Special issue: Innovative Teaching Personal Essays

Replacing Torture Tests with Broccoli Quizzes

Daniel Showalter

Eastern Mennonite University

"Are your exams really worth it?" my colleague Jess whispered to me during a professional development workshop. Jess worked in the Academic Success Center and had apparently been cleaning up the emotional havoc wrought by my stats exams for several semesters. The simple question sparked a journey that has transformed how I view exams, especially in my statistics class.

It was an elementary stats course for non-majors, and I based their grade primarily on two midterms and a final. Students could do almost whatever they wanted on a daily basis, as long as they could demonstrate their knowledge on the exams. This professor-student contract has withstood the test of time in the ivory tower, right? Jess recounted horror stories connected to my exams: anxiety, sleeplessness, depression, breakdowns, self-doubt. It's not that I was unaware of how stressful exams could be; I had spent a year in counseling during grad school for test anxiety. I, like many of us, just hadn't found an alternative assessment that would accurately check knowledge and be practical to implement in a large class. I gave some points for homework, investigations, and a project, but I didn't want to weight those areas more. I had seen too many students find crafty ways to minimize learning through such assessments. When I relayed these excuses to Jess, she said, "Then I, at least, hope the exams are worth it."

Were my exams worth it? Was there another way that would give me similar info but inflict less suffering? By serendipity, I had some National Science Foundation (NSF) funding to support conducting a research project in one of my classes. The grant was designed to support under-represented minority students and first generation students, two populations who seemed to suffer the most at the hands of my stats exams. Moreover, the timing was perfect to overhaul my stats exams because the grant would pay for the use of instruments to measure any changes that occurred in the process. I just needed to find a new vision for the course! I spent that that summer mulling over the question and reading up on some stats education literature. I was particularly inspired by Garfield and Ben-Zvi's (2009) Statistical Reasoning Learning Environments. The respected duo painted a warm picture of a statistics class heaven, where students engaged in stimulating dialogue as they investigated issues dear to their hearts through a statistical lens. The formulas and procedures that confound students took a back seat to understanding the deeper concepts and reasoning that could feasibly improve the quality of the students' lives. It was almost too good to be true -- a stats class that might actually fulfill those crazy institutional learning outcomes stamped onto my syllabus! Although I had been trying to use more authentic datasets, I doubted that most of my students would say the class was improving their lives.

I eventually set two goals for my stats tests: (a) evoke less anxiety and (b) be more meaningful (beyond measuring knowledge). Jess claimed I could accomplish the first goal simply by replacing my high-stakes midterms with more frequent quizzes. I felt fine with this, as long as I kept my final exam as a security blanket for course grade validity. Another colleague warned that frequent quizzes could reduce content coverage since I would designate more days for testing. Garfield and Ben-Zvi's (2009) article helped me justify removing some focus on

formulas because these new assessments, in theory, could address my subtler objectives of empowering students to apply statistics to everyday issues.

To craft purposeful quizzes, I appealed to the vision of Garfield and Ben-Zvi (2009) as well as similar articles. I split the course content into eight chunks, each of which would go with an "authentic quiz." But, rather than start with the statistical content of each chunk, I selected a relevant topic as determined by past conversations with students. These included child poverty, global female education, compulsory math courses, healthy eating, racism, quality of life, depression, embracing diversity (of religion, race, politics, and socioeconomic status), and gender identity. *Get this: In trying a new teaching intervention, I had accidentally carried out the Dean's encouragement to brand my course content with elements of the university's mission.* Of course, there was also student interest in topics like video games, dating, and alcohol, but as a parent of two young kids, I had enough "eat your broccoli" mentality in me to prioritize topics of societal impact. The bulk of my work came through pairing up each "broccoli" topic with the corresponding statistical content in that chunk of the course. I laboriously sought ways to present my students with authentic data so they could legitimately explore each topic. Luckily, it was summertime, and we professors don't have *any* work to do in the summer…

The format went like this. On an authentic quiz day, I randomly paired students up. I then facilitated a 15-minute discussion on the topic. The discussion always involved personal reflection, peer sharing, and debates over how to measure the issues. The students then had 35 minutes for the quiz. The first half of each quiz began with authentic data that allowed them to use statistical methods we had recently learned to explore that day's topic. I had students exploring gray areas that pushed them to think carefully and sometimes uncomfortably. For example, on a test about unschooled children, I asked them to create an accurate but potentially

misleading data visualization to overemphasize the gender inequities in the schooling of low-income countries. The next question asked them to discuss the ethics of using that visualization if they were in charge of fundraising for a nonprofit organization focused on raising money for female education worldwide. This format finally allowed me to sneak in some questions that were higher up on Bloom's taxonomy, something I had previously neglected to do in stats beyond the obligatory "justify your answer." The second half of each quiz was routine statistical content that allowed me to address learning objectives not dealt with by the context-based exploration.

The implementation was surprisingly smooth. The lone hiccup was that some students struggled to prepare for the first half of the quiz since I asked questions unlike anything we had done in class. I reasoned that life is filled with statistical issues that don't appear the same way they do in a stats class. However, I did start giving more practice with spontaneous statistical reasoning; I began each class period with a short discussion on something I had come across and asking the class how we could use stats to inform the issue (I even dropped my broccoli standards and included "discussion candy" to get the entire class engaged).

In sum, the innovation process started with years of growing awareness on my part that many students did not enjoy my statistics course. When a trusted colleague told me how much test anxiety surrounding my class was impeding students' learning, and even well-being, I finally resolved to change my exams. I decided to find and implement a plan immediately so that I could leverage NSF funding to capture any changes to students' affect and knowledge. I surfed around Google Scholar with terms like "elementary statistics" and "test anxiety" until I came across an article by Garfield and Ben-Zvi (2009) that captured what I hoped to achieve. As articulate as their theory was, putting it into practice was by no means trivial. I put myself into my students'

place to list some topics they would find interesting, then whittled the list down to the more meaningful topics, and finally sought out data and ways in which the statistical content could inform these topics. Using student feedback, I tinkered with the logistics until I had a set of assessments that I felt good about; seeing the students engage with the class discussions, and then the quiz itself, was a welcome contrast to the clouds of despair that used to fill the room on exam days. That alone convinced me to sustain the innovation, although this decision was bolstered by their performance; students handled procedural questions on the final exam about as well as they had under the old exam structure, but their responses to the open-ended questions with authentic data were of much higher caliber. If the students' stress hadn't decreased or their performance on the final exam had dropped, I probably couldn't have justified the extra class time the authentic quizzes required. But with happier, more knowledgeable students, deciding to keep the innovation has been a no brainer – even if it would have required substantial time and energy (which, thankfully, it hasn't after I made the original set of assessments).

While students probably didn't look forward to the quizzes, they were less stressed and most students sincerely bought into them. Frequently, students gave touching accounts of how the topic related to their lives – with their discussion partner, the entire class, or (most commonly) in their responses to the interpretation questions on the quiz. Sometimes, a quiz's topic would initiate a series of emails or office hour talks with an interested student (prior to the authentic quizzes, students had only emailed me to debate the grading of a question!). Many students have commented on the relevancy of the course to their lives in their end-of-course evaluations. Even better, they have shared stories of using stats to inform decisions about exercising, eating, family health issues, and career paths. In short, the cost of my exams is lower

and the benefit is higher. I still have far to go in shaping my class, but I can say in good conscience that my exams are worth it.

Acknowledgements: Funding provided by the National Science Foundation, DUE-1611713.

References

Garfield, J., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, *31* (3), 72-77.