New Perspectives: A Statistician and a Statistics Educator Discuss the Lessons Learned from Cross Disciplinary Sojourns

Jennifer J. Kaplan[1,3] and Vincent Melfi[2]
[1] University of Georgia, Athens, GA, USA
[2] Michigan State University, East Lansing, MI, USA
[3] Corresponding author: Jennifer J. Kaplan, e-mail: jkaplan@uga.edu

Abstract

Within science and mathematics, including the field of statistics, there is increasing evidence that although instructors are aware of research on teaching and learning, changes in instructional practices are uncommon or implemented haphazardly with a lack of persistence and/or evaluation. Briefly, there is a substantial gap between the research and its large-scale implementation. In addition, most researchers who study the teaching and learning of statistics, statistics educators, are trained in departments of educational psychology, mathematics education and/or mathematical sciences and not in departments of statistics. As such they may not have been fully immersed in the discipline of statistics. Having one or more statistics educators on faculty in a department of statistics can address both of these issues simultaneously. There is the opportunity for disciplinary statisticians who instruct students to discuss research on teaching and learning with experts, to see research-informed teaching in practice and to evaluate such teaching practices on student outcomes. The statistics educators not only inform the department of innovations in teaching and learning of statistics, but also learn the true nature of statistical research, both in terms of the cycle of enquiry in which statisticians function and the types of problems valued by the discipline. Statistics education research becomes part of the culture of the department and statistics educators are acculturated into the domain of statistics. This paper will illustrate what one statistician and one statistics educator, working together in a department of statistics, learned from talks given in the other discipline. The paper will conclude with results from a research project on student motivation completed by a research team of statisticians and statistics educators.

Key Words: Collaborative research dissemination, cross-discipline professional development, Undergraduate learning.

1. Introduction

In STEM disciplines there is increasing evidence that although instructors are aware of research on teaching and learning, changes in instructional practices are uncommon. In addition, it is common that even when instructional practices are changed in response to research, these changes typically are not persistent (Henderson, Dancy & Niewiadomska-Bugaj, 2012; Fairweather, 2005; Ebert-May, Derting, Hodder, Momsen, Long, & Jardeleza, 2011). Briefly, there is a substantial gap between the research and its large-scale implementation. Upon reflection this is not surprising, since although reading a research report conveys some understanding, it is unlikely to convey the depth of understanding that a researcher in the field would have. If a department of statistics has one or more statistics educators on faculty, there is the opportunity for disciplinary statisticians to discuss the research with experts, to see research-informed teaching in practice and to evaluate such teaching practices on student outcomes. Put another way, statistics education research (and its implementation) become part of the culture of the department.

Similarly, most statistics educators are trained in departments of educational psychology, mathematics education and/or mathematics sciences and not in departments of statistics. As such they may not be acculturated to the discipline of statistics. Having a faculty appointment in a department of statistics allows a statistics educator not only to inform the department of innovations in teaching and learning of statistics, but also to learn the true nature of statistical research, both in terms of the cycle of enquiry in which statisticians function and the types of problems valued by the discipline. This paper will illustrate what one statistician and one statistics educator, working together in a department of statistics, learned from talks given in the other disciplines. The paper will conclude with results from a research project on student motivation completed by a research team of statisticians and statistics educators.

2. A Statistics Educator becomes a Statistician

For the first author of the paper, an active statistics education researcher who completed inter-disciplinary training from within a department of mathematics, her first academic appointment was her first opportunity to engage academically with statisticians as a member of their department. In order to properly enculturate into the department, she attended all functions, including departmental colloquia. From sitting in on statistics colloquia, she learned the story arc of a typical statistics talk. In general the story begins with a motivating example, which may be a new type data set or collection method and/or a research question in a new field of study, generally outside the domain of pure statistics. The speaker then explains why the existing statistical methods are not sufficient in the new setting and follows with an explanation of a new method, which is the subject of the talk. Speakers describe the conditions under which the method can be applied, the corresponding mathematical assumptions and generally provide asymptotic results and/or simulation studies that compare the results of the new method to those that had been used previously. The speaker tends to then explain what was found in the motivating example and concludes with further directions the statistical research might take, for example, if data become available that do not meet one of the assumptions of the new methodology.

There are two ways in which learning the story arc and nature of statistical research informed the way the first author teaches undergraduate introduction to statistics courses. The first is the position and emphasis of experimental design in the curriculum. Prior to joining a department of statistics, she found the section on experimental design, or data collection, to be rather dry. She actually dreaded that portion of the course, wondering how to make a discussion of the difference between observational and experimental studies or stratified and cluster sampling interesting. Once she became integrated into the community of statistical practice, she understood how integral experimental design and data collection are to informing analysis. While she had previously been exposed to the cycle of empirical enquiry (Wild & Pfannkuch, 1999) placing the research question as the first stage of statistical inquiry followed by design and data collection, it was not until after listening to statistics colloquia that she fully internalized this cycle. As a result, she moved the experimental design and data collection material to the beginning of the semester, even when that was not the order in the textbook, spent a lot more course time on the material and developed more interesting class activities for the material.

The second way in which her teaching was informed was in the area of statistical inference. Previously, she had given cursory attention to the checking of conditions when performing inferential procedures. While she knew of its importance and was sure to “cover” that material, she found it to be dry and did not appreciate its integration into the cycle of empirical enquiry. Through colloquium attendance, she learned to explain to her students that when conditions for the test under discussion in class were not met, it was not the case that one would have to throw up our hands, abandon the analysis and make no conclusions. She could now explain that, in fact, this is what research statisticians do; they design new methods of analysis for cases in which the conditions for known analyses are not met. As a simple illustration, it provided context from the domain of statistical research for the story of William Gosset and the development of the t-distribution. In this way she was able to help her students to understand what it means to "do statistics."

One final way in which the first author benefited from her faculty position in statistics was the availability of consulting advice when analyzing quantitative data collected in an educational setting. The data in question were collected using a repeated measures design with several predictor variables, some categorical and some quantitative and an outcome variable that was dichotomous, whether a student had answered the question correctly or incorrectly. In the original manuscript, the repeated measures data were analyzed independently. The reviewers sent back the manuscript asking that a correct analysis of the data, taking into account the repeated measures design, be done. By this time the first author was faculty in statistics. She was advised by her colleagues to hire one of the Ph.D. students in statistics who worked at the statistical consulting center. This was done, and the paper was published successfully with the Ph.D. student as the second author. The collaboration, therefore, benefited not only the first author, but also a doctoral student while expanding the research base in the field of statistics education, both methodologically and based on the findings (see Kaplan & Du, 2009).

3. A Statistician becomes a Statistics Educator

Just as regularly attending disciplinary statistics colloquia can benefit a statistics educator's research and teaching, the same is true when a disciplinary statistician regularly attends talks on statistics (and more generally, mathematics) education. For example, these talks often provide examples of the use of data in studying the effectiveness of instructional practices. Now it might seem counterintuitive that a statistician would need to attend a statistics education talk to think more carefully about using data(!), but as Gelman and Loken (2012) point out, statisticians typically do not make use of this opportunity to practice their craft. The day after the first author gave a colloquium in the Department of Statistics on student uses of the word *random*, research statisticians reported having asked undergraduate students in their classes to use the word random. They reported that the findings from their students matched the findings presented in the talk. While this may not represent a sustained change in the department, it is a first step toward helping disciplinary statisticians think about using data collected from their students to inform changes in teaching and curriculum. Another example from our department is the use of personal response systems (PRS or clickers), which was initiated by the first author and adopted by other faculty members in the department. Using the automatic collection of easily accessible data from students, statisticians began tracking things like attendance, something that had previously been an issue in the large lecture courses, and began to think about studies they could do using the student level data to improve the quality of the courses.

The research tradition in education is quite rich, including not only the use of quantitative methods familiar to the statistician, but also qualitative and mixed methods. Learning about these diverse data collection and analysis methods is another benefit from the attendance of education colloquia by disciplinary statisticians that also helps the statistician to think more carefully about the role and relevance of statistical methods more generally. Consider as an example an education colloquium in which two education researchers, one who works mainly in the qualitative tradition and the other in using quantitative methods, had a dialogue. The colloquium provided both a window into the rich methods used in qualitative research and an opportunity to think about the undue respect sometimes given to quantitative research because of its seeming objectivity. At the colloquium the second author found it necessary to voice concerns about this, resulting in the somewhat odd spectacle of the only statistician in the audience being the one who was arguing for some caution in applying statistical methods!

As a third example of the benefits of interdisciplinary colloquium attendance, the culture at statistics education talks seems to be more encouraging of active participation by the audience and robust discussion and debate than is the case for statistics talks. A statistician can again think about how such active participation and discussion might enrich not only teaching but also research in statistics. As an example of the former, at a recent mathematics education colloquium participants were given a rubric for scoring a brief classroom experience, shown a five minute video of a classroom, and then asked to score the video using the rubric. It was no surprise that the mathematics education faculty were quite engaged in the task and ensuing discussion. More interestingly, a group of undergraduate elementary education majors who attended the colloquium were also very engaged and willing to join in the discussion. Most statistics instructors claim to want more interaction and discussion in their classrooms, so examples like this are both inspiring and educational. For the latter, the discussion at statistics colloquia and more informal statistics gatherings (e.g., guidance committee meetings) often focus on narrow technical points. Statisticians could learn from education colloquia, at which the discussion is typically broader and aimed more a fundamental questions. This sort of discussion within statistics could help to focus some of the attention on more fundamental issues.

4. A Joint Research Project

The authors of this paper were members of the Quantitative Literacy (QL) Working Group, a group tasked with creating QL instruments to be used by the university to assess the QL level of incoming students and to assess the effectiveness of the university QL program. One of the statisticians on the QL working group was concerned that the results of the assessments were underestimating the actual QL competency of the students. The researcher felt that the low scores on the assessments could be due in part to lack of effort by the subjects. This researcher thought that paying students for correct answers, rather than for their time, would provide more motivation for students to answer correctly and would produce responses more representative of the actual level of QL competency of students at the university.

The statistics and mathematics educators in the QL working group disagreed with the statistician, arguing from the perspective of motivation theory. According to the Self-Determination Theory of motivation (Ryan & Deci, 2000), providing payment for completing the task meant the students already had extrinsic motivation to complete the task. Extrinsic motivation has been shown to be less powerful in general than intrinsic motivation, particularly in education related settings (ibid). Motivation, however, is a complex construct. Extrinsic motivation has many levels. Different types of rewards, such as rewards for completion and rewards based on performance similar to those suggested by the statistician on the project, have been shown to have independent and negative associations with performance on academic tasks (Deci, Koestner, & Ryan, 2001). Furthermore, associations have been found between students' self-efficacy, identification with the academy, level of motivation and level of engagement with academic tasks. In particular, higher levels of self-efficacy, identification with the academy and motivation contribute together to predicting higher levels of task engagement (Walker, Greene & Mansell, 2006). Using monetary rewards further complicates the issue. In a meta-analysis of experimental papers in which the level of monetary reward was the manipulated variable, Camerer and Hogarth (1999) found that monetary incentives do not always lead to improved performance. This was true in particular when the tasks were easy, requiring little cognitive engagement. Camerer and Hogarth also note that there were few studies with tasks that required a high level of cognitive engagement so a complete model relating reward structure and cognitive engagement level of tasks could not be hypothesized.

The QL working group designed an experiment to test the hypothesis of the statistician. Subjects for the experiment were students recruited from a computer science course focused on problem solving. Eighty four percent of the students in the class were students in either the College of Engineering or the College of Natural Sciences and nearly 70% had second year or higher standing at the university so these students are considered to have high self-efficacy and identification with QL-type outcomes. Prior to volunteering all students in the class were told that they would be compensated for their participation, but not the amount of the compensation and 65 of the 107 students registered for the course volunteered to participate. The students were asked to go to one of two rooms at the end of class based on the last digit of their student identification number. The rooms had been assigned as “control” and “experimental” based on a coin flip prior to the start of the experiment. Students in both rooms were given the same 16-question QL assessment. The directions on the assessments were identical except that the directions on the assessment for the control group stated that students would receive $10 for completing the assessment and the directions for the experimental group stated that students would receive $1.50 for each correct response. For both groups, after the subject completed the assessment s/he was directed to a table in the hall outside the rooms where the subject was given an envelope with $24 and a written explanation of the deception. Two response variables were measured for each subject, the time to complete the QL assessment and the number of correct responses on the QL assessment.

Table 1: Results of Compensation Experiment

		Time to Complete (min)		Score (out of 16)	
	N	Mean	SD	Mean	SD
Control	31	22.6	4.08	8.71	3.13
Experimental	34	25.8	4.61	8.29	2.68

The results of the experiment, shown in Table 1, indicate that the students in the experimental condition spent more time, on average, than the students in the control condition. This difference was statistically significant (p-value < 0.01, two-sided, *t*-test). While the experimental group scored lower, on average, than the students in the control condition, the difference was not statistically significant (p-value $= 0.57$, two-sided, *t*-test). In the language of motivation theory, these results suggest that while the students who were being compensated based on performance had higher cognitive engagement, this did not lead to higher performance than that of the students being compensated for completion of the task. In terms of the hypothesis put forward by the statistician who suggested that low scores were due to a lack of motivation to do well that could be addressed through a performance payment structure, the results of the experiment did not support his hypothesis.

5. Discussion

In this paper we have discussed a number of benefits to including statistics educators as faculty within a department of statistics. For the statistics educator, there is the opportunity to be enculturated into the discipline of statistics. This enculturation can benefit the educator in the classroom and in research endeavors. For the statisticians, particularly those who regularly attend education colloquia, there is not only the opportunity to learn about research that will inform and improve teaching, but also how to do such research and the myriad of research methods available for such research. Finally, working together in interdisciplinary research groups ensures that proper care is taken in all aspects of research, from the literature review, through the research design and data collection, analysis and presentation of findings appropriate to the audience, be it a journal article or research talk. We conclude this paper with one caveat, because the benefits of including statistics education research in a statistics department is not without challenges. Care needs to be taken to be sure that the statistics educator is getting proper mentoring from both statisticians and educators and that the mentoring is aligned with the standards for promotion of the department. In addition, some education at both the department and college level should be done so that faculty who make promotion decisions are prepared for and feel qualified to judge the promotion documentation of a statistics education researcher. We do not mention these challenges to discourage others from proceeding down this path, but rather to inform other so that they may reap the benefits described above.

References

Camerer, C.F. & Hogarth, R.M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19 (1-3), 7 – 42.

Deci, E.L., Koestner, R. & Ryan, R.M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Education Research,* 70(1), 1 – 27.

Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *Bioscience, 61*(7), 550 – 558.

Fairweather, J. (2005). Beyond the rhetoric: Trends in the relative value of teaching and researching faculty salaries. *Journal of Higher Education, 76*, 401 – 422.

Gelman, A. & Loken, E. (2012) Statisticians: When We Teach, We Don’t Practice What We Preach. *Chance*, 25(1). http://chance.amstat.org/2012/02/ethics-25-1/

Henderson, C., Dancy, M., & Niewiadomska-Bugaj, M. (2012) The Use of Research-Based Instructional Strategies in Introductory Physics: Where do Faculty Leave the Innovation-Decision Process?, *Physical Review Special Topics - Physics Education Research,* 8(2), 020104

Kaplan, J.J. & Du, J. (2009). Question Format and Representations: Do Heuristics and Biases Apply to Statistics Students? *Statistics Education Research Journal, 8(2),* 56 – 73. http://www.stat.auckland.ac.nz/serj

Ryan, R.M. & Deci, E.L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54 – 67.

Walker, C.O., Green, B.A. & Mansell, R.A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences*, 16, 1 – 12.

Wild, C.J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223 – 248.