What do our students value in instruction?
Structural topic modelling of student evaluations of teaching

JEON Ah Jung[1] and Greg TUCKER-KELLOGG[1, 2] *

[1]Department of Biological Sciences, Faculty of Science, National University of Singapore
[2]Computational Biology Programme, Faculty of Science, National University of Singapore
*dbsgtk@nus.edu.sg

Student evaluations of teaching (SET) are widely used tools for assessing teaching quality. SETs are almost always based on surveys in which students provide numerical ratings of teachers or classes, while providing open-ended text feedback at the same time. SET scores—calculated from the numerical ratings—are used for a wide variety of purposes by educational institutions, including decisions on hiring, promotion, and tenure. Despite their widespread use, SET scores have been the subject of intense criticism (Spooren *et al.*, 2013). A long-recognised weakness of SET scores is that they may not reflect student learning outcomes, but SETs have other problems as well: they are confounded by gender and racial biases held by students, have low and variable response rates, and are easily misused. Indeed, some researchers have recommended abandoning SET data entirely because of these issues.

Despite these problems, SET data continues to be almost universally collected and used, at least in part because it is possible to collect SET data for virtually all courses and teachers, while measuring learning outcomes is difficult. Rather than abandoning SET data because of its documented flaws, or trying to adjust SET scores to account for student biases, we assume that the unstructured, open-ended free text comments raised by students mean *something*, and that topics freely raised by students in feedback on teaching are of interest for educational institutions. In this study, we undertake a combined analysis of numerical and open-ended feedback from SET data. We use a recently developed statistical learning method, structural topic modelling (Roberts *et al.*, 2014), to simultaneously identify topics whose prevalence in free text feedback data is associated with numerical SET scores. We ask the following questions: what themes are identifiable in student comments about teaching, and can we identify associations between the prevalence of those themes and student perceptions of teaching quality, as represented in SET scores? In other words, what topics are students more or less likely to raise when commenting on teaching that they perceive as better or worse?

We analysed three years of SET data from the Department of Biological Sciences at the National University of Singapore (NUS), representing 75 teachers in 355 module-teacher-semester combinations. The data included over 45,000 feedback comments for two common questions on "teaching strengths" and "suggested areas of improvement". We found that free text response rates were associated with numerical teaching effectiveness scores; teachers rated as most effective showed the greatest excess of "strength" comments over "improvement" comments. Further, structural topic modelling identified many topics clearly representing student perceptions of teaching quality, including some topics associated with numerical teaching scores. "Strength" comments which were more prevalent in highly rated teachers included "connecting learning to real-world situations". There was a striking increase in prevalence of the topic we labelled "life-long learning" for the very highest rated teachers. Positive student feedback for highly rated teachers was specific, detailed, and highly related to student learning; "strength" topics more prevalent for less well rated teachers tended to be more vague.

While the current work is restricted to a single academic department, the results suggest that large scale structural topic modelling of student feedback data may provide opportunities to design data-driven professional development programmes. Student evaluations of teaching may help individual teachers understand the most efficient paths for their own development, and the most important professional development needs of departments and institutions.

References

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand., D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science 58*(4), 1064–1082. http://dx.doi.org/10.1111/ajps.12103

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research, 83*(4), 598-642. http://dx.doi.org/10.3102/0034654313496870