampling strategies in class, no one suggested that he improve his design through random sampling. I used the final response shared during the sequenced discussion (i.e., the one indicating that Johnny should do a school-wide survey) as an opportunity to introduce the idea. To do so, I asked what Johnny could do if it was not feasible for him to conduct a school-wide survey. This led to a discussion of taking a random sample from the school population to improve the study design.

## EXAMPLE 2: MEASURING VARIABILITY

Another important topic during the semester was measuring variability. Our study of variability began informally, with students analysing data sets and then devising their own ways to describe the spread or variability. Eventually, we incorporated formal measures such as range, interquartile range, and mean absolute deviation (MAD) as ways to provide one-number summaries of variability. After we dealt with these formal measures in class, I assigned a task to determine how students might draw upon them to make sense of data. I presented the data shown in Figure 1 and asked them to decide which team was more consistent in terms of goals scored. Students were also required to explain their answers. I *anticipated* that students would use their own invented strategies for measuring variability as well as one or more of the formal measures of we had studied in order to respond to the task. Ideally, I wanted students to justify their choice of formal measure by making reference to their initial informal analysis of the data.

**Figure 1.** Data set to accompany task on comparing consistency of two teams

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
Bluebirds	3	3	5	4	3	1	6	4	5	2
Redbirds	1	1	0	5	7	2	3	4	7	6

As I *monitored* students' strategies, I noticed an incredibly broad array of approaches to the task. Some students who employed informal strategies showed evidence of reasoning about data as an aggregate, and others used point-by-point reasoning strategies. Those using point-by-point strategies compared scores for the teams from individual weeks against one another, or they looked at week-to-week fluctuation in scores. Students using informal strategies discussing the width of the central cluster of data for each team appeared to be reasoning about data as an aggregate. Students using informal strategies that set a cut score (e.g., 3 goals) and discussed data points in comparison to the cut score made use of deviations from a set point to describe variability. Some students used the formal measures of range and MAD to respond to the task. In many cases, students employed more than one formal or informal strategy to analyse the data and make a judgment about consistency. Finally, I noticed that some students interpreted *consistency* to mean the presence of high scores rather than a set of scores with little variability.

The great diversity of responses to the item made *selecting* responses for class discussion a challenge. Although I considered an ideal response to be one that dealt with the data sets as aggregates, I felt a need to make point-by-point strategies part of the conversation as well, since a number of students had used this sort of reasoning. I also wanted to select responses that would illustrate multiple ways to measure variability. Some responses made reference to multiple measures of variability, but many did not. I conjectured that including multiple strategies in the conversation could help students broaden their array of tools for measuring variability. Another priority was to establish a common understanding of the meaning of "consistency" in the given context, since some interpreted it in an unintended manner. One final goal in selecting responses was to connect informal strategies for describing variability to formal measures. Many responses used one of these two modes of reasoning, and in several cases the choice of formal measure of variability was not justified by reference to an informal analysis of the data.

In *sequencing* the responses for class discussion, I decided to begin with a student who used a point-by-point strategy. This student reasoned that the Redbirds team was less consistent, stating, "They have one week where they scored zero goals and three weeks where they scored either 1 or 2 goals. They have two outliers of 7 which also show their inconsistencies." I thought this response would be a good start for the conversation because it pointed out potentially important features of the data. Next, I selected a student whose response also used informal reasoning, but moved toward discussing the data as an aggregate. This student reasoned that the Bluebirds were more consistent because, "most of the scores are packed together." The third student I selected to present used a cut-score of 3 in her response (since 3 was close to the mean of each data set), and then discussed the number of data points on each side of the cut-score as a measure of variability. I believed this sort of strategy would lay groundwork for discussing the MAD as a measure of variability later in the conversation, since the response incorporated a score close to the mean as a reference point. The next student I selected used the range of the data sets and their central clusters to make a decision about consistency. She stated, "The Bluebirds, again, had a smaller range of 5, while the Redbirds had a range of 7. In only 3 weeks

had the Bluebirds gone out of their typical 3-5 goals, therefore they were the most consistent.” The final student selected to present used the MAD and range together to decide that the Bluebirds were more consistent.

As I *connected* the ideas in students’ responses during the class discussion, I aimed to make several ideas explicit. I asked students to notice that each presenter connected the term *consistency* to the amount of variability in each data set rather than to the typical score. I also asked students to compare and contrast the strategy of counting the number of points on each side of a cut score to the use of the MAD. This provided an opportunity to discuss the advantages of the MAD, as it provides the typical distance of scores from the mean rather than just a count of the number of data points on each side. Perhaps most importantly, the conversation provided a chance to link informal strategies for measuring variability with formal ones. I noted that the informal strategies used by students presenting at the beginning of our class discussion could inform the choice of formal measures. Clusters in the data and outliers can make one formal measure of variability a more accurate one-number summary of a data set than another (e.g., extreme value can easily make the range a misleading measure of variability).

## CONCLUSION

The five practices model offers distinct advantages for managing and guiding statistical discourse. In ordinary classroom conversations, teachers often call on students at random or ask for volunteers. Although such practices are satisfactory in some situations, it is often not optimal for discussions of tasks that can be approached in more than one way. The five practices model provides a structure for making decisions about which strategies are most important to share, the order in which they should be shared, and how they can be connected. Without these considerations, teachers may miss opportunities to draw students’ attention toward thinking that is different from their own, and important ideas may be overlooked. Using the five practices model to organize discourse about complex tasks can help teachers weave classroom narratives that attend to diverse patterns of student thinking while simultaneously addressing important statistical ideas from multiple perspectives.

## References

- Ball, D.L. (2001). Teaching, with respect to mathematics and students. In T. Wood, B. Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 11-22). Mahwah, NJ: Erlbaum.
- Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning*, 1, 5-43.
- 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.
- Morrone, A. S., Harkness, S. S., D'Ambrosio, B., & Caulfield, R. (2004). Patterns of instructional discourse that promote the perception of mastery goals in a social constructivist mathematics course. *Educational Studies in Mathematics*, 56, 19-38.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

- Silver, E.A., & Smith, M.S. (1996). Building discourse communities in mathematics classrooms: A worthwhile but challenging journey. In P.C. Elliot (Ed.), *Communication in mathematics: K-12 and beyond* (pp. 20-28). Reston, VA: National Council of Teachers of Mathematics.
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*, 313-340.