EXPERIENCES WITH RESEARCH TEAMS COMPRISED OF GRADUATE STUDENTS, FACULTY RESEARCHERS, AND A STATISTICAL CONSULTING TEAM

Heather S. Smith and John H. Walker
Department of Statistics, California Polytechnic State University, United States of America
hsmith@calpoly.edu

Each year at Cal Poly, statistics faculty provide consulting services to over 100 non-statistician graduate students and research faculty from across the university as part of our Statistical Consulting Service. In addition, all undergraduate/tertiary statistics majors take a capstone course titled 'Statistical Communication and Consulting.' This course is a blend of the theoretical and practical aspects of statistical consulting; helping our statistics majors develop the tools necessary to successfully participate as statistical consultants. Following this training, many of these statistics majors work collaboratively with clients in their research efforts. These research teams are often comprised of a faculty statistician, a trained statistics major, non-statistician graduate students, and research faculty. We report on the key aspects of the statistical consulting course, and provide examples of these research efforts, emphasizing the learning benefits of this research arrangement to both the statistics majors and the graduate students.

BACKGROUND

The California Polytechnic State University (Cal Poly) is located in San Luis Obispo, California, United States of America (USA). It has 18,516 undergraduate students, 788 graduate students, and 1,121 faculty. The academic year is divided into three 10-week quarters, plus two 5-week summer sessions. While education, rather than research, is the school's primary mission, faculty and students do engage in many research projects. Part of the mission of the Cal Poly Statistics Department is to provide statistical consulting to faculty, staff, and students. Since the department does not offer a graduate degree, this consulting is done by our 16 full-time faculty members with the assistance of some of our 60 statistics majors.

Each Cal Poly undergraduate student must complete a senior project in order to graduate. While the requirements of these projects vary from major to major, many include data collection and analysis. These senior projects create an additional need for statistical consulting on campus—over and above that required to support faculty research and graduate student theses. This makes it even more important to train and use the abilities of our undergraduate statistics majors to fill this need.

The GAISE guidelines (American Statistical Association, 2005) list working with real data as an important aspect of undergraduate statistics education. You can't get more "real" than first-hand experience as part of a statistical consulting team. For those students who participate, these consulting opportunities help prepare them to enter master's and doctoral programs and to enter the workforce with the experience of working on an interdisciplinary research team.

Cal Poly's B.S. in Statistics requires 180 units of coursework. These include 90 units of major courses, 16 units of support courses, and 64 units of general education courses. The required statistics courses include 3 quarters of introductory statistics, 1 quarter of applied regression analysis, 1 quarter on design of experiments, 1 quarter of SAS programming, 3 quarters of mathematical statistics, 3-6 quarters of upper-level statistics electives (time series analysis, survival analysis, categorical data analysis, multivariate statistics, survey sampling, advanced design of experiments, and R programming), and a capstone course on statistical consulting.

TRAINING STATISTICS STUDENTS TO SERVE ON RESEARCH TEAMS

In addition to traditional coverage of statistical methodology, our majors get experience in statistical consulting in three ways. All statistics majors take our *Statistical Communication and Consulting* capstone course. In addition, some students participate in senior projects that pair them with a faculty member to participate in a single large interdisciplinary research project. A few students (1 to 2 per year) spend a year assisting the faculty in our department's Statistical Consulting Service. These students participate in consulting sessions on many different research projects for senior project credit. We describe these experiences below.

Cal Poly's Course in Statistical Communication and Consulting

Cal Poly's 4th-year statistics majors take a capstone course entitled, *Statistical Communication and Consulting*. The course is designed to introduce students to statistical consulting while helping them integrate their knowledge from previous statistics courses. Students learn how to conduct initial and follow-up consulting sessions, how to communicate with clients while avoiding statistical jargon, and how to manage difficult consulting situations.

The class meets 4 hours per week for 10 weeks and is built around 4 or 5 statistical consulting projects—each lasting about 2 weeks. For each project, the 12 to 20 students in the class are divided into 3 to 4 person teams. Usually, each team works in parallel on the same project, although occasionally larger projects are divided into several pieces—one for each team. Since this is the first consulting experience for many of our students, the first few projects are selected from previous consulting projects that have already been analyzed with faculty playing the role of the client. This provides a safe environment for the students to learn and practise consulting skills. The later team consulting projects of the quarter are typically done for real clients.

In addition to the project assignments, students read and discuss journal articles on statistical consulting. They also practise statistics skills important in consulting, such as choosing the most appropriate statistical method and computing power and sample size.

For each project, the initial and follow-up consulting sessions for each team are recorded on video. Faculty coach the students on their communication skills as well as their consulting report and statistical methodology. The student teams are assessed on their communication with the client, their consulting reports, and the technical aspects of their analyses. A final individual project and written exam assess each individual student's abilities.

Cal Poly's Statistical Consulting Service

The Cal Poly Statistics Department runs a consulting service to provide free statistical advice to any client within the university. Each quarter, one or two rotating faculty members receive a course reduction to participate in the consulting service. Typical clients are undergraduate students completing senior projects, graduate students completing master's theses, faculty researchers, and university staff. The consulting issues vary widely in their complexity and in the time required to address them. Some questions are simple enough to be answered immediately during the initial consulting appointment. Others require extensive analysis with multiple follow-up sessions that stretch throughout the quarter and sometimes beyond. Because the consulting is free and faculty resources are limited, the consultants require the client to do their own statistical analysis whenever possible. In situations where long-term collaborations are needed and funding is available, the client is advised to place a member of the statistics faculty on their grant for appropriate compensation.

In recent years, several undergraduate statistics majors have participated in the consulting service to earn senior project credit. One or two students per year work with the faculty consultants. Early in the year, the students observe consulting sessions and discuss the consulting process with their faculty supervisors. Later, as students become more comfortable with the process, the faculty consultant may hand-off some analyses to the student. The student performs the analysis with faculty supervision and presents the results to the client with the faculty member present. By the end of the year, the student may meet one-on-one with client depending on the nature of the project.

By participating in the consulting service, the student sees many more types of projects than would be possible in a single course or a typical senior project. The student learns how to identify which statistical method to use in many different situations as well as how to communicate effectively with many different types of clients. The assistance of the student allows the faculty member to concentrate on more complicated projects that require advanced expertise. This division of labor means that the client receives more assistance in a more timely manner than would be possible otherwise.

EXAMPLE RESEARCH EFFORTS

In the last few years our faculty has worked on dozens of research projects with statistics students trained using the methods described earlier. There have been great benefits to all research

team members. The clients with whom we work benefit by having their statistical needs met in a timely manner and by obtaining high quality consulting advice for a fraction of the cost that might otherwise be required. The statistics faculty member benefits by fulfilling his responsibility to service the research needs of the university, by developing professionally through continued consulting, and by learning how best to teach and collaborate with students of different skill sets. Our faculty can also better assess the statistical knowledge and skills that our students obtain in their coursework. This has led to several curricular changes in our department including:

- providing more early and frequent opportunities for statistics students to explain technical statistical ideas to non-statisticians;
- providing opportunities for students to videotape their statistical explanations so that they can review them and learn how to improve them; and
- revising the breadth and depth of the content of several courses, including our regression and design of experiments courses.

The Cal Poly Statistics Department has been able to implement these improvements because of the information obtained through the research team collaborations.

We believe that the greatest benefits of these consulting teams are experienced by the students involved in these projects, both trained statistics majors and non-statistician graduate students. Many of our statistics students choose to attend a master's or doctoral program upon completion of their statistics training at Cal Poly. Other students often begin their working life employed as entry-level junior statisticians. Whichever road they take, they have already had an opportunity to develop their consulting skills while working on real projects, with real clients, meeting real timelines, with real (and messy) data, and writing and communicating study results so that they can be understood. Our graduates report that these experiences are invaluable in their future endeavors.

The non-statistician graduate students with whom we work have an opportunity to learn how teams collaborate to address a research question, how to identify, understand, and implement appropriate statistical methods for use in their projects, and how to write and speak about these research methods. Because these graduate students are keenly interested in completing their projects successfully, we have observed that they are fully committed to working with both faculty and trained statistics majors, utilizing each team member's skills to help them reach their research goals.

Below we describe three examples of these collaborative research efforts. They illustrate the different types of projects that are commonly seen in our consulting environment. They involve clients from the public sector, the private sector, and the university. The selected projects also demonstrate a variety of study types: a longitudinal survey, a multi-phase experiment, and a cross-sectional survey.

Research Effort 1: The Infant Feeding Practices Project

In the *Infant Feeding Practices Project*, we assisted two Cal Poly Food Science and Nutrition researchers with the planning, design, implementation, analysis, and reporting of a longitudinal survey of mothers in the community who had just given birth to a baby. The purpose of the survey was to determine how new mothers feed their babies throughout the first nine-months of life. This project was partially funded by local government agencies. Because of this there were two distinct types of clients whose needs were important: the funding agencies' researchers and Cal Poly nutrition researchers. The project spanned a five-year period, including planning, data collection, analysis, and reporting. Because of the long project timeline, this project required that a series of trained statistics majors be involved. Each year a statistics student worked on the research team, and each summer that student overlapped with the next student to ensure that training and project continuity was achieved. Similarly, several non-statistician graduate students in the Nutrition Department did their master's work based on this project. While the students rotated throughout the life of the project, the government clients, the nutrition faculty, and the faculty statistician worked on the project throughout its duration.

Many project tasks, statistical and otherwise, were undertaken by these students under the guidance of the faculty researchers. Some of these tasks were:

- sample frame development and survey sample design and implementation;
- questionnaire development (five different surveys in total, each in English and Spanish);
- development of procedures for managing the field-work and data quality;
- interviewer training for both in-person interviews and telephone interviews;
- the conduct of interviews with mothers in the hospital as well as telephone interviews;
- interview document management, data coding, data management and quality control;
- data analysis planning, both for interim analyses and final analyses;
- data analysis implementation, requiring sophisticated logistic regression models, using weighted data; and
- development of written reports (for government clients, Cal Poly senior projects, master's theses, and manuscripts for publication).

The scope of this project was immense, and it required that the students learn, not only how to collaborate with many people, but also how to implement scientific research methods, many of which were new to them. As a side note, of the five statistics students who worked on this project, two went straight to work upon graduation and the other three went on to graduate school in statistics.

Research Effort 2: The Gas Purification Project

In the *Gas Purification Project*, we assisted a local manufacturing company in their efforts to improve the quality of their production process for a particular product. The client was a senior R&D engineer who had worked with Cal Poly engineering graduate students in the past. This project was the first time that the client had decided to include a statistician in his research efforts. The team was comprised of the client, a graduate student in engineering, a trained statistics student, and a faculty statistician. The client's research goal was to better understand how to maximize product quality by manipulating several production variables. Thus, methods of experimental design were required. Over six months the two statisticians worked with the two engineers to understand the underlying relationships between the variables of interest. A multi-phase approach to experimental design was utilized, conducting screening experiments first, followed by more extensive follow-up experiments. The team worked together to plan each experiment, the client ran the experimental runs, and the two students analyzed the data, presenting the findings to the client, and proposing the next experimental design. The statistics student was instrumental in helping the engineering graduate student understand how the experimental design and data analysis should be done, how to explain the results in practical terms, and how to use the current findings to plan the next experiment. Upon completion of this research effort, the client had a much better understanding of how to improve his production process and how to utilize experimental design methods to achieve such improvements. The engineering graduate student was able to take his new skill set and continue to assist the client with future design needs. The statistics student completed her senior project based on this research and went on to graduate school in statistics.

There is an interesting side note to this project. The statistics faculty found that there were several aspects of this research project that were intriguing, so much so that we chose to reuse this project as a team consulting project in our statistical consulting course. In the consulting course, the students are asked to design an appropriate experiment and then analyze data simulated from their design in order to provide advice about how best to improve the production process. The quality of advice given by the students has become a benchmark for the statistics department faculty in our ongoing assessment of how well we are teaching experimental design principles. We have made curricular changes in response to these assessment results.

Research Effort 3: The Library Project

In the *Library Project*, we assisted two Cal Poly library researchers with the analysis of a large survey concerning faculty satisfaction with university library services. Prior to coming to our Statistical Consulting Service for help with this project the clients had contracted with a private

company to implement a pre-designed survey. Upon receipt of a data file from this company, the clients approached the statistical consultants for help in analyzing the data and understanding the results. At this point the project was identified by the statistical consulting course instructors as one that could be integrated directly into the course. The clients agreed to allow the students to analyze their data as long as each team's final results were later screened by the faculty. The statistics students in the course worked directly with these clients over several weeks, allowing the students to utilize their newly developing consulting skills. The students analyzed the client's data and wrote them a report. They also commented on their concerns regarding the survey design implemented by the outside company. The project demonstrated how, with carefully selected clients and appropriate oversight, undergraduate statistics students can successfully collaborate on a research project while currently enrolled in a statistical consulting course.

CONCLUSION
         The main idea of this paper is that undergraduate statistics students can successfully participate as consulting statisticians in collaborative research teams. These projects must be carefully selected by a supervising faculty member so that the project requirements do not exceed the students' expertise. The students should also receive some training in statistical consulting through either a consulting course or supervision by a faculty member. This collaboration benefits the entire team. The researcher receives more service than would be possible from a faculty statistician alone. Non-statistics graduate students get more interaction with the statisticians, which leads to better understanding of the statistical methods and results. The statistics student gains valuable experience in statistical consulting. Finally, the faculty statistician helps educate students in a way that is not possible in a classroom setting and also is relieved of some of the consulting workload.

REFERENCES
American Statistical Association (2005). *GAISE (Guidelines for Assessment in Statistics Education) College Report*.