# PROJECT-BASED CURRICULUM AS A GUIDE FOR THE RE-SEQUENCING OF DISCIPLINE-SPECIFIC STATISTICS COURSES

Veronica Fruiht
Department of Psychology
Dominican University of California, San Rafael, CA 94930
veronica.fruiht@dominican.edu

*In discipline-specific statistics courses the pedagogical challenges of providing conceptual statistics training are compounded as students lack the background in mathematics necessary for understanding probability and fail to see the applicability of statistics to their field. To address these challenges, a statistics course targeted for health and behavioral science students was redeveloped utilizing a project-based curriculum. Per the American Statistical Association's call for increased integration of technology, real data, and conceptually-focused active learning in introductory statistics, students interacted with real data to learn concepts necessary to answer discipline related questions. Because topics in the course were re-sequenced to be more in line with the process followed by scientists, statistical inference was covered earlier in the curriculum and revisited multiple times in various contexts to reinforce understanding.*

## BACKROUND

In addition to traditional statistics courses taught in mathematics departments, many universities also offer introductory courses in applied statistics to serve the needs of students who will use statistics for research in their disciplines. Most of these introductory statistics courses for non-majors follow a relatively prescribed structure covering the exploration and summarization of data, probability, sampling, and sampling distributions before discussing the basics of hypothesis testing. Students often learn to calculate parameters and statistics by hand using examples that may be relevant to their fields, but spend little time learning how to apply these concepts to real world data or problem solving.

While this technique can be effective for engaging students who come to the course prepared with adequate understanding of mathematics to make sense of these concepts, the pedagogical challenges of conceptual but accessible statistics education are compounded in applied statistics courses targeted at students from specific majors and disciplines. Students from outside traditional mathematics, science, and engineering disciplines are less likely to have a background in advanced mathematics necessary for understanding probability. With recent trends in higher education to cut developmental mathematics courses from the curriculum in favor of supplemental instruction (e.g., California State University, 2017), underprepared students are increasingly likely to lack the algebraic skills necessary to understand the logic behind conceptual formulas and to carry out computational formulas. In addition to the issue of preparedness, students may not see the utility of statistics to their fields or careers and therefore disengage from the material.

To address these and other challenges, a revision of the introductory statistics curriculum guidelines presented by the American Statistical Association (ASA) called for increased integration of technology, real data, and conceptually-focused active learning into high quality introductory statistics courses (Guidelines for Assessment and Instruction in Statistics Education; GAISE, 2016). The report lists six key recommendations for statistics education including emphasizing the role of statistics as a process of decision making and embedded formative and summative assessment to measure student learning. While these suggestions should improve the learning of students in any introductory statistics course, they are of particular importance for applied statistics instruction. The present manuscript describes one curricular model developed for an applied introductory statistics course for students in the health and behavioral sciences to better meet the guidelines set forth by the GAISE report (2016).

## REVERSING THE CURRICULUM

Recent publications discuss strategies for adapting statistics courses to better meet the ASA guidelines. One innovation in statistics education that may address some of these issues is the re-sequencing of topics to cover concepts in a way more in line with the process followed by

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org 

scientists (Malone, Gabrosek, Curtiss & Race, 2010). This re-sequencing allows the topic of statistical inference, an essential learning outcome of the course, to be covered earlier in the curriculum and thus revisited multiple times in various contexts. The opportunity for repetition throughout the semester helps to address the issue of students' struggle to understand the logic of inferential statistics by allowing for multiple iterations of similar content in slightly varied contexts.

Malone and colleagues (2010) also advocate for reducing the burden of aiming for full understanding of probability and instead teaching "just enough" for students to understand *p*-values. Rather than spend many weeks developing an understanding of sampling and probability at the outset, students can be introduced to statistical inference early in the semester, and probability can be addressed in the context throughout the semester. Therefore, students are exposed to a variety of statistical analysis techniques necessary for research in their fields and build on their understanding of concepts in tandem to learning these skills. In order to highlight the process of data analysis used by statisticians and scientists, course content can then be arranged around a semester long project in which students apply their skills to a real world problem.

In my course students in a discipline-specific statistics course targeted at students in the health and behavioral sciences were introduced to a real data set during the second week of instruction. A truncated and cleaned version of the publically available National Longitudinal Survey of Adolescent and Adult Health (Harris & Udry, 1994-2008) data set was provided to students as a tool for learning how to ask and answer questions with statistics. Students were introduced to discipline-standard data analysis software (Statistical Package for the Social Sciences, Version 22) and given the opportunity to explore the variables and response options while learning the differences between nominal, ordinal, interval, and ratio data. During this initial familiarization with the data set, students worked in small groups to develop questions that matched their interests and field specific expertise, ranging from nursing and occupational therapy, to political science and psychology. While students did not yet have the skills to answer their questions, these served as a framework for the skills that they would master over the course of the semester. Students began to understand the goal of statistical inference from their first days in the classroom, discussing how data from a sample might allow them to make predictions about the population the sample was drawn from. Topics were then covered in the following sequence, spending about 2-3 weeks on each topic:

- Central tendency and variability (2 weeks)
- Comparing one to many: Percentiles, z-scores, and one sample t-tests (2 weeks)
- Statistical inference and probability basics with correlation and regression (3 weeks)
- Other inferential parametrics (independent and paired samples t-tests, ANOVA; 3 weeks)
- Non-parametrics (2 weeks)
- Data visualization and presentation (2 weeks)

## PROJECT-BASED LEARNING

Project-based learning, in which students construct their own learning through student-inquiry driven projects, has become a well-respected classroom practice that deepens student learning and real world problem solving (Barron & Darling-Hammond, 2008; Finkelstein, Hanson, Huang, Hirschman, & Huang, 2010). In alignment with the ASA guidelines (GAISE, 2016), project-based learning in introductory statistics would allow students to interact with real data to learn in a hands-on format, answer real world questions using the software and technology of the industry, and see the utility of statistics in understanding their fields (Dierker, Kaparakis, Rose, Selya, & Beveridge, 2012; Wardrop, 1999).

In my course, project-based learning provided a conduit for demonstrating the value of statistics to as a decision-making and problem-solving tool as well as an opportunity to for technology supported learning as students conducted and learned to graphically represent their analyses. Students responded enthusiastically to the opportunity to utilize real world data and expressed excitement about considering the relationships between various health outcomes and larger social contexts. Projects culminated in a presentation in which students described their research questions, analyses, and conclusions as they would for a panel of policy decision makers.

In addition, the use of small workgroups provided students an opportunity to engage in collaborative learning. The Association of American Colleges and Universities have identified collaborative projects as a high impact practice which contribute to student learning and retention (Kuh, 2008).

## FOUR-PRONGED FOMRATIVE AND SUMMATIVE ASSESSMENT

The final recommendation provided by the GAISE report was the utilization of high quality assessment to measure student learning (GAISE, 2016). Given the non-traditional structure of this project-based course, students demonstrated their developing expertise and conceptual understanding of course content through multiple forms of formative and summative assessments.

### *Formative Assessment: The Portfolio*

To track progress through the course, each student created a statistics portfolio in which he or she kept track of each key skill built in the course. Portfolios were structured by the instructor, but each student was responsible for providing a definition, conceptual explanation, conceptual formula, and example of computerized output for each major statistic or parameter covered in the course. This binder was assessed periodically by the instructor to gauge conceptual understanding and skill-building. In addition to its value as an assessment tool, the portfolio was intended to serve as a tool for students to use in their later methodology or capstone courses. To encourage student effort on this assignment, given the low-stakes and formative nature of it, students were allowed to use this resource as a reference during summative examinations.

### *Summative Assessment: Written Exams*

Two traditional summative written exams were used to assess student learning in the course. Students were asked to hand-calculate measures of central tendency and variability and provide written descriptions of the conceptual processes behind inferential statistics. On the final exam students demonstrated their ability to select the correct inferential statistic to answer a research question, and interpret output from statistical software explaining not only the result of the test, but what that result means. Students were asked to present results both in formal APA reporting format, and in an informal narrative as they would explain it to a statistics novice.

### *Formative Assessment: Sample analyses with a shared data set*

In place of traditional homework problem sets, students demonstrated their learning and competence building by conducting analyses by hand with a computer program using the class shared data set. For each new skill or analysis introduced in the classroom, students ran one analysis using the variables the instructor suggested and a second to answer an empirical question for their semester-long project. For example, the entire class conducted a Pearson r correlation to assess the relationship between hours spent watching TV in adolescence and body mass index (BMI) in adulthood. Then, each student conducted and interpreted one additional Pearson r correlation pertaining to his or her area of interest.

### *Summative Assessment: The Presentation*

At the end of the course, students presented the findings of their semester-long research projects to their classmates and instructor. Students used descriptive statistics to describe their samples and demonstrated appropriate use of at least two inferential statistics that had been conducted throughout the semester. In addition to demonstrating their ability to describe a sample set and select, conduct and interpret an appropriate analysis, students were also required to present their findings visually to their classmates as they would to stakeholders and decision makers in the fields of behavioral and physical health policy. Most student projects showed a good deal of applied understanding of how scientists use data to make decisions.

## CONCLUSION

Students from varied mathematics backgrounds and with an array of different learning and professional goals enroll in courses in applied statistics to supplement their training in methodology and practice. A project-based learning approach following the sequence of description and data analysis followed by scientists is one promising approach to address the complex issue of student engagement and preparedness. This approach capitalizes on students' interests and desire to master skills that will allow them to be more competent professionals and problem solvers. Utilizing a real data set and allowing students to apply the skills they learn in the classroom to answer their own questions using statistical software, students learned the theory behind the procedures in parallel to application.

Students demonstrated enthusiasm for learning and were notably engaged with their projects, proudly sharing their findings on a broad range of topics including substance use, sexual literacy, and income inequality. From the perspective of an instructor, I found that backing off from the expectation that students would master summation notation, successfully perform algebraic computations, and fully understand the complexities of sampling and probability gave me the space to support student mastery of the application of statistics to their fields. Similarly, this approach helped provide a more equitable experience for students with varying levels of mathematics expertise and did not put students without a solid foundation in advanced algebra at such a stark disadvantage. The difference in course grades between students who had previously taken calculus and those who had not was somewhat smaller in this course than in the traditionally instructed sections. Just as I had more flexibility to focus my instruction on helping students to build the most critical skills for interpreting statistics in scholarly research, students appeared to have more cognitive resources available to grapple with the idea of using statistics for problem solving and decision making without as much mathematics anxiety. However, future research should aim to use a mixed-methods approach to understanding the utility of this approach in promoting students conceptual understanding, confidence in applying skills, and ability to succeed in subsequent courses that draw upon the skills developed in this course.

REFERENCES

Barron, B., & Darling-Hammond, L. (2008). Teaching for meaningful learning: A review of research on inquiry-based and cooperative learning. In L. Darling-Hammond, Brigid Barron, et al. (2008). *Powerful Learning: What We Know About Teaching for Understanding.* San Francisco: Jossey-Bass. Retrieved on March 14, 2012 from http://www.edutopia.org/pdfs/edutopia-teaching-for-meaningful-learning.pdf

California State University. (2017). *CSU makes sweeping changes to developmental education policy* [Press release]. Retrieved from https://www2.calstate.edu/csu-system/news/Pages/CSU-Makes-Sweeping-Changes-to-Developmental-Education-Policy.aspx

Dierker, L., Kaparakis, E., Rose, J., Selya, A., & Beveridge, D. (2012). Strength in numbers: A multidisciplinary, project-based course in introductory statistics. *The Journal of Effective Teaching, 12*(2), 4-14.

Finkelstein, N., Hanson, T., Huang, C. W., Hirschman, B., & Huang, M. (2010). *Effects of problem-based economics on high school economics instruction.* (NCEE 2010-4110). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ncee.ed.gov

GAISE College Report ASA Revision Committee. (2016). *Guidelines for Assessment and Instruction in Statistics Education College Report 2016*. Retrieved from The American Statistical Association. http://www.amstat.org/education/gaise

Harris, K. M. & Udry, J. R. (1994-2008). National Longitudinal Study of Adolescent Health (Add Health). ICPSR21600-v9. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: Association of American Colleges and Universities.

Malone, C. J., Gabrosek, J., Curtiss, P., & Race, M. (2010). Resequencing topics in an introductory applied statistics course. *The American Statistician, 64*, 52-88.

Wardrop, R. L. (1999). *Student projects in an introductory statistics course* [White paper]. Retrieved Feb 16, 2017, from University of Wisconsin Department of Statistics. www.stat.wisc.edu/~wardrop/papers/tmoore.pdf