Which of These Is True? Validity and Ethics in Scientific Experimentation

by
Fred B. Schnee, Department of Biology, Loras College, Dubuque, IA
Andrea Bixler, Department of Biology, Clarke University, Dubuque, IA

Introduction

Below are descriptions of three different research studies. At least one of them is fictitious. Your job is to determine which one(s) is/are fake. To do so, you will have to use your knowledge of the scientific method and ethical research practices as well as your critical thinking skills. We have listed specific questions for you to consider. Read the research studies below and determine the following:

1. Is the research described scientifically valid?
   a. Identify the goal/question of the study. Is the hypothesis falsifiable?
   b. Identify any alternative hypotheses being studied.
   c. Identify any limitations or flaws in the research design. (For example, is the study repeatable and objective?)
2. Is the described research ethical? Using the Georgetown Mantra of Bioethics (explained below), determine whether the study adheres to the principles of:
   a. Beneficence
   b. Non-maleficence
   c. Autonomy
   d. Justice
3. How could you improve the described study and/or what further research is needed?
4. Do you think the described research really happened?

Was It Ethical?

There are many theories that could be used to determine if research is carried out in an ethical fashion. The Georgetown Mantra of Bioethics (also referred to as Principlism; see https://www.nwabr.org/sites/default/files/Principles.pdf) is often used as a framework for examining the morality of scientific procedures. It arose in response to the horrific ethical violations of the Tuskegee Syphilis Study and includes the following principles, illustrated here with examples from that research:

a. *Beneficence*—do good. For example, in the Tuskegee Syphilis Study, the possibility of better understanding the disease, including progression of symptoms, illustrates beneficence.

b. *Non-maleficence*—do not intentionally create a needless harm or injury. Together, beneficence and non-maleficence help researchers formulate a cost-benefit analysis. In the Tuskegee Syphilis Study, the principle of non-maleficence was blatantly ignored when subjects (poor African American men) were not given appropriate treatment for their disease, even after a cure became available; in fact, local doctors were specifically asked not to treat the men. Furthermore, this allowed the men to pass the disease to their wives (who could

infect their babies during delivery), causing egregious harm even beyond the subjects of the study.

3. *Autonomy*—individuals have a right to freely choose their actions. In determining the morality of a study involving human subjects this often takes the form of informed consent. In the Tuskegee Syphilis Study, subjects agreed to participate but were misinformed about the study. They were told they were being treated for "bad blood" and they were not given the opportunity to withdraw from the study, even when a cure for syphilis became available.
4. *Justice*—fair distribution of goods and services in society. This factor is not going to come into play in evaluating the ethics of most studies, although it should be noted that finite research dollars provided to one type of project or one group of people are not available to fund other projects or groups. In the Tuskegee Syphilis Study, it could be argued that money spent to research the disease in African Americans was unfairly distributed since whites and other races might not also benefit, or not to the same extent. This would miss the point, however, that the study took advantage of African Americans, and that it has caused longstanding distrust of medical professionals within this community.

Using the Georgetown Mantra as a guideline, determine whether the studies described on the following pages were ethical. One limitation of this approach is that the principles above are often in conflict with each other. For example a research project will provide information to help individuals in need but it may come at a cost to research subjects. The question then becomes which principle(s) take(s) precedence. In deciding whether you think the studies below are ethical, it is important that you identify what principles are relevant and justify how you reached your conclusion.

It may also help you to use the Nuremberg Code, developed in 1949 in response to the atrocities committed by Nazi doctors (see http://history.nih.gov/research/downloads/nuremberg.pdf).

The Georgetown Mantra (developed in the 1970s) provides the same overarching guidelines, but the Nuremberg Code is more detailed and provides greater guidance in cases when the four principles conflict.

Research Study I

Chester Southam worked in the Division of Experimental Pathology at the prestigious Sloan-Kettering Institute for Cancer Research in the 1950s. At this time, many scientists did not understand the genetic basis of cancer. It was known that cultured human cancer cells could be injected into animal models ("heterotransplantation") to cause cancer. Southam wondered what would happen if he injected such cells into other people ("homotransplantation"). He hoped to learn more about how the immune system functioned in the presence of cancer; he was especially interested in whether researchers who worked regularly with these cultured cancer cells, or patients given vaccines developed using these cancer cells, might be in danger of developing cancer.

In the first part of his study, Southam solicited 14 volunteers who already had incurable cancer (which was most cancer at the time) and "very short" life expectancies. He told the patients he was giving them an injection to test their immune systems and made certain they were willing to undergo later biopsies for the purposes of the study. He did not tell them they were being injected with cancer cells because he did not want to confuse them about their own diagnosis or frighten them unnecessarily.

Southam, Moore, and Rhoades made "[t]wenty-four homologous implantations of seven cancer cell lines … in 14 cancer patients" (159). "Usually a single preparation was inoculated at one or two sites, but a few recipients received two to four cell types simultaneously, and one received a total of seven preparations on two occasions" (ibid., 158). He also gave them a tiny tattoo at each injection site for later reference. Each injection caused slight reddening and swelling that subsided within 3 days. All 7 different lines of implanted cells resulted in formation of "palpable nodules" (i.e., lumps that could be felt through the skin). Most of these nodules were removed and studied; researchers found actively dividing cancer cells in all of them. If the nodules were not removed, they disappeared spontaneously in 4–6 weeks. However, they grew back in 4 patients, and in 2 patients, persisted until death (6 or 8 weeks after injection). In one of these, the injected cancer metastasized to the patient's axillary (underarm) lymph nodes (because the injected cancer was of a different type than the patient's original cancer, it was possible to identify that in the lymph node as originating from the injection).

No growth was detected from injections of normal cells, even though cancer cells injected into the same patients at the same time did lead to the formation of nodules.

Southam and his colleagues then wondered whether the cancer cells would be able to grow in patients who did not already have cancer. For this study, they performed the same procedures on 14 volunteers from the Ohio State Penitentiary in June 1956.

The results in these subjects differed: the initial reddening and swelling persisted for a week or more. Nodules grew and were removed for study, but cell division and other evidence of cancer were present in only 4 of 15 biopsies. Nodules that were not removed disappeared within 3 weeks and did not recur within 5 months.

The researchers stated in their 1957 paper that they could not yet conclude whether the differences between the cancer patients and normal subjects were due to the earlier presence of cancer itself or to the fact that the cancer patients were already debilitated by their disease. However, the cancer patients did produce antibodies to viruses with which they were injected around the same time, suggesting that the growth of nodules at the injection sites was not simply due to a complete lack of immune system activity.

Research Study II

In the 1970s and 1980s, Elaine Hatfield and her colleagues conducted a number of studies on romantic and sexual behavior of undergraduates at the University of Wisconsin, Madison. Of particular interest was the effect of "equity" on various aspects of the subjects' relationships. An equitable relationship was one in which both partners felt they contributed as much as they derived from it. An individual who received more than s/he contributed was "overbenefited," and one who contributed more than s/he received was "underbenefited." The researchers predicted that if individuals felt that their relationship was equitable, they would feel more contentment and have more satisfying sexual relations. Individuals who were overbenefited or underbenefited would not be as content or find their sex lives as satisfying.

Subjects were drawn from introductory human sexuality courses and filled out anonymous surveys about their relationships. Those who were currently dating in a "casual" or "steady" manner were included in the study. These included 70 men and 119 women in one study, 227 men and 310 women in another.

The equity of the relationship was determined through four questions:

1. How would you describe your contribution to your relationship?
2. How would you describe your partner's contributions to your relationship?
3. How would you describe your outcomes from your relationship?
4. How would you describe your partner's outcomes from your relationship? (Hatfield, Walster, and Traupmann 325).

Subjects answered on an 8-point scale ranging from "extremely positive" to "extremely negative." Mathematical formulas were used to convert these answers to an equity measure of "greatly underbenefited," "slightly underbenefited," "equitably treated," "slightly overbenefited," and "greatly overbenefited" (Hatfield, Walster, and Traupmann 325).

Contentment with the relationship was measured in a similar way, by asking subjects to respond on a 4-point scale how "contented," "happy," "angry," and "guilty" they felt about "what you put into [your relationship] and what you get out of it—and what your partner puts into it, and what s/he gets out of it" (Hatfield, Walster, and Traupmann 326). Positive responses (happy and contented) were summed; negative responses were subtracted from positive to give one overall score. The same was done with satisfaction/happiness with subjects' "life in general" (Traupmann, Hatfield, and Wexler 39).

Researchers measured sexual satisfaction in two ways. The first was overall sexual satisfaction. "How satisfied are you with your sexual relationship with your partner? (Possible answers ranged from: 8=extremely satisfied, to 1=extremely dissatisfied.)" (Traupmann, Hatfield, and Wexler 39–40). The second referred to how the subject felt immediately after sex and involved two questions, the scores for which were summed. "After sex with my partner, I usually feel … (possible answers ranged from 8= extremely loving and close, to 1=extremely distant and angry). After sex with my partner, I usually feel … (this time, possible answers ranged from 8= extremely sexually satisfied, to 1= extremely sexually frustrated)" (Traupmann, Hatfield, and Wexler 40).

Results strongly supported the hypothesis that those who view themselves as slightly under/overbenefited or in equitable relationships were more content than those who saw themselves as greatly under/overbenefited. Comparing just those in equitable relationships to the greatly under/overbenefited, the difference in contentment was significant with $p<0.001$.

The results of the surveys on overall sexual satisfaction versus equity were not significant. However, subjects did report more positive feelings immediately after sex when they were in equitable relationships than when they were under/overbenefited.

The researchers concluded that "equity considerations might be an important determinant" of sexual satisfaction (Traupmann, Hatfield, and Wexler 38). They pointed out that "our data are correlational, so it is possible that some unknown variable X might be causing college men and women to (1) rate themselves as Overbenefited, Equitably Treated or Underbenefited, and (2) make them more or less enthusiastic about … sex. … . In addition, the causal sequence might be the opposite to that we suggest. "Men and women's sexual experiences might determine their perception of Equity/Inequity, rather than the other way around" (Hatfield, Walster, and Traupmann 333).

Research Study III

In 1995, Rhonda Daley and Robert Martin, a husband and wife research team at the Institute for Cancer and the Environment, wondered if women who wore underwire bras all day, every day would have a higher risk of developing breast cancer than those who did not. The researchers theorized that wearing underwire bras would inhibit lymphatic drainage, trapping toxins produced by the wire in the breast tissue, causing cancer. To test their idea they surveyed the bra-wearing behaviors of 4,700 US women in 5 major cities (2,056 women in the cancer group and 2,644 women for the standard group). Individuals who used softcup bras part of the time or exclusively were not included in the study. Women in the control group were asked how often they wore their underwire bras; women who had had breast cancer were asked about their bra-wearing habits prior to their diagnosis of cancer. The women were then classified into one of four groups (women who wore underwire bras 24 hours a day, women who wore underwire bras more than 12 hour per day but not to bed, women who wore their underwire bras fewer than 12 hours per day or women who wore bras rarely or never).

All participants signed informed consent forms but were not told the true nature of the study. In order not to bias the findings, participants were told that the research was conducted to determine bra-wearing habits of the general population. The protocol and consent form were reviewed and approved by the participating institutional review boards.

Daley and Martin found that the odds of getting breast cancer dramatically increased with underwire bra-wearing for more than 12 hours per day. Women who wore their bras 24 hours per day had a 3 out of 4 chance of developing breast cancer, women who wore bras more than 12 hour per day but not to bed had a 1 out of 7 risk, women who wore their bras fewer than 12 hours per day had a 1 out of 152 risk, while women who wore bras rarely or never had a 1 out of 168 chance of getting breast cancer. The overall difference between 24 hour bra-wearing and not wearing bras was a 125-fold difference ($p < 0.001$).

The researchers recognized that while their study did not control for known risk factors for breast cancer such as the BRCA1 gene and environmental risk factors that affect endogenous hormone levels, they felt their results were compelling and warranted further research. The authors also argued that many of the known risk factors can be related to bra-wearing behavior and/or the lymphatic system. "For example, breast feeding and pregnancy cause full development of the mammary lymphatics. Also, women of higher economic status have higher breast cancer rates, and one would expect that they would wear underwire bras more hours per day. Women who exercise have lower risk," and are more likely to wear softcup bras when exercising, thus improving lymphatic circulation and reducing toxic exposure (Daley and Martin 468).