TESTING, TESTEES, AND TESTED: PRACTICAL LESSONS FROM THE FIRST YEARS AT A SMALL TEACHING-FOCUSED UNIVERSITY

R. Adam Molnar
Bellarmine University, United States of America
amolnar@bellarmine.edu

Changes in rank, responsibility, and duties make the first few years as a University professor substantially different from graduate student life. The shift comes with several stumbling blocks. This paper describes issues encountered at a small teaching-focused university in the United States. The shift to full responsibility for multiple courses, including syllabus and test design, involves skills acquirable through training and study. Less bookish, yet no less important, is learning about the students; their motivations and styles differ from those of graduate students and professional scholars. Then, dealing with academics through University politics has potential pitfalls. Also discussed are topics particular to small schools, such as the wide variety of requests for general statistics advice.

Signing a contract for the first full-time academic position is an understandably joyous occasion. Generally, it grants a substantial increase in economic status. It also provides a better title, replacing "Graduate Student" with the more impressive "Lecturer" or "Assistant Professor." After the struggle of a dissertation defense, it might seem like a bed of roses. Unfortunately, as in gardening, with roses come thorns. This paper looks at thorns that I have encountered in my first three years of full-time teaching. By doing so, I hope to remind mature faculty of potential problems, so they might better assist new faculty. Some of the issues, but not all, can also be incorporated into the transition from doctoral programs.

The title, "Testing, Testees, and Tested," gives the three areas in this paper. Testing refers to the mechanics of administration for a course, such as books, schedules, and tests. Here, formal guidance can be provided. The second section, Testees, focuses on the student population. After years working with professors and graduate students, there are things to learn not only about pupils' experiences, but also their motivations. The final section, Tested, describes the struggles of negotiating a new work environment, particularly the strange workings of academic politics.

All three sections contain personal anecdotes, so I'll begin by describing my workplace. In January 2007, I began in the Mathematics department at Bellarmine University in Louisville, Kentucky. Originally founded by the Catholic Church, it's now independent but still draws a fairly high proportion of Catholic students. Compared to most countries' state schools, it's small, with roughly 2200 undergraduates. About 700 graduate students attend Masters and professional programs. Most students grew up less than 150 km away. About half the undergraduates live in on-campus dormitories, a percentage rapidly increasing.

Bellarmine advertises small classes, and the university generally draws above average students looking for individual attention. Compared to research universities, professors teach many sections–I have seven this school year–but courses rarely contain more than 35 students. Most majors require introductory statistics, with 15 sections in 2009-10. Of these, I have a plurality, but not a majority. The department also offers three or four courses each year that prepare mathematics majors for actuarial work. I teach half the upper division courses, while a Math PhD who worked as an actuary handles the other half. As you might expect, faculty have specialized roles. An important point is that I am the only faculty member with degree work in Statistics.

TESTING AND COURSE DESIGN

The small school experience includes a lot of grading without assistance, since students expect frequent assignments and rapid feedback. My courses typically include homework, two to four hour exams, a project, and a three-hour final exam. A large portion of my time is taken with writing and grading tests. At first, this was surprisingly difficult. I had taken many tests during my 20 years of schooling, but tests are like vaccines; even though we've received many shots, that doesn't mean we know how to find a vein. Mistakes were made. In my second lectureship in graduate school, a young professor and I wrote six thoughtful, balanced questions. Unfortunately,

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

in the time provided the average student could only attempt four. The first time I taught probability, I tried to define a bivariate area of integration, but forgot one limit and turned a diamond into an infinite space. Then, there was my first attempt at writing multiple-choice questions, because I thought the provided text bank had too many "all of the above" answers. Out of ten questions, I had to throw one out and count multiple correct replies on two others.

Some of my errors arose from carelessness, like my failure to properly write the limits of integration. Those will just happen. Others, though, could be avoided through instruction. There are many resources on test construction. A 2006 list from UMDNJ in New Jersey provides links to many online, including those at Tennessee-Chattanooga (Adsit, 2003) and BYU (Zimmerman, Sudweeks, Shelley, & Wood, 1990). I wish I had seen them before I hit the thorns. New professors could receive this information alongside the handbook on work rules, benefits, and policies. Ideally, this would be part of graduate training, but it's rare enough to warrant mention in Meng's 2009 article on the future of statistics. "Is it true that we used to put up teaching fellows without any training?" a student asked. It happens, but that's a different subject.

Another subject for new professors is course design. I've been asked to design two new courses, statistical literacy for undergraduates and intermediate methods for graduate nursing students; I've also revamped the undergraduate stochastic processes course. As students, we have seen how courses flow, some smooth, others lumpy. Moving to the captain's seat is another jump. Some things require live practice, like flowing from lecture to lecture and handling disruptions. Other things can be provided. I've seen seminars for humanities students on syllabus design. Resources are plentiful; BYU (2006) provides a good starting point. It would greatly help new faculty have them inside the new faculty handbook. I would also include department standards on grading and reasonable student expectations.

A last useful addition would be a list of textbook publishers in the region. For each course development, the science division secretary and I have had to work through a large list of publishers. Until there is a worldwide online list of publishers and texts by subfield, the list of relevant statistics booksellers would save time.

TESTEES, MOTIVATION, AND LANGUAGE

Earlier, I mentioned student expectations as something that needed explanation. Someone established at a school knows the level of the students. For a new professor, it is very easy to assume that these undergraduates will be like the last place, the graduate school, or like some other known experience. Sometimes that's true. Often it's not, such as my shift from the world-aspiring University of Chicago to a regional Catholic college. I solved the problem by pestering other faculty for guidelines. In 2008, I aided an adjunct night professor with this problem. She thought that college students would be stronger than her high school pupils. Since statistics is an advanced elective in Kentucky secondary schools, while most students have to take college statistics, the opposite was true. She could adjust after that realization, and things went much more smoothly.

Besides the difference in knowledge levels, there can be cultural differences. I had a bigger gap than most. The University of Chicago has a fairly well deserved reputation as the place "where fun goes to die," enough to make Wikipedia (n.d.). After two weeks at Bellarmine, I wrote to my friends about a nursing student who dotted her I's with hearts (Molnar, 2007).

Culture speaks to motivation, another point of divergence. As academics, generally we've gone through the starkness of doctoral programs because we like the subject. We proceed in the face of the negative reactions mentioned by Meng (2009) and others, like the common "Sorry, but I really hated my statistics course" (p. 203), Graduate students avoid much of this because they spend lots of time around other people who like academics, and have motivation for learning. In general, undergraduates are less driven by class activities, as many others have noted. For instance, in 2005 Nathan (2005) wrote in *My Freshman Year* that for her subjects "Non-class-related learning was reported as high as 90 percent for some, and few students ranked class activities as constituting more than 50 percent of what they learned in college" (p. 101).

Different motivation is not wrong motivation. For instance, to maintain my overall weight and health I work out at the gym. The basketball team also trains there, so they can play better. Both are completely acceptable. Just as the basketball team should not criticize why I sit on the bicycle as long as I exercise (yet occasionally comments leak out), professors should not critique

why pupils sit on the chair as long as they participate. I occasionally forget this, and I hear others do the same. This is particularly true for young faculty, who are moving from an environment where people have high interest. It is not fair to me at the gym, and it's not fair to the students.

Another subtly unfair point is to rely too heavily on symbols. In graduate school, around other mathematical types, abbreviations were commonplace. Students, particularly math-phobic ones in introductory classes, struggle with translation. Croxton wrote about the symbol problem in 1941; his words still hold today. To quote, "there must be scores who are interested in obtaining a working knowledge of the basic principles of statistics, so that they may occasionally read the results of statistical studies or occasionally apply modest statistical procedures in their every-day work, to every one who enters the field intending to specialize." For this majority, too many symbols turn them away. It's little more work to write Null instead of H_0, or false rejection for Type 1 error, yet it adds a great deal of meaning.

I was reminded about abbreviations at a recent mathematics talk. The speaker began by saying "poset" several times, not immediately defining it as a partially ordered set. I began the talk confused. Midway through the talk, references to "P" appeared on several slides. I was lost. Where did P come from? Eventually, the speaker casually remarked that P stood for poset, and confusion cleared, but I had spent a few minutes puzzled. It has been suggested (Limerick 1993) that such times of confusion are intentional, with specialized language designed to make academics feel separate and important. I disagree; most of the time, as with this speaker, it's unintentional complication. Whatever the reason, vigilance to avoid jargon is still necessary.

TESTED, POLITICS, AND GENERAL KNOWLEDGE

Another necessary part of academic life is dealing with other departments and the administration. At a large school, a new professor is part of a large group, and tends to work inside the department and division. At Bellarmine, I interact with faculty from all divisions, and received a University-wide committee assignment in my first full year. Other scholars have different perspectives. For instance, a sociologist wondered about the length of time taken for statistics projects. I explained that not only do statisticians write about subjects, we frequently have to write new code, and that coding can take extremely variable amounts of time. While other statisticians would know this, and other scientists understand, this was new to her.

Professional isolation leads to other types of questions. As the only statistician on campus, people come to me for advice on a potpourri of topics. I've been asked about regression, experimental design, content analysis, SPSS, SAS, and many other issues not generally included in dissertation research. Of special note is that I get questions on study design often enough that I've had to review things. People rarely ask me about neural networks, even the Bellarmine professor whose paper included a non-Bayesian neural network (Sinski, Smith, & Smith, 2001). Being a generalist is much more important at a smaller school. Tasks come from many different places.

A good example occurred in 2007, when at the start of my second semester, the administration announced hiring priorities for the upcoming year. One position was for psychology, potentially in psychometrics. This surprised the psychologists, as they hadn't asked for that specialty. Inquiries revealed a complaint from the Communication department. This division had required Elementary Statistics for their mathematics course. Bellarmine's course took, and still takes, a relatively rigorous approach; at the time, the book was the fifth edition of Moore and McCabe's *Introduction to the Practice of Statistics* (2005). For weaker students studying Communication, the challenging content led to complaints. Stronger students handled the work, but didn't see the relevance of computation and probability. The grumblings of the department, and to a lesser extent the Nursing department, were heard by the central administration.

The suggested resolution was a "social science statistics" course, which led to the decision by management to hire a new faculty member. However, this was not the best option, Psychology and Math agreed. I wound up on the search committee for the Psychology position, which hired a specialist in child development. I also planned to introduce a course focused on statistical literacy, with more writing, descriptive statistics, and computer support. That summer, I had heard about such a course at USCOTS (Rossman, 2007). My trip report included the idea, as something to consider when we had time. That time had been moved up. To plan the course, I talked to people from across the university, including Communication, Nursing, Psychology, Political Science,

Global Languages, and Biology. After the discussions, the Math department decided to offer Nursing, Communication, and humanities majors the new course called Statistics and Society, using *Seeing through Statistics* by Jessica Utts (2004). We implemented it for the fall of 2008. The Communication and Nursing students find it more relevant, and faculty members are pleased with what the students are learning.

This story illustrates the complexity of working at a small school. At a larger school, it would be very unlikely for a junior instructor to serve on a search committee for faculty outside his own area, or design a new introductory course. Here, there's not much choice. Additionally, I had to determine the motivation behind the comments. There were solid points about content, not just complaints about difficulty. Fortunately, I had departmental support, and helpful people in other faculties; nevertheless, I was tested.

CONCLUSION

The tests and thorns as a new professor differ from those as a graduate student. While my story about social science statistics might be unusual, it illustrates struggles not expected while poring over issues of *Biometrika*. Some issues require support; others require experience; still others improve with training. By writing about my mistakes and troubles, I hope to improve training and programs for new professors, keeping the new job closer to rose petals than thorns.

REFERENCES

Adsit, K. I. (2003). Designing Test Questions. Retrieved November 30, 2009, from: Teaching Resource Center at the University of Tennessee at Chatanooga website. Online: www.utc.edu/Administration/WalkerTeachingResourceCenter/FacultyDevelopment/Assessment/test-questions.html

BYU Center for Teaching and Learning. (2006). Syllabus Design. Retrieved November 30, 2009, from http://ctl.byu.edu/home/tools/syllabus-builder/.

College of the University of Chicago. (n.d.). In Wikipedia. Retrieved November 29, 2009, from http://en.wikipedia.org/wiki/College_of_the_University_of_Chicago.

Croxton, F. E. (1941). Towards Standardized Symbols for Basic Statistical Concepts. *Journal of the American Statistical Association, 36*, 426-428.

Limerick, P. N. (1993, October 31). Dancing with professors: The trouble with academic prose. New York Times Book Review. Reprinted in Something in the Soil (2000). New York: W. W. Norton & Company.

Meng, X.-L. (2009). Desired and feared – what do we do now and in the next 50 years? *American Statistician, 63*, 202-210.

Molnar, A. (2007, January 13). Hearts [Web log message]. Retrieved November 28, 2009, from http://mathematicsafterthefall.twelvefruits.com/archives/66

Moore, D., & McCabe, G. (2005). *Introduction to the Practice of Statistics, 5th edition*. New York: W. H. Freeman.

Nathan, R. (2005). *My Freshman Year: What a Professor Learned by Becoming a Student*. Ithaca, NY: Cornell University Press.

Rossman, A. (2007, May 18). Seven Challenges for the Undergraduate Statistics Curriculum in 2007. *U. S. Conference on Teaching Statistics*.

Sinski, J.F., Smith, G.C., & Smith, J. (2001). Spectroscopy based neural networks: using 3-dimensional fluorescence spectra and artificial neural networks to match petroleum contaminated ground water with possible sources. *Scientific Computing & Instrumentation, 18*(9), 22-30.

UMDNJ Center for Teaching Excellence. (2006). Student Evaluation – Test Construction and Item Writing. Retrieved November 29, 2009, from: http://cte.umdnj.edu/student_evaluation/evaluation_constructing.cfm.

Utts, J. (2004). *Seeing through Statistics, 3rd edition*. Florence, KY: Cengage Press.

Zimmerman, B. B., Sudweeks, R. R., Shelley, M. F., & Wood, B. (1990). How to prepare better tests: Guidelines for university faculty. Retrieved from Brigham Young University Testing Services website: http://testing.byu.edu/info/handbooks/bettertests.pdf.