

Chapter 2

The Cutting Edge

Medical science continues to advance, and all the new things—new procedures, new imaging machines, new drugs and theories—gives rise to anxious questions. Have we, this time, gone too far? The human being does not accept significant change easily, and medical science has always had the power to set the nerves on edge. At one time organ transplantation was the stuff of science fiction; now it is almost commonplace. Meanwhile, new knowledge can only come from research, experiments on humans like us or on animals like enough to us (or our pets) to worry us when they suffer. Research with human beings became much too urgent in the 1950s and 1960s; eagerness to get useful results quickly spawned abuses. These abuses led to the formation of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, mentioned above, whose reflections yielded some of the best public moral thinking this nation has allowed itself to enjoy, concluding with the articulation of the Belmont Principles, now central to ethical reflection not only in bioethics but in all fields of applied ethics.

2.1 Organ Transplantation

The God Committee (1962) and the kidney transplants: Kidney disease is serious and becoming more so: One in six Americans have some form of chronic kidney disease. 400,000 patients (according to eMedicineHealth) are on dialysis or have received kidney transplants; about 67,000 people die each year from kidney failure. Almost 40 % of adults over 60 years of age have some level of kidney failure. Causes of kidney failure, absent some exotic disease, include diabetes mellitus, high blood pressure, and obesity; all of these are on the increase, and the population is aging; no wonder kidney disease is up 16 % in the last decade.

For most of the human experience, there was little that could be done for a person with failing kidneys. Then in the 1940s, a machine that could accomplish hemodialysis, filtering toxins from the blood the way the kidneys do, was developed, but could only be used for a short time with any patient; it required careful reattachment to new veins and arteries with every use, and the patient's blood vessels quickly collapsed. In 1960 Belding Scribner, of the Swedish Hospital in Seattle, developed a "permanent indwelling shunt,"

a tube permanently attached to one artery and one vein, so that blood could flow continuously through the machine; the tube had a valve that could be closed between uses. Now hemodialysis could keep a patient alive indefinitely.

But very expensively. At the time Scribner developed his shunt, dialysis cost about \$20,000 a year, and insurance companies would not pay it; the number of patients who could benefit from it was effectively infinite, and they would die if they did not get it. When Scribner moved his dialysis center to an outpatient setting, the cost went down a little, but not enough; and only 17 patients could be served. Who should receive treatment when not all can? Scribner, Swedish Hospital and the Kings' County Medical Board established a committee, the Admissions and Policy Committee, to make the decisions. (By this time insurance covered the patients' treatments.) The Committee set criteria for admission to treatment: the patient had to be from the state of Washington, had to be under 45 years of age, had to have insurance or otherwise be able to pay. Still they had many too many candidates. They drafted new criteria: family situation, employment, value to the community, ultimately character and personality. They did their work objectively and well; they were never accused of bias or corruption.

But still: how can one person "deserve" life more than another? The issue came to a head when dialysis and the Committee hit the front page of *The New York Times* in May, 1962. Scribner had taken a dialysis patient to a newspaper convention in Atlantic City to lobby for more media attention to kidney disease, and the Committee had been mentioned; the *Times* picked it up. Shana Alexander, a reporter for *Life* magazine, did a feature story on the Committee, which she called the "God committee," since it played a "godlike" role in deciding who would live and who would die at least in part on grounds of "social worth," the worth of a person to society. The article appeared in November, 1962; in spring 1963, the *Seattle Times* ran a front page story on several patients who had not been accepted for dialysis—"Will These People Have to Die?" Temporary financing for more machines was found from Seattle industry. In 1965, a television special on NBC featured the God committee, and raised popular sentiment for national funding for dialysis, along the lines of future programs like Medicare and Medicaid. By 1971, everyone in the country knew of the plight of rejected kidney patients, and after some more media dramatics, Congress enacted the End State Renal Disease Act (ESDRA), mandating the federal government to pay for dialysis for anyone who needed it. That ended the need for the God Committee.

ESDRA solved some problems, but led to others. By 2006, despite the efforts of several cost-containment mechanisms, dialysis was costing Medicare \$16 billion a year. In a country with 45 million citizens unable to obtain medical insurance for even basic medical care, that seems disproportionate. And of course, sufferers from other diseases wanted to know why their disease could not also have mandatory federal funding; no good reasons have been offered. The use of media and of public awareness on behalf of health care decisions also pioneered in this case. Scribner's trip to Atlantic City was the first time the medical profession had broken out of confidentiality to enlist the public for help; inevitably, it launched the first public discussion of the terms and conditions of health care, the ethics of medical choices, and the insertion of politics into health care decisions. This last development has been a very mixed blessing (see *Terri Schiavo*, above).

Dialysis is not an ideal solution for sufferers from kidney disease; it is time-consuming, exhausting, occasionally dangerous (it is very hard to exclude completely disease-causing organisms), and a drain on the insurance. Far better would be a kidney transplant. But then, of the available recipients, who should get the kidney transplant?

The first recorded kidney transplant happened in 1945 in Peter Bent Brigham Hospital in Boston, when a kidney from a recently deceased patient was transplanted into a woman in kidney failure following the birth of her baby. It functioned badly, did little good but apparently little harm, and the woman's own kidneys recovered. In 1952, another transplant, from deceased mother to her son, functioned for 22 days until it was rejected. The first successful kidney transplant was accomplished in 1954, also at Peter Bent Brigham, between identical twins, Ronald and Richard Herrick.

Why did that one work so well? As most of us now know, the reason why we can't treat our organs as interchangeable parts, freely exchanged according to need, is that each of has a built in physiological hostility to anything recognized as strange or alien, a hostility that forms the essential center of our immune system. Without it, we're dead at an early age, invaded by any and all passing organisms. With it, we reject transplanted organs, unless we figure out some way to turn it off. Steroids will turn it off, or down, temporarily; but steroids have long-term side effects that we don't want. Various immunosuppressive drugs were used in the early days of transplants, 1950s and 1960s, but none worked very well until the discovery of cyclosporin(e) (spellings differ) in 1971. Since then the success rate of organ transplantation has soared. Identical twins, of course, have identical immune systems, and do not recognize each other's organs as alien. That's why the Herrick transplant worked in 1954.

Kidney transplants have one other advantage over other organ transplants: the kidneys are paired organs where each member of the pair can perform the whole task of filtering the blood by itself—as a matter of fact, if you remove one kidney, the other kidney will promptly expand to fill its space, and the health of the donor will not be affected. So “live-donor” organs are preferred for kidney transplants, the closer the genetic relationship to the recipient, the better. With the donor alive, the question of deterioration of the organ does not arise, as it does when the donor is deceased. With advances in immunosuppressive drugs, “cadaver-donor” organs can achieve a decent success rate, but the drugs have a universal disadvantage—they suppress the immune system, increasing the probability of lethal infection and cancer.

Live-donor sources are possible for kidneys, and the practice is extending; at the turn of the millennium, it had become standard practice to look for live donors among the relatives of those suffering from a variety of liver failures (one lobe of a healthy liver can often take the place of a failed liver, especially for children), and lung failures, like cystic fibrosis. The lungs are not exactly paired organs, as not being separated in nature, but a lobe of a lung can often be attached in a sick person's body, and take over the role of the original pair of lungs. The heart remains the most important organ that can only be obtained from cadavers.

The first heart transplant, mentioned briefly above, took place in December of 1967, before immunosuppressive drugs were fully developed. Christiaan Barnard found himself with a long-term patient, Louis Washkansky, who was clearly dying from heart failure. Talk of heart transplants was widespread in the profession, and many of the necessary techniques were known; Norman Shumway, cardiac

surgeon at Stanford University, had announced a month before that he was ready to do a heart transplant when a suitable donor could be found. Then the body of Denise Darvall was brought into Barnard's hospital from an automobile accident; with her father's consent, he removed her heart, placed it in Washkansky's chest, and spent the next 17 days trying to keep him alive, an attempt that ultimately failed. But there seemed to be hope that later transplants might work; after the first week's struggles with multi-organ failure, steroids had bought Washkansky five good days, during which the heart beat well and there was real hope of recovery. (Rejection set in from that point.) Barnard tried again, with Philip Blaiberg, who may have been in better health than Washkansky at the time of the surgery. Blaiberg walked out of the hospital on his own, living nineteen months before dying from complications of heart disease. The world was elated at the possibilities, and surgeons in many nations tried transplanting hearts, livers, lungs. Most patients died before long, the transplanted organ rejected; long-term success awaited the development of much better drugs.

(In an attempt to get around the scarcity of cadaver hearts for heart patients, physicians and engineers collaborated through the 1970s to develop a "totally implantable artificial heart," a small but efficient pump that would simply take the place of the heart of flesh. Never mind the Biblical echoes, please. By 1982 Willem Kolff and Robert Jarvik of the University of Utah thought they had a model ready to go, and William DeVries, a cardiac surgeon at the medical center, implanted it in the chest of cardiac patient Barney Clark on December 1. It didn't work very well; Clark died in quite a bit of discomfort. Later recipients fared worse. One of them, William Schroeder, lived 21 months, most of them miserable, and talked to President Ronald Reagan on the telephone. He died of a series of strokes. What the physicians had not known, or had forgotten, is that the heart is much more than a pump; it is a highly sensitive chemical control mechanism, a source of hormones, responding in multiple unknown ways to stimuli from the brain, and communicating complex system adjustments to the rest of the body. Engineers will design a computer capable of taking the place of the brains of experimental surgeons before they find a machine capable of replacing the heart. Latest news is that the Left Ventricle Assist Device (LVAD), supposed to be a partial temporary artificial heart, isn't very good at keeping people alive, either; research continues and improvements are expected.)

For a recipient of a heart to live, the donor has to die. Not just any cadaver will do for an organ transplant. First, consent must be obtained. If the deceased has indicated, in a living will or other document, a willingness to be an organ donor, that solves most problems (but not all; the family can still protest). If there is no document, and only 20 % of adults have signed such, the family can be asked to allow harvest of viable organs. The usual circumstances surrounding the death of a potential donor (sudden, unexpected, violent death of a young and healthy person) make this request very difficult. Second, the organs must be healthy and strong enough to be useful to the recipient. That rules out the organs of most victims of disease, including all victims of cancer, for the disease or the cancer may well have spread to the organs. It also rules out the most willing of organ

donors, the very elderly; no one wants an 86-year old heart. In practice, the eligible donors turn out to be young (under-40) victims of traffic accidents or gunshot wounds to the head. These are difficult to come by; gunshot victims do not end up in hospitals in time, and the victims of traffic accidents have steadily decreased in number, thanks to seat belts, air bags, tougher laws against drunk driving, and required helmets for motorcyclists. (Citing respect for liberty and self-determination, states occasionally repeal the helmet laws, with the verifiable effect of making many more organs available for transplant. The ethical trade-off here has not been adequately explored.) The problem created by these restrictions is that the next of kin, usually parents, have to be asked to allow organ donation as they stand by the body of their dead son or daughter, a young future brutally cut short, in the throes of shock and grief. It is a scene that no one enjoys, and many cannot bear.

How can we make more organs available for transplant? One attempt, in use in some states, requires any hospital with a possible organ donor—the brain-dead body of a young accident victim—to notify a state organ procurement team, made up of carefully trained professionals, who will arrive and conduct the interview with the next of kin. Life support for the patient could not be removed until the team arrived. The rationale for such teams is that the staff at the hospital may be unwilling to ask for organs due to the sensitivity of the situation, and the organs will be lost as the potential donor dies. The availability of a trained team may increase the quantity of organs donated. Another possibility, in practice in France and other European nations, is “presumed consent”: when a viable donor comes along, whether or not next of kin are available to consult, it is presumed that the deceased wanted to donate organs, which are harvested immediately. That practice does indeed raise the number of organs available; it is not clear that such a practice would work in the United States, where trust in the medical profession is much lower.

Once the organs have been obtained, how shall they be distributed? The usual method of “distributing” medical resources is no method at all: physicians simply assign the medical resources under their control to the patients in their care, and if two patients need the same resource, the first to need it is the first served. (And the patients who have no insurance, or no access to medical care, are not served at all.) Somehow that chaotic method did not seem adequate to the problems of organ distribution: for an organ transplant to work, the donated organ must be placed in the recipient’s body within a very short time after removal, well oxygenated, under near ideal operating conditions; and the recipient must have been carefully screened for medical fitness, that is ability to tolerate and support the donation.

Further, donor and recipient must be histocompatible. “Histocompatibility,” as described on the website of the United Network for Organ Sharing (UNOS), the mechanism created to solve the distribution problem, is the condition in which a donor and recipient share antigens so that a graft (donated tissue) is accepted and remains functional. Histocompatibility antigens, called HLA (for *human leukocyte antigens*), are proteins on the surface of the cells in the body. Their main function is to help the immune system defend against invaders such as bacteria, viruses, and parasites. Unfortunately, the immune system can also recognize as foreign

the histocompatibility antigens of other people's cells and will fight them, causing rejection of grafts and donated organs. Because HLA antigens can be recognized as foreign by another person's immune system, transplant professionals try to match as many of the HLA antigens as possible, between the donated organ and the recipient. That way, there is less of a chance that the recipient's body will reject the organ. In order to do this, the HLA type of every potential organ recipient is determined before they are placed on the waiting list. When a potential organ donor becomes available, the donor's HLA type is determined as well. A match program is run through UNOS and the best possible recipient for each organ is chosen. Further tests, known as crossmatches, are performed to make absolutely sure that the donor organ is suitable for the recipient. Above all, for the transplant to work, someone in the system has to know where the possible donors and recipients are—they could be anywhere, and are very unlikely to be in the practice of a single physician. The possibilities for failure are legion, and the transplant is of infinitely high value—it means life for the recipient. We cannot leave this practice to chance.

So a national system was necessary just to sort out the logistics of organ transplants. That system already centralizes the decision-making, and makes it visible, but there still were no criteria to decide who gets the very scarce and very valuable organ. The God Committee was one answer: appoint a group of prominent citizens, trusted by all, to decide who is most deserving of the pool of medically appropriate recipients. But the God Committee, and the "social worth" criteria they had incorporated into their thinking, had come under vigorous attack from the bioethicists: Sanders and Dukeminier argued in a 1968 *UCLA Law Review* article that social worth criteria should never have been used, saying that the *Life* article "paints a disturbing picture of the bourgeoisie sparing the bourgeoisie, of the Seattle committee measuring persons in accordance with its own middle class suburban value system: Scouts, Sunday School, Red Cross." Nonconformists seem to have no chance for treatment, as they see it; the article concludes with the famous tag, "the Pacific Northwest is no place for a Henry David Thoreau with bad kidneys." Following on that critique, medical sociologists Renee Fox and Judith Swazey pointed out in 1974 (*The Courage to Fail*) that the selection criteria mirrored "the middle-class American value system," where the preferred candidate was characterized by "decency and responsibility," with a strong family life, who had "demonstrated achievement through hard work and success at [his or her] job, who went to church, joined groups, and was actively involved in community affairs." The fact that these accounts were judged to be warranted, and devastating, criticisms of the Committee throws interesting light on the ethical tendencies of bioethics in the late 1960s and 1970s. (What's wrong with decency, responsibility, hard work, church attendance, and active citizenship? If we have to select on some criteria, we could do a lot worse than these.)

Justified or not, these criticisms ensured that no more God Committees would show up on the American scene. Some bioethicists had suggested random distribution: that after recipients had been matched to the available organs medically and genetically, the system should simply draw straws to see who should get the organ.

That didn't seem like a good idea, either. Congress had had a few efforts at governing organ transplantation. In 1987 these attempts, the National Transplantation Act and the Task Force on Organ Transplantation, were combined to create UNOS, which at least did away with regional competition (and hoarding of organs to reserve them for local patients), and adopted some generally fair criteria for allotting them. It created waiting lists: each medical center that was prepared to transplant organs sent a list of available recipients. When a donor became available, the organs were distributed to the recipients according to how long they had been on the list (first listed first served) and by seriousness of illness (the sickest get to the front of the line). That's probably as fair as it's going to get. There are a multitude of difficulties: a person too poor to have insurance or access to medical care will not get on a list at all; a person wealthy enough to devote full time to acquiring a needed organ can go to every center in the country and get on their list, multiplying his chances; there are centers abroad that can supply organs to the very wealthy; an influential surgeon can often put one of his patients in front of others in the line; and there is the unpredictable legacy of the "publicity" factor.

Recall, the whole problem of allotting scarce medical resources first came to public attention when Scribner brought a patient to a newspaper convention. Since then, UNOS or not, every time a patient, or his family or his physician, discovers that he needs a transplant, the newspaper is the first to know, then the talk shows, then local fundraisers to cover medical expenses. The hope is not to persuade the managers of UNOS to put the endearing child or brave mother at the front of the lines they control. The hope is that the next-of-kin of a newly deceased potential donor, seeing the appeals on the local news, will insist that the required organ be donated only to that recipient. Since the implication is that if the request is refused, the organ will not be donated, it's generally granted, restricted only, as above, by histocompatibility.

No one has figured out a way to get around the inherent unfairness of the publicity factor—nor the similar distortion of the public figure. Consider the liver transplant: it is the most difficult of the transplant surgeries, requiring highly skilled teams to make it work at all. Considerations of allotment of livers, among the scarce organs, are complicated by the fact that the most common reason why livers fail is alcohol abuse. Alcohol-related end-stage liver disease (its own diagnosis: ARESLD) is the condition most likely to raise the need for a transplant, but it raises a serious question of justice alongside: this patient did this to himself. Does he deserve an organ transplant after destroying his own liver? While this debate continued, beloved retired Yankees' baseball star Mickey Mantle showed up at his hometown hospital with ARESLD and liver cancer. He was put on the UNOS list, and it seemed a mere matter of days before he got a donated liver. Was that fair? UNOS insisted its guidelines had been followed, but its guidelines insist that cancer is a disqualifying factor; in fact, after receiving the transplant, Mantle died in a matter of weeks from cancer.

When we take up the larger problems of justice in health care, we will have to confront the general problem of the responsibility of the individual to care for his own health and the appropriate response of the health care system to evidence that

he has not. Should smokers be charged more for surgery—or refused treatment altogether? At least, should smokers and heavy drinkers be declared ineligible for heart transplants? This question may not be resolved in our generation.

2.2 Problems in Research

Much medical research is done on animals. Do we have any right to cage and abuse animals in this way? Does it matter if the animal suffers extensively, or just participates in the experiment and is then killed? One of the first targets of the reformers was the Draize test:

The Draize Test: Humans like to rub, or spray, cosmetics on their faces and hair, and some of the cosmetic substance can get into the eyes. For years, the manufacturers tested their new cosmetics by introducing them into the eyes of rabbits, after clipping the rabbits eyes open to prevent self-defense and clamping their heads into stalls so that they could not move. If the rabbits' eyes did not seem to be badly affected, it was concluded that the substance was safe enough for human use around the eyes. The test was terribly painful for the rabbit, and eventually blinded them. Was this necessary? Much fruitful work is being done to develop cloned tissue models to replace the rabbits.

Then the reformers pointed out that beyond the individual procedures, the whole method of counting success was unnecessarily cruel:

LD-50: How toxic is a substance—a new agricultural product, a new floor cleaner, especially new pharmaceuticals? A standard way of discovering the toxicity of something, is to feed it to (or spray it on) a large population of test animals, mice or rats or rabbits or beagles. The question asked in this test is, what quantity of the substance is sufficiently toxic to bring about the deaths of 50 % of the test animals? That's quite a bit of animal death, and the ones that do not die, of course, are very sick and miserable. Is there any other standard that can be used? By reformulating their counts, it seems to be possible to use an LD-10 standard instead, which will not only kill off fewer of the test animals, but will leave the remainder in much better shape.

As experimentation with animals began to accumulate publicity, a new generation of activists closed in on the research to gather the materials for public information:

Edward Taub's "somato-sensory deafferentation" studies in monkeys: Alex Pacheco, an activist for PETA (People for the Ethical Treatment of Animals) volunteered in Edward Taub's neurology laboratory in 1981, pretending to be a neurology student, but really in order to videotape the activities of the lab. Taub had been cutting nerves to the limbs of his subject monkeys to simulate the injuries suffered in brain injury and stroke. He hypothesized that one reason that people had such trouble regaining the use of their limbs after a stroke or brain injury was that since the affected limb did not respond in any way they were used to, they essentially gave up on it, and moved all activity to the uninjured limb, a situation he called "learned helplessness." He thought that if the good limb were restrained—tied down—during the rehabilitation period, the brain might rewire itself to recover the use of the affected limb. The experiments cannot have been pleasant to watch, and Pacheco got some damning film to show the world. To the whirl of TV cameras, police raided Taub's laboratory, while PETA handed out literature. Taub was briefly convicted of failing to give the animals proper veterinary care; the conviction was reversed; he went on

to become one of the most admired neurologists in the field, as it turned out that he had indeed brought about more recovery of function in the monkeys than any had thought possible. His Constraint-Induced Movement Therapy (CI) is now widely used, not only for stroke victims, but for military personnel returning from Iraq with traumatic brain injuries, and is widely praised for obtaining marked increase in their ability to use paralyzed limbs.

2.3 Philosophical Issues of Animal Research

What right do we have to do research with animals anyway? How are animals like ourselves? How different? What is important—their reason, their communication, or their ability to suffer? When we do research with humans, we have to get consent from the subject; how do we get consent from an animal?

There are two major lines of thought to protect the animals: Utilitarians concentrate on the balance of pleasure of pain, claiming that the central issue is the suffering of the animals. This line is argued by Jeremy Bentham and Peter Singer, who are focused on what should be done for the animals' welfare. The other line, led by Tom Regan and others, holds that all creatures who are subjects of a life—a life that can go better or worse, whether or not they are aware of it—have a right to be left alone that that life may flourish; this is an animal rights argument.

Is there some obligation to let animals live in their own place, their own natural habitat, according to their own laws? If there is such a general duty, does it follow that there is an obligation to release to the wild all restrained or tamed animals? Would that be good for them? If not, are we obligated to keep the animals we have? Or may we humanely dispose of them? Meanwhile, a new moral issue is raised by the tactics of Animal Rights activists. Is there a right to destroy laboratories, burn data, when animals are being made to suffer? Who authorized this? Whatever our sympathies, must we not hold animal rights activists responsible for the harm that they do?

As a practical matter, if we do not decide to abolish research with animals altogether, can we adopt guidelines to minimize the harm done? If we decide that the enormous value of using animals for research—there is very little we know about medicine, surgery, or pharmaceuticals, which we did not discover through tests or practice on animals—outweighs the claims animals have to be left alone, we will still have the obligation to minimize the suffering of the animals in our labs. The morally responsible program seems to involve three imperatives:

1. **Replace:** to the extent possible, replace the use of live animals with lab-cultured human tissue, grown from surgical detritus available from hospitals.
2. **Reduce:** to the extent possible, reduce the total number of animals used in the tests.
3. **Refine:** modify investigative procedures to cause less pain and suffering, for a shorter time, leaving the animals healthier; or euthanize them earlier.

2.4 Research with Human Subjects

If we must conduct research with human subjects, how can we do it ethically? Do the subjects have rights that should be protected? The question had been a staple of science fiction for a century and more, but did not arise in any serious way until the end of World War II, when the documents from the concentration camps showed us that many Nazi physicians had used concentration camp populations for experiments that might have been of use to the German military—on, for instance, the best way to revive flyers who had ditched in the waters of the North Atlantic, or the human tolerance for very low atmospheric pressure. Most of the unwilling subjects died, or were severely injured. The cruelty of the experiments shocked the world, comprising a good part of the case of “crimes against humanity” in the Nuremberg Trials that followed the war. The international community adopted a set of “Nuremberg Principles” to govern human subjects research, requiring the voluntary consent of each subject, complete information of risks given to the subjects, overall, a careful balancing of risks and benefits (not so easy, since the subjects took the risks and the patient population as a whole reaped the benefits), and the right to withdraw from the experiment at any time. It further specified that the question the research was exploring had to be worthwhile, and that the investigators had to be qualified to carry on the research and understand the results. The Nuremberg Code became the ethics guide for human subjects research everywhere—not always observed.

One of the first cases of human subjects research in the United States that came to our attention was the Tuskegee Syphilis study. It’s important enough to include in detail.

The Tuskegee Syphilis Study: In the late 1920s, the U. S. Public Health Service undertook a study of the occurrence of the venereal disease syphilis among the male African American employees of the Delta Pine and Land Company of Mississippi, and found an appalling prevalence of 25% of 2000 men tested. The PHS brought in the Julius Rosenwald Fund, a health-oriented charity based in Chicago, to treat these employees, and later expanded the treatment protocol to include five more counties in the South. In 1928 the PHS and Rosenwald decided to collaborate on a study of the course of syphilis in men and the effectiveness of the existing treatments.

The “treatment” for syphilis from 1900 to 1940 was an amalgam of heavy metals—mercury, arsenic and lead—called “salvarsan.” When it became clear that salvarsan was killing people, it was replaced by “neosalvarsan,” a synthetic drug known as an “organoarsenic” compound. It became available in 1912, was equally good at suppressing the symptoms of syphilis, and superseded the more toxic and less water-soluble salvarsan as an effective treatment for syphilis. It too had bad side-effects; incidentally, the side-effects of mercury poisoning are the same as the medical description of the last stages of syphilis. Even as the protocol started, no one really knew whether the drugs being given for syphilis were doing more harm than good.

Meanwhile, some intriguing work by Norwegians Boeck and Bruusgaard between 1890 and 1930 suggested that maybe syphilis should be untreated altogether—that those who had had syphilis for 30 or 40 years seemed to be symptom

free, and of those who were clearly positive for syphilis, many never developed the symptoms at all. So as the protocol started, no one really knew if it was necessary to treat syphilis at all.

The setup of the study reflects the strong suspicion that syphilis followed a different course in blacks than in whites. The speculation was undoubtedly tinged with racism, but not implausible: people who look radically different from ourselves often share a different genetic heritage and show up with different diseases or diseases in different forms: from Tay-Sachs disease, to sickle cell anemia, to the alarmingly heightened prevalence of hypertension in the African American population. On the other hand, most diseases are not different among us. So as the protocol started, no one really knew whether to look for different effects, black and white.

The original protocol called for six groups, matched by uncertainty:

a. Black, syphilitic, treated	b. Black, syphilitic, untreated	c. Black, not syphilitic
d. White, syphilitic, treated	e. White syphilitic, untreated	f. White, not syphilitic

Comparing a/d with b/e, they might find out if treatment did any good over the long run, or maybe did more harm than good;

Comparing b/e with c/f, they could find out if syphilis untreated really shortened the lifespan, or if treatment might safely be foregone;

Comparing a/b/c with d/e/f, they could find out if there was a real difference between blacks and whites.

Let us continue with the study. Early in the study period, the Great Depression hit the research community. The Rosenwald Fund had to withdraw its support. Without the Rosenwald Fund, the PHS did not have the resources to implement treatment (they could not purchase the drugs). White subjects, with or without syphilis, could not be recruited—there was no money to compensate them. Dr. Taliaferro Clark of the PHS suggested that the project could be continued with group (b) alone, the documented observation of African American men with syphilis, not given any treatment at all. At least they could find out if syphilis shortened the life by comparing the experimental group with the surrounding population. So that’s what they did.

With Rosenwald out, PHS enlisted the support of the Tuskegee Institute, a prestigious and trusted college for African American students, whose reputation would recommend the study to the African American community of the area. In return, Tuskegee Institute received money, training for its interns, and employment for its nurses; it was joined by black church leaders and other community leaders, including plantation owners, who also encouraged participation.

For the participants, the project seemed to be a reasonably good deal. At this time, African Americans had very little access to medical care. For many of the subjects, the examination provided by the PHS physician was the first medical examination they had ever received. Along with free examinations, food and transportation were supplied to participants; burial stipends were used to get permission from family members to perform autopsies on study participants. The controversy over “Tuskegee” comes from the information supplied, or rather denied, to the participants.

While study participants received medical examinations, none were told that they were infected with syphilis; the reason for the examinations and the “treatments” they were

receiving was held to be “bad blood.” PHS officials not only denied study participants any treatment there was, but also prevented other agencies from supplying treatment. As World War II gathered steam, many participants were drafted, and after blood tests, ordered to be treated for syphilis; the PHS managed to block the treatment. In 1943, the PHS began to administer penicillin to patients with syphilis, but study participants were deliberately excluded. Through cooperating local health departments, the PHS managed to keep penicillin from the subjects until 1972, when a whistleblower, PHS physician Peter Buxton, broke the story to the Associated Press. Incredibly, the study continued even after its details were splashed all over the nation’s newspapers.

The popular reaction to the Tuskegee Study—astonishing to the investigators who had carried it out—was outrage that heartless scientists could take this population of vulnerable people and turn them into lab rats (or guinea pigs) just to advance Science (and their own careers). Comparisons to the horrendous “experiments” of the Nazi concentration camps surfaced immediately. The public demanded action from their lawmakers, and got it. The Congressional subcommittee on Health, Education and Welfare, chaired by Senator Edward Kennedy, held hearings on clinical research in 1973, resulting in a rewrite of the Health, Education, and Welfare regulations on working with human subjects. Turning to the private law, the study subjects filed a \$1.8 billion class action suit in U.S. District Court, also in 1973; in December of 1974, the U.S. government paid \$10 million in an out of court settlement. Nor did the outrage subside at that time: years later, when the plague of HIV-AIDS hit the African American community, public health workers had a great deal of trouble getting African American men to consult physicians, because of suspicions (and superstitions) traceable to the Tuskegee scandal. Tuskegee still stands as a paradigm case of an immoral experiment. Is this entirely just?

Let us consider. When Rosenwald pulled out, and the federal government pulled most of its money out of the project, a few stalwarts were left to decide what to do. They decided that since they had no more money for treatment, groups a & d were out; since they had blessed little money to follow up on people who were off to California as fast as they could go, following groups c & f would not be practical; and all of the people in group e were off to less drought and dust-stricken parts also. But it would be possible to follow group b, mostly sharecroppers who were stuck in the area; since many of the members of the crew had already contributed time and effort, it was decided to keep the project going just following group b. Was that research unreasonable? Not necessarily *at that time*. Given the outcome—that members of group b were still alive to receive compensation from President Clinton, while no one originally from groups a and d came forward—it would seem that they demonstrated in fact that the treatments available at that time, amalgams of heavy metals, probably did do more harm than good. That justification disappears in 1943, when penicillin became available—no matter how much they wanted to find out if syphilis needed to be treated at all, when they had a safe and effective remedy, they should have used it, and the PHS’s efforts to keep their subjects away from penicillin was inexcusable.

Was anything learned in the Tuskegee Study? Yes, that the medications that incorporated heavy metals did more harm than good. But more importantly, the reaction to the study alerted the scientific community to the importance of two

factors that they had completely ignored—that subjects had a right to be informed about the nature and purpose of the study, and of the risks to which they would be subjected, and that issues of social justice—of the vulnerability of their subjects, especially vulnerability resulting from differences of race, gender, and socioeconomic status, should be considered in the design of the study.

There were other cases of human subjects research that caused controversy. In 1963, a cancer team at the Jewish Chronic Diseases Hospital injected cancer cells into a goodly number of elderly demented patients. The point was to see if the elderly body would reject the cells in the same way younger bodies had been shown to do. The investigators claimed that all the patients had been asked if they would like to participate, and all had agreed. Of course they had omitted to mention that the injection was of cancer cells, for they did not want to alarm the patients. Those who reviewed the study felt, correctly, that omitting the information about the content of the cells made “informed consent” a sham. The investigators were disciplined, and in the end, no patient contracted cancer. Then, in 1966, Henry K. Beecher of Harvard Medical School documented 22 cases (down from 32 in an earlier version) of apparent abuse of subjects published in the medical journals of the time; his article, published in *The New England Journal of Medicine*, brought the medical profession to its feet and evoked a call for much more careful regulation of clinical research. Ironically, neither the Tuskegee study, the Jewish Chronic Diseases Study, nor the Beecher article actually precipitated federal action to regulate research. That stimulus came from opposition to fetal research, which had begun with the *Roe v. Wade* decision, establishing the permissibility of abortion (and therefore of the availability of fetuses to study). In the wake of that decision, several lines of research immediately suggested themselves, especially the testing of drugs to discover whether they crossed the placenta in pregnancy. From 1957 through 1960, a drug called thalidomide was sold as a sedative effective in pregnancy. Unfortunately, it turned out to be a teratogen; if taken when the fetal limbs were forming, the child was likely to be born with grotesquely deformed limbs. No one wanted a repeat of that. So after *Roe v. Wade*, a woman seeking an abortion could earn a small stipend by taking an experimental drug before the abortion and consenting to the examination of the fetus afterward to find out if it contained traces of the drug. If it did not (in repeated trials), then the drug did not cross the placenta and was safe for pregnant women to take. When the political furor over abortion extended to fetal research, the U.S. Congress was pressured to take up the cause of regulation of research with human subjects, which it did with the 1974 formation of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, a committee specifically charged to come up with sensible regulations in an ethically problematic field, research with human subjects.

The Commission did its work well, and in April of 1979 issued the Belmont Report, summarizing the principles that seemed to have been at work in the decisions on regulation: beneficence (including non-maleficence), justice and respect for persons. These have been the foundation, not only of the ethics of human subjects research, but of most of the fields of applied ethics, ever since.

What happens when the part of the human subject that is being investigated is his mind, his opinions, his set of beliefs that prepare him to work in the world? Consider the famous Milgram investigations into obedience to authority.

Stanley Milgram and the Research on Obedience to Authority: When the volunteers arrive to participate in Milgram's study, he puts them in teams of three: a "teacher", a "learner," and of course the "experimenter," in charge of the study, a biologist in a white coat, and an electric generator with wires. Only the "teacher" is an actual participant, unaware of the arrangements for the experiment; the "learner" is a confederate of the experimenter, an amateur actor trained to act for the role. The experiment is described to teacher and learner as part of a study of memory and learning: the study is designed to find out if "negative reinforcement," i.e. punishment, is helpful in getting people to learn. The participant and the actor are then told to draw slips of paper to determine who will take the role of teacher and who the learner; in actuality, both slips of paper are marked "teacher." (The actor claims that his says "learner.") At this point, the "teacher" and "learner" were separated into different rooms where they could talk to each other, but not see each other. (In one version of the experiment—it had multiple iterations—the actor mentioned to the participant that he had a heart condition.)

The teacher, the actual subject of the experiment, was given a 45-volt electric shock from the generator as a sample of the shock that the learner, wired to the generator, would receive during the experiment. The teacher was then given a list of word pairs to teach the learner. After reading the whole list to the learner, now hidden behind a screen and apparently wired to the generator, the teacher would then read the first word of each pair and read four possible answers. The learner would press a button to indicate his response. If the answer was correct, the teacher would read the next word pair. But if the answer was incorrect, the teacher would administer a shock to the learner, with the voltage increasing for each wrong answer. (In reality, there were no shocks, of course. After the actor disappeared behind the screen, he set up a tape recorder integrated with the electro-shock generator, which played pre-recorded sounds—grunts, groans, increasing to screams of pain—for each shock level. The use of the pre-recorded sounds made sure that each subject heard exactly the same reactions, to make the results comparable.) After a number of voltage level increases, the actor started to bang on the wall that separated him from the subject, and, in some iterations, complain about his heart condition. After several shocks of increasing ferocity, all responses by the learner would cease.

Many of the subjects, ordinary citizens all, worried about the pain apparently felt by the learner, and questioned the purpose of the experiment and asked permission to stop. Most continued after being assured that they would not be held responsible. Some exhibited signs of extreme stress as the screams of pain increased. But if the subject expressed a desire to halt the experiment, he was given a succession of verbal prods by the experimenter, in this order:

1. Please continue.
2. The experiment requires that you continue.
3. It is absolutely essential that you continue.
4. You have no other choice, you *must* go on.

If the subject insisted on stopping at this point, the experiment was halted. If he continued, it was stopped after the subject had given the maximum 450-volt shock three times in succession. Had these shocks been real, the "learners" would have been dead.

Would we, any of us, have administered shocks we thought were lethal to a total stranger, just because the professor in a white coat told us to? Before conducting the experiment, Milgram polled Yale University psychology majors as to what they thought the results would be, and went on to poll his colleagues in

the psychology department. All of the poll respondents believed that only a few subjects would be prepared to inflict the maximum voltage. But in Milgram's first set of experiments, 65 % (26 of 40) of experiment participants administered the experiment's final 450-V shock, though many were very uncomfortable doing so; at some point, every participant paused and questioned the experiment. Some said they would refund the money they were paid for participating in the experiment! But no participant steadfastly refused to administer shocks *before* the 300-V level. Later, Milgram and other psychologists performed variations of the experiment throughout the world, with similar results, although unlike the Yale experiment, resistance to the experimenter was reported anecdotally elsewhere. The only consistent variation in results stemmed from the place the experiment was conducted: the greater the prestige and respectability of the locale (Yale laboratory vs. office in the inner city), the higher the level of obedience.

Later analyses of the results of the experiment show a remarkably consistent percentage of subjects willing to inflict the final voltages, 61–66 %, regardless of time or place. (There is a little-known footnote to the Milgram Experiment, reported by Philip Zimbardo: none of the participants who refused to administer the final shocks insisted that the experiment itself be terminated, nor left the room to check the health of the victim without requesting permission to leave, as per Milgram's notes and recollections, when Zimbardo asked him about that point.)

The Milgram Experiments were promptly criticized on ethical grounds, because of the extreme emotional stress suffered by the participants, during and after their participation. Had the subjects been asked beforehand if in obedience to orders they would hurt or kill a stranger, they would surely have said they would not. But look, when the choice was theirs, they did it. They are not the persons they thought they were; the kind of suffering they went through has since been called "inflicted insight." Yet despite the suffering, 84 % of former participants surveyed later said they were "glad" or "very glad" to have participated.

The experiments provoked emotional criticism more about the experiment's implications than with experimental ethics. Joseph Dimow, a participant in the 1961 experiment at Yale University, wrote about his early withdrawal as a "teacher," suspicious "that the whole experiment was designed to see if ordinary Americans would obey immoral orders, as many Germans had done during the Nazi period." In fact, that was one of the explicitly-stated goals of the experiments. In the preface to *Obedience to Authority*, the book recounting the experiments, Milgram had written "The question arises as to whether there is any connection between what we have studied in the laboratory and the forms of obedience we so deplored in the Nazi epoch."

It is unfashionable to defend the Milgram experiments at this time. But it should be pointed out that just prior to these studies, in the 1950s, the psychological literature had been full of speculation on "the authoritarian personality," a personality type that was supposed to predispose its owner to obey authority unthinkingly, jump off a cliff if commanded, that sort of thing. The nation was in agreement that Germans had such personalities, which is why they could imprison and abuse Jews and carry out unlawful executions. Americans, we were told,

had “democratic personalities” and would *never* obey unjust or unreasonable or harmful orders. This literature, part of a decades-long effort to explain the Third Reich to the stunned nation, was ended by the Milgram experiments; Americans, it turned out, were very likely to engage in hurting people on command, just like Germans. And the Americans in question were not soldiers, trained to obey, for whom disobedience would mean instant death, but ordinary citizens under no compulsion stronger than the instructions of an unarmed man in a white coat.

The Stanford prison experiment, conducted in 1971 by social psychologist Philip Zimbardo at Stanford University, was supposed to find out how prisoners and prison guards adapted to their roles in any prison. Twenty-four male undergraduates were selected, carefully screened for mental stability and general health, and assigned the roles of guards and prisoners in what was described as a “two-week prison simulation.” They were to be paid \$15 per day for the two weeks. The study took place in a mock prison in the Stanford psychology building. Zimbardo himself took on the role of “superintendent,” while his research assistant became the “warden.” Contrary to the predictions of their professors, the participants rapidly took on their roles. The guards became authoritarian to the point of sadism, while the prisoners exhibited all of the stress symptoms that are found among all victims in institutional settings. After sensing that everyone had been too absorbed in their roles, including himself, Zimbardo terminated the experiment after six days.

The experiment had set out to test the idea that the inherent personality traits of prisoners and guards led to observed abusive prison situations. In order to duplicate what he took to be the “dehumanizing” characteristics of prisons and other total environments, Zimbardo provided the guards with weapons, uniforms, and mirrored sunglasses to prevent eye contact.

Prisoners wore ill-fitting smocks and stocking caps. Guards called prisoners by their assigned numbers, sewn on their uniforms, instead of by name. A chain around their ankles reminded them of their roles as prisoners. They were fingerprinted, photographed, then taken to their mock prison where they were strip-searched and given their new identities.

The experiment quickly got out of hand. Prisoners suffered—and accepted—sadistic and humiliating treatment from the guards. They became upset, sick, and by the termination of the experiment, many showed signs of severe emotional disturbances. The prisoners had staged riots and rebellions, and spent much of their time in schemes to outfox the guards. The guards worked continually to keep the prisoners from organizing, to prevent escapes, and quickly resorted to fire extinguishers to quell insufficiently submissive behavior. Guards forced the prisoners to count off repeatedly as a way to learn their prison numbers, and to reinforce the idea that this was their new identity. Guards soon used these prisoner counts as another method to harass the prisoners, using physical punishment such as protracted exercise for errors in the prisoner count. Several guards became increasingly cruel as the experiment continued. Interestingly, most of the guards were upset when the experiment concluded early.

The prisoners weren’t. By the end of the six days, many “prisoners” had asked to leave the experiment (be granted “parole”) even at the cost of denial of their stipend. (This result puzzled Zimbardo; after all, the money was the major reason they stayed to that point. He concluded that the prisoners must have “internalized” their roles.) It was clear when this experiment ended that the situation, not their inherent psychological dispositions, governed the behavior of the participants; the hypothesis was disconfirmed.

The experiment was, and is, widely criticized as being unethical and bordering on unscientific. Current ethical standards of psychology would not permit such a study to be conducted today. The study would violate the American Psychological

Associate Ethics Code, the Canadian Code of Conduct for Research Involving Humans, and the Belmont Report. Critics including Erich Fromm challenged how readily the results of the experiment could be generalized. Fromm specifically writes about how the personality of an individual does in fact affect behavior when imprisoned (using historical examples from the Nazi concentration camps). This runs counter to the study's conclusion that the prison situation itself controls the individual's behavior. Fromm also argues that the amount of "sadism" in the "normal" subjects could not be determined with the methods employed to screen them. Further, because it was a field experiment, it was impossible to keep traditional scientific controls. Zimbardo was not merely a neutral observer, but influenced the direction of the experiment as its "superintendent." Conclusions and observations drawn by the experimenters, it was said, were largely subjective and anecdotal, and the experiment would be difficult for other researchers to reproduce.

Some of the experiment's critics argued that participants based their behavior on how they were expected to behave, or modeled it after stereotypes they already had about the behavior of prisoners and guards. In other words, the participants were merely engaging in role-playing. Another problem with the experiment was certain guards may have changed their behavior because of wanting to conform to the behavior that they thought Zimbardo was trying to elicit. In response, Zimbardo claimed that even if there was role-playing initially, participants internalized these roles as the experiment continued.

The larger question raised by the prison experiment, the same one raised by Milgram's study in obedience to authority and possibly more forcefully, is why on earth the participants, college students all, did not remember, and act on, their roles in the larger system, which surely should have been more solidly "internalized" over a lifetime? This was a voluntary study. Unlike German citizens threatened with loss of employment, houses, liberty and life if they disobeyed, or prisoners in the California prison system, put there against their will, these participants could have, and clearly should have, walked out at any time. They were not surrounded by barbed wire. They were in the basement of the Psychology Building. Nor were they isolated; they had each other to talk to. Why did not just one of them—especially when the study really got going and it was evident that some participants were getting sick—point out to the others that this game was not as much fun as they'd thought it would be, and that it was time to gather the bats and balls and go home to supper?

Given the universal condemnation of the Milgram and Stanford experiments, it would seem that we are not ready for the "inflicted insight" contained in both their results. We are not the people we thought we were, as individuals or as a nation. We will need a much more determined effort to internalize ethics if we are to become the people we ought to be.



<http://www.springer.com/978-3-319-00362-7>

The American Experience in Bioethics

Newton, L.H.

2013, VIII, 81 p., Softcover

ISBN: 978-3-319-00362-7