

Ethics of Experimentation

Scientists in most fields have rules or guidelines that limit the research they can perform. These principles may be enforced by:

- professional societies and journals;
- government or funding agencies;
- universities;
- individual good sense.

Corporate scientists, and scientists in other jurisdictions, have fewer restrictions than government or academic researchers.

Statisticians need to understand the rules that apply in different situations.

Review Procedures

The National Institute of Health requires that all grantees must have their research approved by an Institutional Review Board (IRB).

Most universities have extended the scope of their IRBs to cover all human and animal research conducted by faculty, staff, and students.

At Duke, there are two IRBs—one for the medical school, and one for all the other departments. The head of the non-medical IRB is Lorna Hicks, and she can be found in the Duke on-line directory.

Several of the statistics faculty are doing experiments to determine ways to improve student learning and retention. It is likely that members of this class will be invited to participate in these studies, and later we shall discuss the approval process for such experiments.

Medical clinical trials are highly regulated. Want to distinguish good therapies from harmful therapies as quickly as possible. Many statisticians are involved in this kind of work, employed by both the pharmaceutical industries and the Food and Drug Administration which regulates new treatments.

In clinical trials, there is also a requirement for an Interim Monitoring Board, which periodically looks at the unblinded data from the trial to see whether there are emerging problems. Examples of problems are:

- Unanticipated differences in efficacy on subgroups; e.g., a drug that helps women but harms men is a pattern that should be noticed early so that the trial can be modified.

- Unanticipated side effects; e.g., a blood pressure drug that causes liver damage.

Important Examples

The establishment of IRBs and other systems for professional ethics was largely a response to three famous experiments in the 1960s and early 1970s that were viewed (perhaps incorrectly) as unethical.

The three experiments were:

- Stanley Milgram's psychology experiments on obedience to authority (sometimes called the Abraham Experiment) in 1960-1963.
- Laud Humphrey's study of the demographics of male homosexuality in The Tearoom Trade, 1965-1968.
- Phil Zimbardo's study of role identification (the "Stanford Prison Experiment") in 1971.

Obedience to Authority

In Milgram's experiment, subjects responded to an ad in the New Haven newspaper which promised a small sum of money in exchange for participation in a psychology experiment.

When the subject arrived, he met an Actor who pretended to be another respondent to the ad, and the Scientist said the experiment needed two people, and through a fake coin toss the Actor was chosen to be the Learner and the subject was chosen to be the Teacher.

The Actor/Learner was given a list of pairs on nonsense syllables and was instructed to spend 30 minutes memorizing the pairs.

The Scientist explained that the purpose of the experiment was to determine whether a small amount of pain, in the form of an electric shock, would encourage better recall on the part of the Learner/Actor. Both the Learner and the Teacher were given a shock on the lowest setting (15 volts) of the machine.

The Teacher/Subject was supposed to read one of the nonsense pairs, and the Learner/Actor was supposed to provide the matching syllable. If the Learner/Actor failed, the Teacher/Subject was supposed to administer a shock of ever increasing intensity. The scale on the machine ran from 15 volts to 450 volts, and at 360 volts the machine dial was marked **DANGER**.

The point of the experiment was to determine how far the subject would go on the apparent orders of the Scientist.

The Learner/Actor went to another room, was attached to the electrodes, and the Teacher/Subject and the Scientist returned to the room with the control panel for the shock generator. The Learner/Actor communicated his answers by intercom.

The Learner/Actor pretended to do a poor job. As the voltage increased, he began to scream, to beg them to stop, and finally shouted that he refused to answer further questions. At 300 volts he pounded the wall frantically, and after 315 volts there was no sound from the room.

Usually, the Teacher/Subject wanted to stop. But the Scientist calmly says (according to a strict script) that the experiment should proceed. Often the subject exhibits great stress, but in many cases follows the instructions of the Scientist.

Results: About 72% of men, and about 63% of women, went to the top of the dial. A much larger percentage went very far on the dial before refusing to proceed. Similar results in replications.

After the experiment, the subjects were debriefed. They were told that their behavior was quite normal, and that they should not take this too seriously. Some of were deeply upset by the realization that they were morally equivalent to Nazis.

The ethical question arose because some of the subjects experienced great distress: None of them had anticipated that their participation would lead them to such painful self-discovery.

The Tearoom Trade

Laud Humphreys was a graduate student at Harvard who wanted to study the demographic characteristics of male homosexuals. At the time, the sociology community generally thought that homosexuals were more likely than the general population to have low intelligence, criminal records, menial jobs, and/or be foreign.

Humphreys went to highway rest stations where gay men would meet. He took down the license plate numbers of the cars that parked there. He became friendly with the community, so that he could be sure that the subjects were gay, and even acted as “Watch Queen” to warn when police entered the parking lot.

Humphreys took his license plate numbers to the Division of Motor Vehicles and paid a small fee to get the registered address (the DMV no longer provides that service).

Humphreys' major professor was performing a public health survey for the U.S. government. With the professor's permission, Humphreys sent the survey form to his list of subjects, so that he could compare their responses to those of a random sample of the community.

Humphreys kept the name list in a locked safe, to which only he and the major professor had the combination. After the experiment was complete, he destroyed it.

Humphreys' experiment found that gays were not different from respondents in any significant way, aside from the fact that slightly fewer were married and they had slightly fewer children. This helped correct misconceptions about homosexuality.

However, the sociology community was disturbed by his methodology.

- He had not obtained informed consent from his subjects for their participation.

- His list of names posed a danger—under unlikely contingencies, it could have led to the subjects being “outed”.
- His role as Watch Queen may have created a social obligation that his actions as a researcher violated.

The Stanford Prison Experiment

In 1971, Zimbardo developed an experiment in which 24 Stanford students were paid \$15/day to play roles as guards and prisoners. They were selected from among 70 potential recruits through a psychological screening system designed to weed out social misfits. From those who qualified, 8 were randomly chosen to be prisoners and 16 were chosen to be guards. Then there was an elaborate simulation of arrest, delousing, imprisonment, and so forth. Details are at www.childsoffortune.com/stanford.prison.htm.

The experiment was supposed to last two weeks. But within a few days, many of the prisoners had broken down and wanted to be released. And many of the guards were humiliating and brutalizing the prisoners.

Zimbardo and his colleagues were also affected. They began thinking like prison wardens, trying to keep the prisoner subjects in detention and tending to side with the guards. They manipulated the visiting hours so as to make the experiment seem less onerous.

Christina Maslach, a recent Stanford Ph.D., joined the team after the experiment was in progress. She was revolted by what she saw, and insisted that they terminate the experiment immediately. The experiment ended after only six days.

Zimbardo wrote: “We had learned through videotapes that the guards were escalating their abuse of prisoners in the middle of the night when they thought no researchers were watching and the experiment was ‘off.’ Their boredom had driven them to ever more pornographic and degrading abuse of the prisoners.”

The Stanford Prison Experiment showed very convincingly (though not statistically) that people adopted social roles according to situational expectations. And that those roles could easily take over, so that normal people behaved badly.

Zimbardo’s early termination obviated much of the ethical criticism that surrounded the Milgram and Humphreys studies. But the issues raised by his experiment are obvious.