

Submission on the Ethical Dimensions of your Statistical Study

Your team should plan to have a meeting of about 1 hour length where all team members participate, to discuss the ethical dimensions of your project. The meeting should take place after your proposal is approved but can be done even near the end of the project. In preparation for the meeting all team members should read the piece below and think about how it relates to your project.

Consider starting your meeting with a period of brainstorming, where participants give their ideas, listen to others' ideas and attempt to build or add on to what they are hearing (to try to get more out of the meeting than just the sum of each individual's thoughts). An important part of brainstorming is that no one can say anything critical about any of the contributions since this will inhibit members from giving their input. Often a bad idea can lead to another bad idea and another bad idea and then suddenly a very good idea. Don't let criticism squelch this possibility.

At least one team member should be the scribe or notetaker at the meeting to ensure things are not lost.

After a period of brainstorming the team can enter evaluation mode where the team identifies the most important issues they have identified. At this point evaluation, criticism and voting may be appropriate. But always be respectful with your team members.

After the meeting you should write up your findings on the ethical dimensions of your statistical study in a 2-page document that you include with your final submission of the project.

Considering the ethical dimensions of a statistical study

In considering the ethical dimensions of the statistical study you are conducting, four areas that are useful to consider are **consent**, **confidentiality**, **avoiding harm** and **transparency**.

If your study involves getting information from human participants, it is important that you explain to them that their consent to participate is voluntary and that they may end their participation in your study at any time. You should be especially careful if you represent any sort of an authority figure with the participants in your study, say for example you are a supervisor in a job and you decide to interview people who work for you. As part of getting informed consent, you should explain how you will use the data you collect and the extent to which the information might be disseminated (e.g., for internal use, for a report visible only to the professor and TA for the course, summarized in a social media post, submitted to an academic journal). Also, you should be clear when obtaining their consent, on whether their identity will be anonymous (i.e., even the people running the study cannot identify them), confidential (you will know their identity but you will not disclose it further), or some identifier like their name will be published. Extra care should be taken if your study involves deception of participants or sensitive personal information. Note that if your project involves human subjects, you must fill out the abbreviated Summary Protocol Form (SPF) and submit it with your proposal. You must then have your form approved before getting information from human participants.

You should have a plan on how you will keep the data you collect confidential and protected from unauthorized access. You should then follow this plan. Think about whether the details you disclose will be enough to identify individuals and if so, take action to mitigate this risk to confidentiality. Because no security methods are perfect you should avoid collecting personal information that you do not need.

Consider the information you collect, and the way you collect: are you avoiding harm to the participants, to others or to the natural environment? Examples can include whether your investigation into traffic patterns makes an intersection less safe, or whether in gathering data about certain squirrels you disturb them and cause them harm. On the other hand, if the data you collect could bring harm to participants if disclosed you should identify that risk and take action to mitigate it.

Transparency is a cornerstone for ethical behavior. When it comes to statistical studies this includes identifying and stating any assumptions you are using, reporting limitations and sources of error in your data or collection methods, identification and mitigation of preferences or bias by the investigators, and disclosure of conflicts of interest. Note that sometimes even identifying your own biases can be difficult and as they may be unconscious bias, i.e., biases that you have that you are not aware of. These biases can manifest themselves in the people you pick to participate in a survey, user groups of the tool you are using to find people, or how you interpret the data/answers you gather.

For example, an online poll would preference age and economic groups (for example a twitter poll would likely target university students but not professionals over 40 nor retirees). Using this medium would likely bias your data.

Another factor is influenced by social pressure. If the subject feels that there may be negative impacts to answering the questions truthfully, the results will suffer from bias. This may also be affecting the subject's free speech and feelings of security (a bit like bullying).

Another issue that comes up in statistical studies is the concept of confounding variables. A confounding variable is something that you are not studying that is likely to impact your study. For example, you are studying the drinking of coffee and its relationship to heart disease. You conclude that there is a link. But it may be that people who drink coffee also smoke a cigarette when drinking their coffee. Smoking is the contributor to heart disease and is the confounding variable; therefore, reporting that drinking coffee is linked to heart disease would mislead the readers of your study. Think carefully about whether your study contains any confounding variables and identify them in your report. Most of what was written earlier in this paragraph could be issues that come up inadvertently through lack of care of the investigators. Aside from these you must eliminate actions like creating questions or data collection methods to obtain a desired result or throwing away data that were collected because they do not conform with the result you were hoping for.

It is often helpful to have people who are not involved in the project or have differing views review your methods and proposed hypothesis to get alternate viewpoints and constructive criticism. This can help mitigate some forms of bias as well as possibly help refine the entire experiment.