

## A Student's Perspective on the LANGURE Ethical Use of Statistics in Research Module

Michael Crotty  
SAS Institute

### Abstract

Graduate students conducting research at land grant universities benefit from training in research ethics, particularly the ethical use of statistics in their research. The ethical use of statistics includes involving a statistician from the design phase of a study through to the analysis and presentation of results. A research ethics course has been designed for graduate research students to learn about conducting research in an ethical manner. This paper focuses on the efforts behind the statistics module developed for the course.

KEY WORDS: statistical ethics, graduate students, consulting.

### 1. Introduction

The LANGURE (Land Grant University Research Ethics) program is an NSF-funded program (NSF Award #0530217) to develop modules for ethical conduct of research in graduate school. The lead institution for the program is North Carolina State University. Dr. Larry Nelson and I worked together to develop the statistical ethics module for the LANGURE program.

In Sections 2 and 3, I discuss my involvement in the program, both as a student in the pilot teaching of the LANGURE course and as a junior LANGURE fellow helping develop the module for statistical ethics. In Section 4, I discuss part of the module we developed, specifically highlights of the responsibilities that statisticians and non-statistician researchers have in the course of academic research.

### 2. Student Experience

As a graduate student in statistics, I was a member of the first offering of the LANGURE course in the Department of Philosophy and Religion at North Carolina State University. The course was comprised of roughly thirty graduate students spanning the spectrum of the sciences at NCSU; there were students from animal sciences, biology, chemistry, physics, as well as a few students from various engineering departments. I was the sole representative from the statistics department. The course was taught in weekly two hour blocks with Dr. Gary Comstock of the Philosophy Department over the span of three months. Active participation was strongly encouraged as the second hour of each class period was devoted to discussion. These discussions were either in small groups of students on a designated topic, classroom-wide discussions or a combination of the two.

The course went quite well, largely because all the students had chosen to be there. Like myself, many of the students were junior LANGURE fellows, so they were also working to develop ethics modules for their particular disciplines. However, others were there simply because they were interested in research ethics. Still others were there because they had already run into some ethically challenging issues in their time in graduate school.

The first few meetings of the course were devoted to ethical theory. While this was an interesting place to begin, I would not want it to be the primary focus of the course. The theory is not what I think most of the students wanted from the course. Luckily, the remainder of the course involved many specific topics in the more applied area of research ethics. I think the course had a good balance of ethical theory and discussion of applied research ethics areas. The topic of research ethics lends itself so well to discussion that the students were generally quite engaged in the discussions.

The course incorporated many guest speakers in the applied part of the course. Topics covered by guest speakers included statistical ethics, diversity in higher education and promotion, human and animal rights in experiments, and authorship responsibilities. I felt that these topics and speakers were well received by the class and generated good discussions.

The course proved to be a valuable experience to me to draw upon later. In fact, the semester immediately following the one in which I completed the LANGURE course, I was tapped to be a statistical consultant for graduate students in other departments working on their thesis and dissertation research projects. In this capacity, I found it quite comforting to have been entrusted with the sense of ethical responsibility I had to these students. There were a couple times where I had to guide students away from particular inappropriate statistical analyses.

As an after-the-fact statistical consultant, I also saw firsthand the importance of a statistician being involved from the outset of a scientific experiment. One student who had worked for two years and made two trips to a field study in Latin America was analyzing her data only to find that the study had been poorly designed and there was little to no power to detect the effects in which she was interested.

The approach of the course is a big part of the value. That is, the course serves as a way for the academic establishment to welcome graduate students into the community of scholars. In doing so, the academic community entrusts the students with a sense of ethical responsibility and gives them a sense of ownership of the issue of ethical conduct in research settings. The course stresses that ethical practice is not so much a long set of rules to be followed as it is an attitude to act in an ethical manner.

## 2.1 Professional Ethical Guidelines for Statisticians

To the end of looking at ethical guidelines in professional settings, one of the assignments of the LANGURE course was to find and print the ethical guidelines (sometimes called Codes of Ethics) of each student's individual professional society. In comparing all of these, we found that the vast majority of the guidelines enumerated are quite similar in nature. Each profession has a few special cases, and certainly we found that many of the professional societies emphasize different items.

The ethical guidelines in place for the statistics profession include those set forth by the American Statistical Association and by the International Statistical Institute. The ASA's "Ethical Guidelines for Statistical Practice" are available on the ASA website; they were approved in August 1999. As our discussant Dr. John Gardenier pointed out in our session at the 2007 Joint Statistical Meetings, these are still a work in progress and all ASA members are encouraged to voice their suggestions for improvement to these guidelines. The ISI's "Declaration on Professional Ethics" was adopted in August 1985.

Dr. Gardenier also made it quite clear that these are both sets of guidelines, not "Codes" of ethics. They are not exhaustive lists of rules, but merely ideals to follow in the pursuit of ethical statistical conduct in all aspects of statistical practice.

## 2.2 Course Suggestions

After completing the course as a student, there were a couple of improvements that I felt could be made. My main suggestion would be to explore other formats for the course. These formats could be either online (there are modules available at the LANGURE website as well as through the OpenSeminar project) or in smaller less formal groups, such as reading groups. I think the reading group format could have many benefits due to the amount of discussion it is helpful to have when pondering ethical issues. Perhaps the reading group could invite experts to facilitate discussions on some of the more difficult or unknown topics, such as ethical theory. However, the majority of the topics are great for discussion in a small group setting with some outside preparatory reading materials provided by the LANGURE modules or other sources. Ideally, I think this structure would include faculty members to uphold the welcoming to the scientific community aspect as well as to provide interesting experiences to the graduate students just getting exposed to the world of academic research.

It is worthwhile exploring the various formats discussed briefly above as well as various compositions of the course. My experience with a cross-section of graduate students was interesting because we were able to share insights across disciplines. However, being the only statistician in the room for all the discussions, the issues discussed were very different than the issues brought up by the audience at our JSM session. Therefore, it could be of interest to the various disciplines to have departmental reading groups that allow members of the group to drill down more deeply into discipline-specific ethical issues.

It is worth noting here that one of the LANGURE course

meetings was an assignment to attend a seminar on research ethics. The student was free to choose if the seminar was in his/her own discipline or another discipline. This assignment was made possible by the fact that there was a week-long seminar series with a focus on ethics in academic research with many departments participating.

Finally, I would recommend that the course be taken by students early in their graduate studies before they get heavily involved in research. This way, they can be thinking about ethical issues from the start of their research careers. It would also be useful to have some graduate students who are near graduation talk with the students in the course to discuss some of the things that they have seen in graduate school.

## 3. Junior Fellow Experience

Beyond being a student in the first LANGURE course at NCSU, I was also a junior fellow for the LANGURE program. Working with Dr. Nelson, the senior fellow in statistics, we developed a module for statistical ethics for graduate students. This module was a draft that was then evaluated by a larger group of statisticians involved as LANGURE fellows at other participating land grant universities. A conference held in April 2007 facilitated creation of the resulting module which is part of the larger set of LANGURE modules available online.

Despite being given few guidelines as to the format of the module we were writing, Dr. Nelson and I came up with a set of responsibilities and case studies. In the following section, I will discuss the highlights of the responsibilities section of the module. Our primary audience is statistics minors, and as such, we do not go into many statistical details of which test should be used in which situation, for example. We assume the overall responsibility for the teaching of the LANGURE modules will be shared between home research departments and the ethics and/or philosophy department on each campus. A separate module could in the future be developed specifically for statistics majors, but that has not been done at this time.

Our primary message for the module is that statisticians in the research setting are more than merely consultants to be contracted to analyze the data after the experiment or study has been conducted. Statisticians should be involved from the planning stages of the study through the end of the study. In this capacity, we view the statistician as a collaborator in the research process, and we urge the non-statistician researchers (the primary audience of the module) to also view the statistician as a collaborator in the research process.

In addition to the content discussed above, we included specific issues that arise in various aspects of statistical consulting such as medical and business applications. I do not go into details of these sections here, however.

## 4. Highlights of Ethical Responsibilities

In the statistics module, we decomposed most research into four stages: planning, execution, analysis and reporting. We

detailed many responsibilities that exist in all four stages that belong either to the statistician, to the researcher or to both. We stress that these are recommendations for ethical practice. In the following sections, I cite some of the most important responsibilities we identified.

#### 4.1 Statistician's Responsibilities

During (and even before) the planning stage, the statistician should stay continuously current with statistical technology. New methods and techniques, as well as more sophisticated software, are important elements of being prepared to be a successful participant in scientific research. It is also beneficial to the scientific community for the statistician to impart basic statistical concepts to others as a matter of course through interactions with scientists. Here, we note that it is beneficial to the community to raise the statistical knowledge in general.

More specifically in the planning stage, the statistician should be careful to recommend designs that minimize the number of humans or animals that will be subjected to potentially harmful treatments. This can be done using accurate power calculations and honing the research questions such that the experiment can be designed to best answer those questions.

The statistician's primary role in the execution stage is to make sure that the researcher has a good roadmap to follow and be available to answer any questions that the researcher may have.

In the analysis stage, the statistician should assist the researcher in choosing the proper procedure and make sure to keep the data confidential. However, this does not imply that the statistician should keep secret the statistical methods and procedures used in the analysis.

The statistician should also assist the researcher in writing the statistical methods section of the paper. Also, depending on the statistical savvy of the researcher, the statistician may need to review the researcher's interpretations of the results and recommend revisions as necessary.

#### 4.2 Researcher's Responsibilities

From a statistical ethics point of view, the researcher's primary responsibility in the planning stage is simply to contact a statistician. The researcher should seek basic training in statistics, but the relationship between the statistician and researcher should be an interactive one. There are also many ethical protocols that must be followed when dealing with human or animal subjects in an experimental setting.

During the execution stage, the researcher is urged to carry out the experiment or study exactly as planned by the researcher and statistician in the planning stage. While deviations from the plan should be minimized as much as possible, the researcher should also log any important data-related events that occur during the execution of the experiment.

The researcher, during the analysis stage, must preserve the integrity of the entire data set. It is important not to discard parts of the data that do not follow preconceived notions. The researcher should also work in conjunction with the statistician to avoid use of result-guided procedures.

In the reporting stage, the researcher has the bulk of the work to do and should strive to provide good descriptions of what was done, both from a subject matter point of view and a statistical point of view. Again, the researcher should seek help from the statistician on the statistical methods and results sections as needed. We also highlight that the researcher should be sure to include variability estimates such as standard deviation or variance when reporting results.

#### 4.3 Responsibilities of Both

Finally, there are some responsibilities we identified that should be shared by both the statistician and the researcher. I highlight three of these responsibilities in this section.

The first is to honor the rights of fellow researchers and give credit to the previous work of others. This includes acknowledgement of contributions to the current work as well as proper citation of prior work.

Secondly, we urge objectivity to be given a prominent role in research. Preconceived notions of desired results should be avoided.

Thirdly, misinterpretations of data should be guarded against. Researchers and statisticians should be mindful of the limits of their results, and they should avoid "overselling" the implications of the results. To whatever extent possible, they should attempt to deter others from overstating the results as well, but this can be difficult to control.

### 5. Conclusion

As part of the NSF-funded LANGURE program, we developed a module for statistical ethics geared towards graduate students at land grant universities. In this paper, I have discussed my experience as a student in the LANGURE course, my experience as a junior LANGURE fellow, as well as the responsibilities of both the scientific researcher and the collaborating statistician identified in the module.

We note that ethics is a vital part of the statistical analysis process, as well as the more general scientific process. We stress that the statistician should be involved with research projects from start to finish rather than brought in at the final stages of data analysis and reporting. The statistician and the researcher *share* the responsibility for ethical research practices.

Finally, we hope to portray that the LANGURE module helps prepare graduate students to use statistics in an ethical manner in their research and their career.

### REFERENCES

- Declaration of Professional Ethics*, International Statistical Institute (<http://isi.cbs.nl/ethics.htm>).
- Ethical Guidelines for Statistical Practice*, American Statistical Association (<http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics>).
- LANGURE Statistical Module*, North Carolina State University, Raleigh (<http://www.chass.ncsu.edu/langure/modules/statistics.html>).
- OpenSeminar — Research Ethics*, Open Courseware Laboratory (<http://openseminar.org/ethics/>).