THE GREAT COURSES®

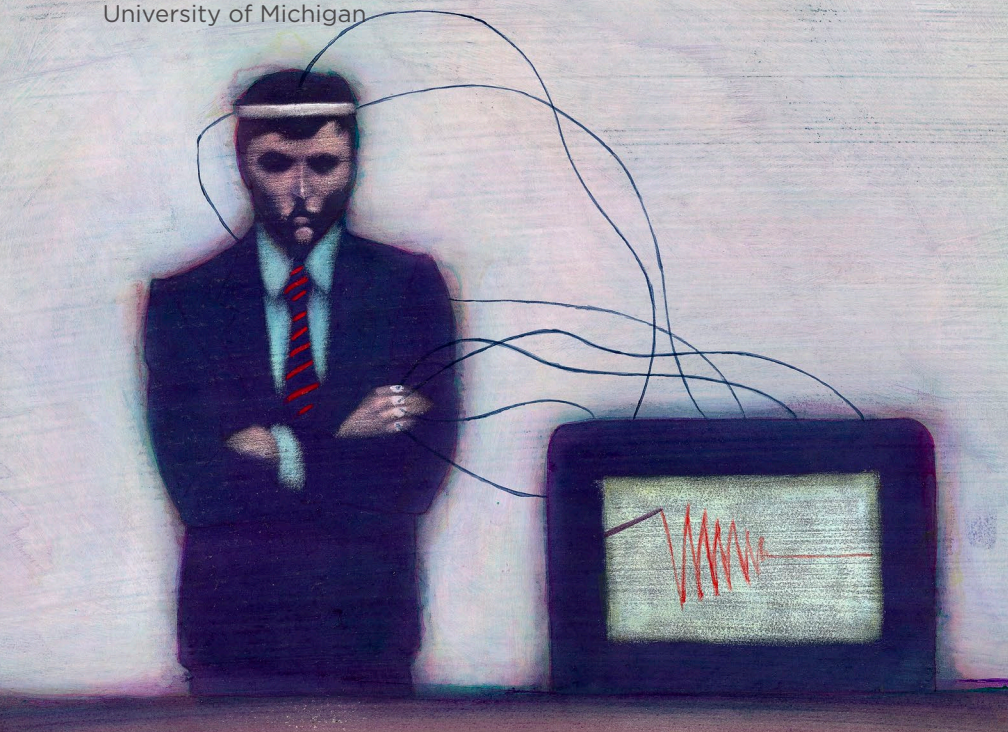
Topic
Science

Subtopic
Neuroscience & Psychology

Shocking Psychological Studies and the Lessons They Teach

Course Guidebook

Professor Thad A. Polk
University of Michigan



THE GREAT COURSES®

4840 Westfields Boulevard | Suite 500 | Chantilly, Virginia | 20151-2299
[PHONE] 1.800.832.2412 | [FAX] 703.378.3819 | [WEB] www.thegreatcourses.com

Thad A. Polk, PhD
Arthur F. Thurnau Professor of Psychology
University of Michigan

Thad A. Polk is an Arthur F. Thurnau Professor in the Department of Psychology at the University of Michigan. He received a BA in Mathematics from the University of Virginia and an interdisciplinary PhD in Computer Science and Psychology from Carnegie Mellon University. He received postdoctoral training in cognitive neuroscience at the University of Pennsylvania before joining the faculty at the University of Michigan.

Professor Polk's research combines functional imaging of the human brain with computational modeling and behavioral methods to investigate the neural architecture underlying cognition. Some of his major projects have investigated changes in the brain as we age, contributions of nature versus nurture to neural organization, and differences in the brains of smokers who quit compared with those who do not. Professor Polk regularly collaborates with scientists at the University of Texas at Dallas and at the Max Planck Institute for Human Development in Berlin, where he is a frequent visiting scientist. At the University of Michigan, he is an associate chair of the Department of Psychology and the chair of the Health Sciences and Behavioral Sciences Institutional Review Boards.

Professor Polk regularly teaches large lecture courses as well as small seminars on topics ranging from the human mind and brain, to cognitive psychology, to computational modeling of cognition. His teaching at the University of Michigan has been recognized with numerous awards, including the Excellence in Education Award from the College of Literature, Science, and the Arts as well as the Arthur F. Thurnau Professorship, the university's highest undergraduate teaching award. He was also named to The Princeton Review's list of the Best 300 Professors in the United States.

Professor Polk's otherGreat Courses include *The Learning Brain*, *The Aging Brain*, and *The Addictive Brain*.

ACKNOWLEDGMENTS

I'd like to express my heartfelt gratitude to Erin Freiburger, Elsy Nouna, Samantha Roy, Bashair Pasha, and Ryan Rich. These amazing students helped identify potential topics to include in the course, performed extensive literature reviews, and provided detailed feedback and suggestions on drafts of the lessons. The course is significantly better than it would have been without their many substantial contributions. Thank you so much!

I'd also like to thank my wife, Norma, and my youngest daughter, Lydia, for their patience and support as I spent many hours working on this course. I love you!

TABLE OF CONTENTS

INTRODUCTION

GUIDES

SUPPLEMENTARY MATERIAL

DISCLAIMER

This series of lessons contains graphic descriptions of violence, sexual violence, sexually explicit language, and offensive and dehumanizing language, which may be disturbing and may not be suitable for minors or other audiences. The opinions and positions provided in these lessons reflect the opinions and positions of the relevant lecturer and do not necessarily reflect the opinions or positions of The Teaching Company or its affiliates.

The Teaching Company expressly DISCLAIMS LIABILITY for any DIRECT, INDIRECT, INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES OR LOST PROFITS that result directly or indirectly from the use of these lectures. In states that do not allow some or all of the above limitations of liability, liability shall be limited to the greatest extent allowed by law.

SCOPE

Psychological studies of human behavior can be among the most fascinating in all of science. We all want to have greater insight into our own thoughts, feelings, and decisions, and psychology offers the hope of providing that insight. But psychology also faces a number of ethical challenges that most other scientific fields don't have to face. After all, when the objects of study are human beings, scientists have to take special care to protect the rights of those participants. Unfortunately, not all psychologists have done so, and this course will review some of the most shocking examples of ethically questionable studies in the history of psychology.

The course discusses studies examining evil human behavior that in some ways were evil themselves. You will learn about controversial studies that were conducted on vulnerable children and others that were secretly conducted by the government and military. You'll also learn about some ethically doubtful studies of sexual behavior and gender identity.

Every study is placed in its historical context, and you will be walked through exactly what was done and what results were obtained. You'll also dive into the studies' ethics and think about the principles that should have been considered, as well as how the studies violated those principles.

The course begins by discussing a controversial study in which the News Feeds of nearly 700,000 Facebook users were manipulated without their knowledge in order to investigate the effects on their posts. You'll

learn why this study led to significant public outrage as well as how the scientists who conducted the study responded.

Then, the course reviews some of the history that led to current human research regulations, including the infamous Tuskegee study, which followed impoverished, uneducated syphilis victims for 40 years without providing them with treatment or even informing them of their diagnosis. You'll also learn about the significant changes in regulations that were put into place after the study was exposed. In particular, you'll be walked through the key ethical principles of respect for persons, beneficence, and justice that must now be considered before any research involving human beings can be conducted.

With that historical background in place, the course turns to 2 of the most famous studies in the history of psychology: Stanley Milgram's obedience study and the Stanford Prison Experiment. Both of these studies shed new light on why people sometimes behave in unethical ways, but ironically, the studies themselves had serious ethical problems that will be considered in some depth.

Next, the course reviews shocking psychological studies involving children. You'll encounter a study in which adopted twins who had been separated at birth were both extensively studied for years without ever being informed that they had a sibling, much less a twin. You'll also learn about a study in which orphan children were repeatedly told that they stuttered in order to investigate the effects on their speech development.

The course also walks you through the history of secret psychological experimentation by the US government and military, including the CIA's infamous MK-Ultra project and the testing of psychoactive chemicals like LSD and PCP on unsuspecting military personnel.

Then, the course discusses some very controversial studies of sexual behavior and gender identity. You'll be introduced to the controversial Tearoom Trade study, in which a social scientist lied about his identity so that he could observe and document sexual acts in public restrooms. You'll also learn about the tragic John/Joan case, which involved a baby boy who was raised as a girl at the urging of a famous psychologist.

The course concludes by addressing current and future ethical dilemmas in the field of psychology. You'll consider ethical problems associated with analyzing the enormous data sets produced by social media, internet searches, and smartphone apps. You'll also learn about some recent cases of scientific fraud, in which psychologists have manipulated—or even completely fabricated—scientific results. Finally, you'll encounter some recent studies that suggest that many published findings in the psychological literature are unreliable and cannot be replicated.

The course will give you new insights into the world of psychological research and make you a more discerning consumer of the studies you hear or read about in the popular press. It will also give you a renewed appreciation for the self-correcting nature of science and for the way that psychology is constantly evaluating its practices and findings to ensure that future studies avoid the problems of the past.

Table of Contents

LESSON 1

LESSONS FROM TUSKEGEE AND FACEBOOK

This lesson will introduce some of the ethical dilemmas that scientists face when conducting research with human subjects. You will discover some of the many ways that research can go wrong by learning about the infamous Tuskegee syphilis study. But the mistakes that happened in Tuskegee are far from the only mistakes that researchers studying human beings have made.

THE FACEBOOK EMOTIONAL CONTAGION EXPERIMENT

The Facebook emotional contagion experiment was conducted during a weeklong period in January 2012 by researchers at Facebook and Cornell University. The scientists were interested in examining whether positive and negative emotions are contagious and specifically whether Facebook users who see fewer emotionally positive posts in their News Feed produce fewer positive posts themselves.

To investigate this question, the researchers made changes to the News Feeds of about 689,000 Facebook users. They randomly deleted 10% to 90% of posts that contained positive words from some users' News Feeds; other users had 10% to 90% of posts containing negative words deleted. Still others had the same proportion of random posts deleted to serve as a control. Then, the scientists analyzed how deleting positive and negative posts affected the status updates of the users.

The researchers found that when positive posts were deleted, the percentage of positive words in people's status updates decreased by a very slight amount relative to the control participants, while the percentage of negative words increased slightly. Conversely, when negative posts were deleted, fewer negative words and more positive words appeared in people's updates. The scientists concluded that emotions expressed by Facebook friends can and do influence our own mood. In other words, emotions that are expressed on social media can be contagious.

This study was published in the *Proceedings of the National Academy of Sciences*, which is one of the most prominent scientific journals in the world. And the study has had a major impact on the field—although it's not the impact that the authors had hoped for.

Almost immediately, a number of people slammed the study as being unethical. Critical articles appeared in academic journals as well as in popular outlets like *Forbes*, *The Atlantic*, *The Guardian*, and *The New York Times*. The study also triggered thousands of protest posts on social media. But why the outrage?

One of the main concerns was that the researchers tried to manipulate the mood of nearly 700,000 people. And for 1/3 of those people, the goal was to make their mood worse rather than better. Many people were concerned that the study could be causing emotional harm to hundreds of thousands of people.

Worse yet, the people in the experiment weren't even aware that they were in an experiment. They were just browsing their Facebook site like they always do, checking their friends' posts and updating their own status. They never gave the scientists permission to

As you learn about each of the shocking studies featured in this course, ask yourself these questions:

What is it about the study that crossed a line?

What line was crossed?

What policies and principles should be in place to make sure that line isn't crossed in the future?

analyze their status updates, much less to mess with their News Feeds. Or did they?

Before people can sign up for Facebook, they do have to agree to certain terms, including a data use policy. And the published paper actually mentioned this fact.[1]

Furthermore, it turns out that Facebook manipulates News Feeds all the time and analyzes the results for internal purposes. Given that Facebook can't feasibly include every post that friends make in a user's News Feed, it uses an algorithm to determine which posts to include and which posts to omit. And different versions of that algorithm are regularly tried in order to improve it.

This study could therefore be seen as a standard algorithm test. The main difference is that this test was published in the scientific literature.

The scientists involved also pointed out that no posts were deleted; they just didn't show up on some loads of the News Feed. But a post that was omitted for one load could show up in the very next load. And all posts were always visible on friends' Timelines.

Furthermore, the effect of the experimental manipulation was very small. In fact, out of every 1000 words in the status updates in the week following the experiment, people in the experimental conditions produced about 1 fewer

emotional word compared to the control subjects. The participants were almost certainly completely unaware of any effect on their mood.

Finally, the scientists pointed out that the motivation for the study was altruistic. Many people are concerned that seeing lots of positive posts from friends might make people feel left out and unhappy. The researchers wanted to find out if this concern was valid, but their results actually suggest the opposite: Seeing positive posts makes Facebook users more positive, not less.

THE TUSKEGEE SYPHILIS STUDY

The Tuskegee syphilis study has had a bigger impact on the regulations governing human research than any other study in history. It also illustrates some of the key principles that studies today must abide by. And that's because it violated pretty much all of them!

Syphilis is a bacterial infection that usually spreads through sexual contact. Today, syphilis can be easily treated with penicillin, but before penicillin came along, the disease was much more common, and the recommended treatments included repeated doses of heavy metals like mercury and arsenic, which are toxic themselves.[2]

There was very little data on how effective these treatments were relative to no treatment, especially in patients who had had the disease for a long time. After all, syphilis can go dormant for decades before becoming active again. It's therefore natural to ask

2 Some doctors even recommended intentionally exposing syphilis patients to malaria because the associated high fever could help in the treatment of the disease.

how much benefit there is in treating late-stage patients with something like mercury compared with not treating them at all. Are people who aren't treated more likely to die? Do they experience significantly more health problems later in life? If so, what are those health problems?

Those are the kinds of questions that the Tuskegee syphilis study wanted to answer. To do so, researchers would have to follow syphilis patients who received treatment as well as patients who didn't receive treatment in order to be able to compare the health outcomes in the 2 groups and figure out how effective the treatments really were.

The problem was finding a lot of people with syphilis who had never been treated. But it turns out that surveys in the early 1900s discovered that there was a large number of African American men in rural areas in the South who had contracted syphilis but were never treated for it. Most of these men were sharecroppers who were typically very poor, were not able to read, and had almost no access to quality medical care.[3]

In 1932, researchers established themselves in Tuskegee, Alabama, and tried to recruit as many of these poor, black syphilis victims as possible to study the effects of the untreated disease.

The researchers took medical histories, performed medical exams, drew blood, and conducted diagnostic tests on more than 1700 people. They also wanted to perform spinal taps to check for evidence of infection in the brain and central nervous system, but spinal taps were painful and could also lead to severe headaches for days if done incorrectly. In order to encourage people to come in, the researchers sent out a letter portraying the spinal tap as a treatment rather than

3 A sharecropper was essentially a person who farmed someone else's land in exchange for some of the crops.

a test. The letter worked, and more than 300 people came in for spinal taps.[4]

By this point, the researchers had a lot of data from a large number of late-stage syphilis victims. Most of the team assumed that data collection was done. But one of the researchers argued that they should continue to study these untreated syphilis victims over the long term. These indigent sharecroppers were going to return to their lives, and almost none of them would get treated for their syphilis. So why not continue testing them over the next 5 to 10 years to see how the untreated disease developed? And if the participants would agree to an autopsy after their death, then the scientists would have much more direct evidence about the ways that untreated syphilis can damage the human body.

The researchers decided to continue the study. And to encourage people to participate, the scientists decided to offer what they called "treatment" but actually consisted of aspirin and tonics. Participants would also be given small amounts of oral mercury if they asked for it, but not at the recommended dosage.

The study team also offered to cover $50 in burial costs for any participant that agreed to an autopsy after his death. This was a significant incentive to the members of this very poor community, who often couldn't pay for a reasonable funeral.

Ultimately, the team managed to recruit about 400 people who had untreated syphilis for at least 5 years and who had not received significant treatment for it. They also recruited another roughly 200 people who tested negative for syphilis

4 The letter said, "You will now be given your last chance to get a second examination. This examination is a very special one and after it is finished you will be given a special treatment if it is believed you are in a condition to stand it." The letter concluded: "Remember this is your last chance for special free treatment. Be sure to meet the nurse."

and 275 people who had been treated for syphilis during the early stages of the disease.

Very quickly, the researchers discovered that treatment made a significant difference, even the clearly inadequate treatment of the time period. For example, in 1936, they published a paper pointing out that the untreated group was 2 to 3 times more likely to have cardiovascular and neurological damage than the treated group.

The team now had a very clear answer to whether people with syphilis were better off in the long term if they were treated with heavy metals or were left untreated: Treatment definitely helps. Today, researchers in such a situation would be required to stop the study and inform the participants about what they had found. And ideally, they would also offer the treatment to everyone.

But the Tuskegee study scientists apparently didn't consider informing the participants about their findings or trying to treat them. And they didn't have the money to provide treatment to everyone even if they wanted to. So they continued the study. They continued to investigate how things progressed and to perform autopsies as participants died.

Then, in the 1940s, penicillin became available, and scientists demonstrated that it was very effective in killing the bacteria that causes syphilis. In fact, penicillin is still the drug of choice in treating syphilis to this day. And by the end of the 1940s, penicillin was regularly being used to actually cure syphilis, at least when caught in the early stages. It was also often used in treating late-stage syphilis, although there was debate

about whether it should be because the long-term effects of penicillin were not yet known.

Today, researchers would be required to tell participants about a new treatment that is potentially more effective than what they're getting. And most researchers would offer the treatment for free if they had the resources to do so.

But that's not what the researchers in this study did. From their point of view, if their no-treatment group started receiving treatment, then that would make it impossible to draw conclusions about the effects of treatment. Essentially, they would lose their control group and wouldn't have anything to compare against. So they didn't tell the study participants about the availability of penicillin. In fact, there's some evidence that they actively tried to prevent the study participants from getting treatment.[5]

The study continued for decades. And despite the very large number of people directly involved in the study over such a long period of time, none of them raised concerns, at least publicly, about whether it was ethical or not.

Likewise, despite decades of publications and public presentations about the study and its results, very few scientists or members of the public questioned the study's ethics or the ethics of the researchers involved.

The person who is most responsible for exposing the ethical problems with the Tuskegee syphilis study is probably Peter Buxtun, a social worker and epidemiologist who was tracking down people in San Francisco

5 The entry of the United States into World War II meant that study participants might get drafted, and if they did, they would get medical exams and likely receive treatment for syphilis. To avoid this possibility, the researchers contacted the head of the local draft board and explained the importance of the study. Almost none of the study participants ended up being drafted.

with syphilis and gonorrhea, interviewing them to try to identify sexual partners, and encouraging all those affected to get treatment.[6]

Buxtun found out about the syphilis study in 1965 and was very concerned. For years, he shared his concerns with friends and colleagues, filed formal protests, and even met with leaders in both the US Public Health Service and the Centers for Disease Control. But nothing was done to stop the study.

That all changed in 1972, when Buxtun gave documentation to an *Associated Press* reporter named Jean Heller. She published the story on July 25, 1972, with the title "Syphilis Victims in US Study Went Untreated for 40 Years."

The reaction was fast—and enormous. The 1950s and 1960s had seen the rise of the civil rights movement, and the fact that the Tuskegee study targeted poor, underprivileged black men was consistent with other examples of institutional racism. Reporters tracked down some of the study participants and discovered that most of them didn't realize what the study was about. The men mentioned being told that they had what was referred to as "bad blood," but many didn't know that this meant they had syphilis. And virtually all of them thought that the aspirins and tonics they had been given were a treatment for their so-called bad blood.

An ad hoc advisory panel was convened to investigate the study, and after reading many of the documents and interviewing some of the people involved, the panel determined that the study had been ethically unjustified. They pointed out that the participants had not been adequately informed about the true nature of the study and had therefore not given truly informed consent. The panel also mentioned that the men should have been treated with penicillin after it became available.

6 Ironically, Buxtun worked for the same government organization that oversaw the Tuskegee syphilis study: the US Public Health Service.

A class action lawsuit was filed in 1973 and was ultimately settled for $10 million. As part of the settlement, the US government agreed to provide lifetime medical benefits and burial services to all study participants who were still alive.

The study also came before a congressional hearing organized by Senator Ted Kennedy, who afterward proposed a National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The commission had 2 major charges: to identify the basic ethical principles that should govern all research involving human subjects and to develop guidelines that all human research should follow in order to be consistent with the ethical principles.

The commission worked on this task for a few years and ultimately drafted a report at the *Belmont* Conference Center in Maryland in 1976. Now known as *The Belmont Report*, it radically changed the way research involving human subjects is regulated.

THE BELMONT REPORT

The Belmont Report identified 3 key ethical principles that should be considered when conducting research with human subjects: respect for persons, beneficence, and justice. It also proposed practical guidelines that researchers should follow based on those ethical principles.

1 Respect for persons. Basically, researchers who do experiments with human subjects should treat those subjects with respect. The report identified 2 key ethical considerations related to this principle. First, human subjects should be treated as autonomous agents who have free will and should be given the opportunity

Most of the psychological studies that will be discussed in this course were conducted before *The Belmont Report* was published. And they all violated the *Belmont* principles and guidelines in one way or another. In fact, the case can be made that the way these studies violated the *Belmont* principles is precisely what makes them shocking by today's ethical standards.

to decide for themselves, without any coercion, whether they want to participate in the research study. Second, research involving people with diminished autonomy—such as children, prisoners, and those with mental disabilities—should provide such individuals with extra protections to ensure that they are not exploited or exposed to unnecessary risks.

In keeping with this principle, *The Belmont Report* recommended requiring informed consent whenever human research is conducted. Basically, potential research participants should be given enough information about the study to make a well-informed decision, that information should be presented in a way that the participants can understand, and participants should be able to make their own decision about participation without any threat of harm or inappropriate reward.

- The Tuskegee study violated this principle. In particular, the participants were not given complete information about the study's goals or their role in the study. And they were never given an opportunity to make an informed decision about whether they would like to participate in the study. You could also make a case that offering burial benefits in exchange for agreeing to an autopsy was coercive.

2 Beneficence. This refers to looking out for the well-being of research participants. In particular, do no harm, just like in medicine. More generally, this principle incorporates the idea of maximizing benefits while minimizing risks. This idea applies to not only the research participants but also society as a whole.

In light of this ethical principle, *The Belmont Report* recommended that all potential research involving human subjects should carefully consider both the benefits and risks of the research. Studies should only be carried out if the benefits outweigh the risks. And researchers are required to inform potential participants about the benefits and risks of a given study before asking for their informed consent.

- The Tuskegee study fell short in terms of beneficence, at least for the participants themselves. They were intentionally denied potentially helpful treatment for a serious disease and were not explicitly informed that this was happening.

3 Justice. Is the research study fair? In other words, does it treat different people equally, both in terms of sharing in the benefits of the research and in bearing the risks? The practical application of justice is that researchers must develop fair procedures for how they select people to participate in their studies. They must be sensitive to issues of gender, race, age, socioeconomic status, and culture and do their best to ensure that they are not being unfair to some social groups relative to others.

- The Tuskegee study would earn a failing score in terms of justice. The study specifically targeted indigent black men to participate in a study with significant risks, despite the fact that syphilis can affect anyone. Furthermore, the knowledge gained from the study largely benefitted people who were much better off financially and who had access to quality medical care.

SUGGESTED READINGS

Jones, *Bad Blood*.

Kramer, Guillory, and Hancock, "Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks."

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report*.

Reverby, *Examining Tuskegee*.

Table of Contents

LESSON 2

PUSHING GOOD PEOPLE TO DO BAD THINGS

Stanley Milgram's obedience study and the Stanford Prison Experiment are 2 of the most famous, but also infamous, studies in the history of psychology. Both studies taught us some very important lessons about human nature—specifically about how authority and power can have a profound influence on our behavior, and not always for the good. These studies significantly changed our understanding of human behavior and inspired entire lines of research that continue to this day. But both studies also raised ethical dilemmas.

STANLEY MILGRAM'S OBEDIENCE STUDY

Stanley Milgram's obedience study was conducted at Yale University in the early 1960s. This was around the time of the highly publicized trial of Adolph Eichmann, who was in charge of organizing the transportation of Jews to concentration camps in eastern Europe during World War II. He therefore played a central role in the death of around 6 million people.

Eichmann's trial was the first ever to be televised, and it generated huge interest all over the world. Before his trial, most people didn't fully recognize the magnitude of the atrocities committed during the Holocaust. But the television coverage of this trial brought home those atrocities to millions of people around the world.

Learning about the scope of the Holocaust led to many questions, including how could all the people involved in carrying out these acts of cruelty go along with it? After all, the Holocaust was a massive logistical undertaking that required thousands of people performing a huge variety of different jobs. Eichmann may have overseen this huge operation, but he obviously couldn't have done it alone. So why did all the German soldiers, and even many civilians, cooperate in this incredibly evil undertaking? Why didn't they simply refuse to participate?

The answer that most of these people gave afterward was that they had no choice—they were simply following orders. But wouldn't ordinary people refuse to obey orders like that? Wouldn't their conscience prevent them from doing so?

That's the question that Milgram was interested in answering. Milgram was Jewish himself and was deeply affected by what he learned about the Holocaust. He decided to design a scientific experiment to determine whether ordinary people really will follow orders even when those orders involve harming someone else.

Of course, most people wouldn't volunteer to participate in an experiment if they knew it involved hurting other people, so Milgram came up with a cover story. He posted fliers asking for men between the ages of 20 and 50 to participate in an experiment on learning and memory. When volunteers arrived, they were paired up with another person and were told that the experiment was designed to test whether people learn better when they are punished for mistakes. One of the participants would be the learner, who would try to learn a bunch of word pairs. The other person would be the teacher, who would test the learner's memory and punish mistakes by administering an electric shock.

But it turns out that only one of the 2 people was an actual participant. The other person was just pretending to be a participant and was actually a confederate who was working with Milgram. This confederate and the real experimental subject then drew a slip of paper out of a hat, supposedly to determine who was going to be the teacher and who was going to be the learner. They were then asked what their paper said. Both pieces of paper said "teacher," but the confederate claimed that

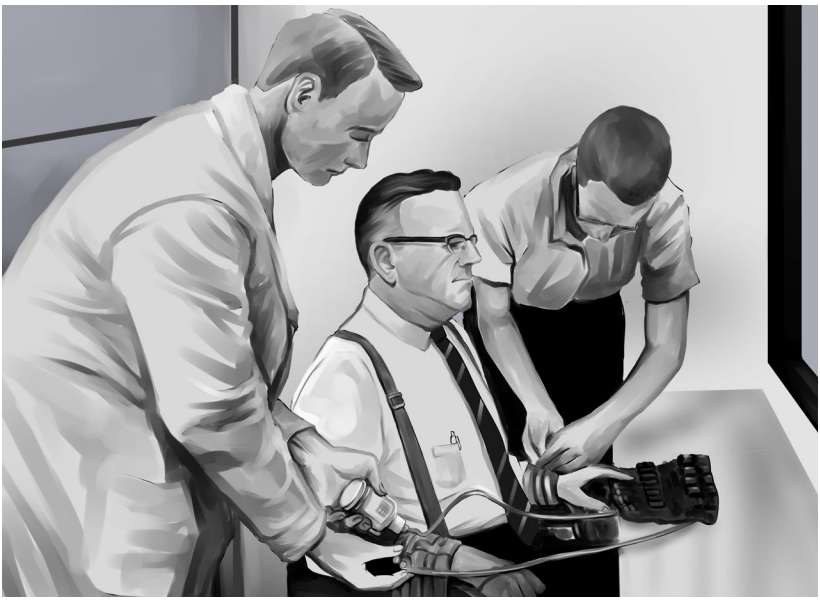
Consider This

In many psychology experiments, knowing the purpose of the study could change the participants' behavior, so it's crucial not to tell them the true nature of the experiment beforehand. In practice, this often means deceiving research participants in some way about the study's goal. Essentially, the psychologist has to lie to the participants or else the experiment won't work. Is that ever ethically justified? And if so, when and how should it be done? Keep these questions in mind as you read about the Milgram obedience study.

his said "learner." In this way, the real experimental subjects were always assigned to be the teacher and the confederate was always the learner.

Next, the teacher and the learner were both led into a room that contained a chair with straps that looked sort of like an electric chair. The fake learner was then strapped into this chair and an electrode was attached to his skin, supposedly to shock him whenever he made a mistake. An electrode paste was also applied with the explanation that it would help the learner avoid blisters and burns. Once strapped in, the confederate asked about the safety of the procedure, and the experimenter told him: "Although the shocks can be extremely painful, they cause no permanent tissue damage."

The real experimental subject was then led into an adjacent room, where he was seated in front of a machine that he was told would generate the electric shocks to punish the learner for mistakes. This shock



generator had 30 switches set in a horizontal line, and each switch was labeled with a voltage in 15-volt increments from left to right, starting at 15 volts and going all the way up to 450 volts. Furthermore, each group of 4 switches had a verbal label on the machine, ranging from "slight shock" to "danger: severe shock" and with the 2 rightmost switches simply labeled "XXX."

When one of these switches was depressed, a red light next to that switch turned on. Then, the machine produced an electric buzzing sound, a blue light labeled "voltage energizer" flashed, and a voltage meter swung to the right.

And all of this was fake—no shock was actually being delivered! The goal was just to convince the teacher that he was indeed administering painful shocks to the learner in the other room.

The teacher was also connected to the machine and given a 45-volt sample shock before the experiment began. And he really did receive a 45-volt shock, although it was generated by a 45-volt battery hidden inside the machine rather than by the machine itself.

After all this setup, the experiment finally began. First, the teacher was instructed to read a list of word pairs, which the learner was supposed to try to remember. Then, the teacher would read the first word from one of the pairs, and the learner was supposed to choose the word that was paired with that word. The learner was given 4 options and supposedly indicated his choice by pressing one of 4 buttons.

If the learner answered correctly, nothing happened, and the next test word was presented. But if the learner answered incorrectly, then the teacher was instructed to punish him by administering a 15-volt shock. And the teacher was told to increase the strength of the shock with each subsequent mistake—until the teacher either could not be convinced to continue or had administered the maximum shock of 450 volts 3 consecutive times.

Of course, the confederate in the other room wasn't really getting shocked. In fact, he wasn't doing the learning task at all. He was simply giving a predetermined set of responses with about 3 mistakes for every 1 correct answer. So the teacher was being asked to give a lot of shocks, and as far as the teacher knew, those shocks were becoming very, very strong.

Worse yet, as the shocks increased in intensity, the confederate in the other room would start to protest. For example, when the 300-volt shock was administered, the confederate would pound on the wall. In some versions of the experiment, the confederate would scream and demand that the experiment be stopped, even claiming that he was concerned about his heart. Typically, the protests would stop as the voltage level continued to increase and the confederate would go completely silent, suggesting that he was now either unconscious or potentially even dead. But the teachers were instructed to treat no response as a mistake and to continue to administer shocks even when the learner in the other room became nonresponsive.

The men who were serving as teachers would often turn to the experimenter to express reservations about continuing to shock the learner under these kinds of conditions. But the experimenter would always ask the teacher to go on using a sequence of prods. First, the experimenter would simply say, "Please continue" or "Please go on." If the teacher still protested, then the experimenter would say, "The experiment requires that you continue." If that wasn't enough, then the experimenter would say, "It is absolutely essential that you continue." Finally, if none of these prods worked, then the experimenter would say, "You have no other choice; you must go on."

The goal was to see whether ordinary people would continue to obey the experimenter even when they personally believed that they were causing the learner very significant pain and perhaps even permanent damage. Would ordinary

American men obey unethical orders like Nazi soldiers did during World War II?

The results were sobering. Out of 40 men who participated in the experiment as teachers, 26 of them—65%!—continued to obey the experimenter's instructions and never stopped administering the shocks. And subsequent experiments that were done with women or in other cultures or that were performed in slightly different ways have found the same thing: Nearly 2/3 of people will continue to obey orders from an authority figure even when those orders are unethical and clearly violate their own conscience.

The Milgram obedience study generated significant controversy when it was first published. And keep in mind that it was conducted more than 15 years before *The Belmont Report* was drafted.

Consider the Milgram study in light of the *Belmont* principles: respect for persons, beneficence, and justice. What most people find objectionable about the study is that it did not demonstrate an appropriate respect for the experimental subjects. In particular, they were told that they were participating in an experiment about learning and memory and were unaware that they were going to be asked to cause significant pain to other people when they volunteered. And as a result, most of them experienced quite significant psychological trauma.

Milgram himself reported that participants were observed to sweat, tremble, stutter, bite their lips, groan, and dig their fingernails into their own flesh. Three of the participants were even described as experiencing full-blown, uncontrollable seizures. That level of psychological trauma would never be allowed today.[7]

7 All legitimate psychological experiments must be reviewed by an institutional review board, which explicitly considers whether the study satisfies all of the *Belmont* principles. The Milgram obedience study clearly doesn't and would therefore not be able to be conducted today.

THE STANFORD PRISON EXPERIMENT

The Stanford Prison Experiment was conducted by Philip Zimbardo at Stanford University in 1971, about 10 years after Milgram's obedience study.[8]

Like Milgram, Zimbardo wanted to examine the causes of unethical behavior. But rather than studying the role of obedience to authority, Zimbardo wanted to examine the influence of the situation or environment—specifically, the environment of a prison.

Prisons are notorious for unethical behavior, and not just by prisoners. Prison guards at numerous correctional facilities have been found to be cruel and inhumane. But why? Is it that the people themselves are evil or unethical, or does the prison environment bring out unethical behavior in ordinary people?

Those are the questions that Zimbardo wanted to investigate. He decided to create a simulated prison, but one in which both the prisoners and guards were regular, everyday college students. He posted a newspaper ad seeking male college students for a scientific study of prison life. The ad mentioned that the study would last 1 to 2 weeks and that participants would be paid per day.[9]

More than 70 people responded to the ad. Zimbardo and his team then interviewed the potential participants and selected 24 men whom they thought were the most mature, the most stable, and the least involved in antisocial behavior.[10] Half these men were randomly assigned to serve

---

8 Zimbardo and Milgram were actually high school classmates!

9 The participants were paid $15 per day, which would be nearly $100 per day in today's dollars.

10 All of the participants were completely normal as far as the psychologists could tell, and none of them had any obvious tendencies toward unethical or antisocial behavior.

Consider This

The scientists conducting a study might have a conflict of interest and might not always consider the best interests of the people they are studying. For example, if a study is producing interesting results, the scientists running it might want it to continue even if the experience is unpleasant for the participants. This was a major problem in the Stanford Prison Experiment.

as prisoners in the study, and the other 1/2 were assigned to serve as prison guards.

Zimbardo and his team constructed a simulated prison in the basement of Stanford's Psychology Building, where the prisoners would live during the study. They built a few cells, each of which contained 3 cots, and the hallway outside the cells served as the prison yard. There was also a closet that was used to put prisoners in solitary confinement. Video cameras and microphones were installed to record the behavior of everyone involved.

The simulated guards were given khaki uniforms, a billy club, and a whistle. They also wore mirror sunglasses so that the prisoners could not see their eyes. The guards were strictly instructed not to physically harm the prisoners in any way, but otherwise, they were invited to use their discretion to keep order in the prison. An undergraduate research assistant served as the prison warden, and Zimbardo himself was the prison superintendent.

To increase the realism of the experience, Zimbardo asked the Palo Alto police to actually arrest the men assigned to be prisoners. The police went to their homes, put them in handcuffs, drove them to the police station, and went through a standard booking

procedure. The prisoners were then blindfolded and driven to the simulated prison, where they were searched, stripped naked, and sprayed with a delousing spray.

The simulated prisoners were then issued a smock with a prison ID number printed on the front and back for them to wear. They were also given rubber sandals and a nylon stocking, which they were ordered to put on their head to cover their hair. Finally, a chain was bolted around each prisoner's right ankle.

The goal was to make the simulated prisoners feel anonymous and humiliated like real prisoners undoubtedly often feel. They were also required to refer to each other by their prison ID number rather than by their name, and the guards referred to them in the same way.

The guards performed prisoner counts that were modeled after similar counts in real prisons. During these counts, the prisoners would be lined up in the hallway and had to recite their prison ID numbers. These counts provided an opportunity for the guards to interact with the prisoners and assert their authority. They happened at all hours, even in the middle of the night, when the prisoners were asleep.

During the first day of the study, not much happened. The prisoners didn't seem to take the study very seriously and even laughed at the guards' attempts to make the prison seem like a real one. But that changed dramatically on the second day.

The prisoners decided to stage a sort of rebellion. They took the nylon stockings off their heads, ripped the numbers off their smocks, and pushed their cots up against the cell doors to prevent the guards from getting in. They also began to curse at the guards and refused to follow their orders.

The guards responded by spraying a fire extinguisher through the bars of the cells to get the prisoners away from the doors. Then, they broke into the cells, removed the beds, took away the prisoners' smocks, and put the leaders of the rebellion in the solitary confinement closet.

Next, they decided to employ psychological tactics in an effort to prevent similar revolts in the future. Specifically, they withdrew privileges from the prisoners most involved in the rebellion. Those prisoners were left in their cells without their beds or smocks. Conversely, the prisoners least involved were given back their smocks and beds and were allowed to brush their teeth and wash themselves. These "good" prisoners were also given food in the presence of the "bad" prisoners, who were temporarily denied food.

After this rebellion, things got significantly worse. The guards became harsher and began to act in ways that many would view as cruel and unethical. For example, they would force the prisoners to urinate and

defecate in a bucket in their cell rather than taking them to the bathroom. Sometimes they even refused to let the prisoners empty the buckets, and their cells began to smell.

Before the end of the second day, one of the prisoners suffered what seemed to be an emotional breakdown, cursing, crying uncontrollably, and exhibiting extreme rage. Ultimately, the researchers decided that they had to let him leave the experiment. But doing so led to new complications, and the guards' treatment of the prisoners got even worse. The prisoner counts turned into multiple-hour ordeals. Guards even made the prisoners clean toilet bowls with their bare hands and simulate sodomy.

Finally, after only 6 days of what was originally planned as a 14-day experiment, the researchers decided to call the whole thing off. But it wasn't because they themselves realized how unethical the experiment had become. They were all fully engrossed in their roles as prison officials and had lost sight of the fact that these were all innocent college students who had volunteered to participate in a psychological study.

Rather, it was Christina Maslach, who had recently graduated with her doctorate from Stanford, who pointed out that the study had gotten completely out of hand.[11] She came to the simulated prison to interview the participants and saw the prisoners being marched to the toilet with their legs chained and with bags over their heads. She opened Zimbardo's eyes to what was going on, and he stopped the study.

The Stanford Prison Experiment had some very significant ethical problems. Notice that unlike the Milgram obedience study, the participants in the prison experiment knew what they were volunteering for, at least initially. The prisoners knew that some of their civil liberties would be temporarily removed. Likewise, the guards understood their expected role in disciplining the simulated prisoners. And everyone agreed to play their part. So what was the problem?

There were a few major ethical oversights with the experiment. First, the researchers continued the study even after it became clear that participants were suffering. Even if they hadn't anticipated all the problems that arose, once they saw prisoners going through significant emotional pain and the guards acting sadistically, they should have immediately stopped the study. Instead, they not only carefully observed all of the disturbing behavior, but they actively sought ways to keep the experiment going. That's a clear violation of respect for persons.

Another major oversight was Zimbardo's decision to get personally involved by serving as the prison superintendent. By his own admission, he started to think more and more like a real prison superintendent. He began to lose sight of the best interests of the students in the experiment, thereby violating the principle of beneficence.

11 Zimbardo ended up marrying Maslach, the whistleblower of his experiment!

SUGGESTED READINGS

Milgram, *Obedience to Authority*.

Oatley, *Our Minds, Our Selves*.

Zimbardo, *The Lucifer Effect*.

Table of Contents

LESSON 3

EXPERIMENTING ON VULNERABLE CHILDREN

Two notorious psychological studies that involved children are the Neubauer twin study and a study of stuttering that has come to be known simply as the monster study. One of the things that makes these studies particularly shocking is that they involved children, some of the most vulnerable people in our population.

THE NEUBAUER TWIN STUDY

In 1980, when Bobby Shafran arrived at Sullivan County Community College in New York to start school, and despite the fact that this was his first day and he didn't know anyone, students at the school were welcoming him back and calling him Eddy.

It turns out that Bobby looked exactly like a former Sullivan County student named Eddy Galland. And that's because they were actually identical twins who had been adopted by different families 19 years earlier! But neither of them knew they even had a biological sibling, much less a biological sibling who was an identical twin.

The story of these identical twins who randomly discovered each other after being separated from infancy generated significant interest, and a newspaper article describing the incredible reunion ran in some of the local newspapers with a picture of the long-lost brothers. And that's when things really got crazy.

Unbelievably, Bobby and Eddy actually had another identical sibling named David Kellman, who had been adopted by yet another family. And David was also completely unaware of the existence of any biological siblings. But that all changed when he picked up a newspaper and saw a picture of 2 men who looked exactly like he did. A few phone calls later and the triplets were reunited for the first time since they were babies.

The boys became overnight celebrities, appearing on talk shows and being featured in numerous magazine articles. They moved in together and opened their own restaurant, named Triplets Roumanian Steakhouse. They were even the subject of an award-

Which is more important to a person's identity and personality: nature or nurture?

winning documentary called *Three Identical Strangers*, which was released in 2018.

And as incredible as their story sounds, it turns out that they were not alone. At least 5 other pairs of identical twins at the same adoption agency had also been split up and placed in separate families. And those families also weren't told that their new child had a twin. Nevertheless, all of these children were carefully followed and repeatedly visited by researchers who tested their cognitive skills and motor performance for years after their adoption.

All of the children were actually part of a scientific study that was overseen by an eminent psychiatrist named Peter Neubauer.

The adoptions of all the children were handled by Louise Wise Services, a highly respected adoption agency for Jewish children and families in New York City that closed in 2004. The agency was advised by a prominent child psychiatrist named Viola Bernard, who

recommended that twins being put up for adoption should be placed in different families. In a recently discovered memo, she explained that she thought "early mothering would be less burdened and divided and the child's developing individuality would be facilitated."

This belief was not based on solid scientific evidence, and in fact, most developmental psychologists today would argue that the benefits of keeping siblings together far outweigh any potential costs. Nevertheless, based on Dr. Bernard's recommendation, Louise Wise Services began splitting up siblings and placing them in different families in the 1950s. Furthermore, in keeping with the closed adoption policy of the time, adopting families were not told about the biological mother or any biological siblings.[12]

Enter Dr. Peter Neubauer, a clinical professor of psychiatry at New York University who was interested in the forces that shape a person's personality and mental health. Neubauer knew about the identical siblings who were being raised in different families and saw it as an extremely rare opportunity to study the effects of nature versus nurture on human development.

The basic question is profound: What makes a person turn out the way he or she does?

Typically, it's impossible to tease apart the effects of environment (nurture) from the effects of genetics (nature). After all, we all have different DNA, and we all go through unique experiences, even if we're raised in the same household. And if both nature and nurture are varying, then there's no way to figure out which is playing the most significant role in shaping our personality and individual characteristics. The only way to distinguish the 2 influences is

12 All adoptions at the time were closed, which means the researchers were legally prohibited from telling the adopting families about the biological parents or any biological siblings. The researchers were therefore complying with New York State law by withholding that information.

to find a group of people who vary on one of those factors but don't vary on the other.

And that's exactly what identical siblings who are reared apart provide. They grow up in completely different family environments with different parents, siblings, and friends—yet their DNA is identical. If they end up having very similar characteristics and personalities as adults, then it provides strong evidence that nature is playing a dominant role. Conversely, if they end being very different as adults, it suggests that nurture is more important.

Neubauer therefore decided to follow at least 5 sets of separated identical twins, as well as the infamous triplets, as they grew up. He recruited some research assistants and instructed them to repeatedly visit the children in their homes and to administer a whole bunch of tests, all while filming the children and observing their behavior. These research assistants all knew that the children had identical siblings growing up in other households.

In fact, the researchers typically tested both twins, just on different days. However, they were under strict orders not to tell the families about the other siblings or about the true purpose of the study. Instead, the study was portrayed as an investigation of the development of adopted children. And the study continued under that pretense for decades until it was finally stopped in 1980. And even after it was stopped, the siblings still weren't informed that they had an identical sibling out in the world somewhere, nor were they told about the true nature of the study.

But some of them found out anyway. Of course, the triplets found each other, but they weren't alone. For example, another person in the study, Elyse Schein, discovered she had a twin sister when she went looking for her biological parents and contacted Louise Wise Services in 2002. She managed to track down

her long-lost sister, and the twins ended up writing a book about their experience, called *Identical Strangers*.[13]

In most cases, people who had been in the study were thrilled to be reunited with the brother or sister they never knew they had. Many of them reported developing a very strong bond with their newfound twin almost immediately. They also often discovered that they had a remarkable amount in common even though they had never met before, such as drinking the same kind of beer or studying the same subject in school.[14] These kinds of similarities suggest that our DNA plays a powerful role in shaping our personality and habits—a conclusion that is consistent with numerous other twin studies, as well as many animal studies.

Initially, most of the twins felt genuine excitement about finally connecting with a potential soulmate and were eager to learn about each other's lives and compare notes. But relatively soon, many of the study participants also began to feel bitter and violated. The more they connected with their newfound brother or sister, the more they realized what they had missed growing up. And most of them deeply regretted being separated as babies at the whim of an adoption agency.

Furthermore, at least 2 of the families reported that their children exhibited signs of significant distress immediately after the adoption, and in hindsight they wondered if this wasn't a severe form of separation anxiety. Some family members were convinced that being separated from their identical siblings at 6 months

13 This title presumably inspired the title of the *Three Identical Strangers* documentary that was made about the triplets in 2018.

14 One set of male twins who were reunited in their late 30s discovered that in many ways they had lived parallel lives. They both coached hockey and had children who played hockey and wore the same jersey number. They both got married in the same year and married women who were runners with type-A personalities. Perhaps most astonishingly, they both carried their wallets in their front pockets.

of age was very traumatic for the children and may have contributed to mental health problems that many of them experienced later in life. In fact, at least 3 of the children who were part of the study ended up committing suicide, and some of the families wonder if being separated from their siblings may have played a role.

Many of the twins also resented being misled over an extended period of time. The researchers had repeatedly visited their homes when they were growing up, and the children and their families believed that the goal was to study the development of adopted children. But all the researchers knew that the children were actually participating in a twin study whose goal was to investigate the role of nature versus nurture in human development.

The twins particularly resented never being told that they had an identical sibling. After all, the researchers interacted with each child and the child's twin many times over many years, yet they never told the children that they had a sibling, much less a twin.

And to add insult to injury, the twins and triplets who were part of the study still don't know what the ultimate scientific conclusions were. In 1990, Dr. Neubauer and the Jewish Board of Family and Children's Services donated all the records from the study to Yale University. But those records are sealed until October 25, 2065. This means that no one can access them without the Jewish Board's approval. And initially, that even included the children themselves.

However, after receiving repeated requests from some of the children involved, and after the study received significant publicity from the *Three Identical Strangers* movie, the Jewish Board agreed that children who were involved in the study should be able to have access to their own data.

As a result, the children now have access to redacted copies of their own records, although they still can't see all the other information about the study, including the final conclusions. The study's results were never thoroughly described in a formal scientific

paper, so naturally, many of the family members feel like all they went through was for nothing.[15] It didn't even help to advance science.

In light of the *Belmont* principles, this study could never be conducted today. In fact, you could make a case that it violated all 3 of the principles.

- It violated respect for persons because it failed to adequately inform families about what was actually going on. Instead, the researchers misled the children and their adoptive parents into thinking that the study was only about the development of adopted children. So, to the extent that the families gave consent, it certainly wasn't informed consent.

- The principle of beneficence, which requires that the risks of a study are warranted based on the benefits, also seems questionable in this case. After all, what were the benefits if the results are under seal and not available to advance scientific knowledge? And even if the results were available, it's not clear what conclusions you could draw from such a small number of participants anyway.

- The study also comes up lacking with regard to the principle of justice, which considers whether the research is fair. Does it treat different people equally, both in terms of sharing in the benefits of the research and in bearing the risks? In this case, the risks were borne entirely by a particularly vulnerable population, namely adopted children and their families.

15 Although Dr. Neubauer never published a complete description of the study and its results, he did write a book arguing that genetics plays a major role in who we turn out to be, entitled *Nature's Thumbprint: The New Genetics of Personality*, which makes frequent reference to the work done in the twin study in order to make the case.

THE MONSTER STUDY

Wendell Johnson was one of the most influential speech pathologists in history. In particular, his work on stuttering has had a profound effect on the field. He was a professor in the Department of Speech Pathology and Audiology at the University of Iowa and had helped to found the American Speech-Language-Hearing Foundation.

One of the reasons Dr. Johnson had an intense desire to understand stuttering and to help those who suffered from it was because he suffered from a very severe stutter as a child and young man. As he put it himself, "I became a speech pathologist because I needed one."

Outside of speech pathology circles, Dr. Johnson is probably best known for a highly criticized study that he conducted with his graduate student Mary Tudor for her master's thesis in 1939. The thesis was never published and sat in the University of Iowa library for more than 60 years. But in 2001, the university publicly apologized for that obscure thesis project. Two years later, some of the people who had been participants in the study sued the state of Iowa for damages resulting from the study. And in 2007, they were awarded around $1 million as compensation. The popular press now typically refers to it as the monster study because some of the children involved were treated in such a monstrous way.

Dr. Johnson, who started stuttering when he was 5 or 6 years old, attributed his speech problems to the fact that one of his teachers told his parents that he was starting to stutter. He believed that others' concern about his speech made him self-conscious and led him to focus obsessively on how he spoke in an effort to avoid stuttering.

And far from preventing him from stuttering, he thought that his obsessive self-consciousness about his speech is actually what made him stutter in the first

place. As he put it, "stuttering often begins not in the child's mouth, but in the parent's ear." Most theories at the time assumed that stuttering was due to a biological abnormality that was primarily genetic. Johnson's radical idea was that the diagnosis of stuttering can actually cause the disorder. He referred to this as the diagnosogenic theory of stuttering.

In 1938, Dr. Johnson recruited Tudor to carry out an experiment to test his theory. Specifically, they wanted to see if telling nonstuttering children that they are starting to stutter would lead them to start stuttering. If so, it would validate his theory and undermine biological explanations.

Tudor studied a total of 22 orphans at the Iowa Soldiers' Orphans' Home. The children ranged in age from 5 to 16 years old, and none of them knew what the study was about. In fact, they probably didn't know they were in a study at all. They just thought they were getting therapy to help them speak better.

Ten of the 22 kids were children whom the teachers at the orphanage identified as already being stutterers. For these children, Johnson wanted to determine if giving them positive feedback about their speech might reduce their stuttering. Five of these stutterers received positive feedback about their speech while the other 5 didn't.

Johnson and Tudor also recruited 12 other children who were not considered to be stutterers. Half of these children were also randomly assigned to a positive-feedback group. The kids in this group received feedback like the following: "Do you enjoy speaking? You speak very well. Your speech is of very good quality. Speak whenever you have an opportunity. You have the earmarks of a fine speaker."

But the other nonstutterers were assigned to a negative-feedback group, and that's where the most significant ethical concerns arise. Specifically, these 6 children, all of whom were judged to speak normally, were given negative feedback about their speech to determine if it would lead to problems with speech fluency. Tudor worked with each of these children every few weeks for about 45 minutes at a time.

She told them things like the following: "The staff has come to the conclusion that you have a great deal of trouble with your speech. The type of interruptions which you have are very undesirable. These interruptions indicate stuttering. You have many of the symptoms of a child who is beginning to stutter. In fact, you are beginning to stutter. You must try to stop yourself immediately."

This kind of negative feedback made the children very self-conscious about their speech, and most of them became reluctant to talk. About one of the children in this group, Tudor wrote:

> It was very difficult to get her to speak although she spoke very freely the month before. She spoke slowly and very distinctly, saying one word at a time. I asked her why she didn't want to talk. She didn't answer. Then I asked her if she was afraid of something. She nodded her head. "What are you afraid of?" After some time she said, "Afraid I might stutter."

Another child in the negative-feedback group became withdrawn and ended up running away from the orphanage a few years later,

although it's hard to know if that was related to the study or not. Decades later, when she was reached by phone, she said, "I couldn't never tell my husband about it. It just ruined my life."

Although the kids in this group became very self-conscious and reluctant to speak, when they did speak, they didn't show much evidence of being more prone to stuttering, so the thesis didn't provide much support for Johnson's diagnosogenic theory of stuttering.

The ethical concerns with this study are pretty obvious.

- The risk of inducing psychological damage seems very real. And one of the primary goals of the study was to try to induce stuttering in children. The study therefore clearly violates the *Belmont* principle of beneficence—that is, minimizing risks while maximizing benefits.
- Respect for persons was also clearly violated. The children were not given appropriate respect as valuable human beings, and their extreme vulnerability was not given appropriate consideration. Not only were they children, but they were orphans who didn't have parents to watch out for them.
- The very high risks of the study were borne entirely by the particularly vulnerable population of orphaned children. That's clearly unfair and is an obvious violation of the *Belmont* principle of justice.

SUGGESTED READINGS

Neubauer and Neubauer, *Nature's Thumbprint*.

Schein and Bernstein, *Identical Strangers*.

Tudor, "An Experimental Study of the Effect of Evaluative Labeling of Speech Fluency."

Wright, *Twins*.

Table of Contents

LESSON 4

TESTING PSYCHOCHEMICAL WEAPONS

This lesson focuses on shocking studies that were conducted by government organizations like the military and the CIA—organizations that are supposed to serve the public, not harm it. In particular, it focuses on 3 men who were tied together by a shocking secret: They had been intentionally exposed to powerful psychoactive drugs as part of confidential experiments that were investigating the potential use of these drugs as mind control weapons.

JAMES STANLEY

In the 1950s and 1960s, the US military and the CIA were both deeply involved in testing the effects of a variety of different psychoactive chemicals, including LSD, mescaline, psychedelic mushrooms, PCP, ecstasy, and marijuana. And they were conducting these tests on human beings, most of whom had no idea what drugs they were being exposed to. In fact, many of the human subjects weren't even aware that they were being administered a drug at all.

Some of these tests took place in Maryland at the Edgewood Arsenal facility, where L. Wilson Greene was the scientific director. He envisioned a new type of war that did not involve killing or destroying property.[16] Essentially, Greene thought it might be possible to win battles by drugging the enemy using chemical gases. But rather than using deadly chemicals, he wanted to use drugs that temporarily disrupted the enemy's ability to think and act rationally.

So Edgewood started a research program to investigate whether psychoactive drugs were suitable to use as weapons in psychochemical warfare. And that meant Edgewood scientists needed to understand how human beings reacted to these drugs so that they could determine whether the drugs would be effective on a battlefield. They therefore began to recruit soldiers through the Medical Research Volunteer Program, which offered soldiers the opportunity to be transferred to Edgewood Arsenal temporarily and to participate in their research.

16 As Greene put it, "Throughout recorded history, wars have been characterized by death, human misery, and the destruction of property; each major conflict being more catastrophic than the one preceding it. I am convinced that it is possible, by means of the techniques of psychochemical warfare, to conquer an enemy without the wholesale killing of his people or the mass destruction of his property."

Advertisements went up around the country promoting the program, and thousands of soldiers volunteered. In addition to helping out their country, being transferred to Edgewood meant they could get out of their normal duties. But the volunteers didn't really understand what the experiments at Edgewood would involve.

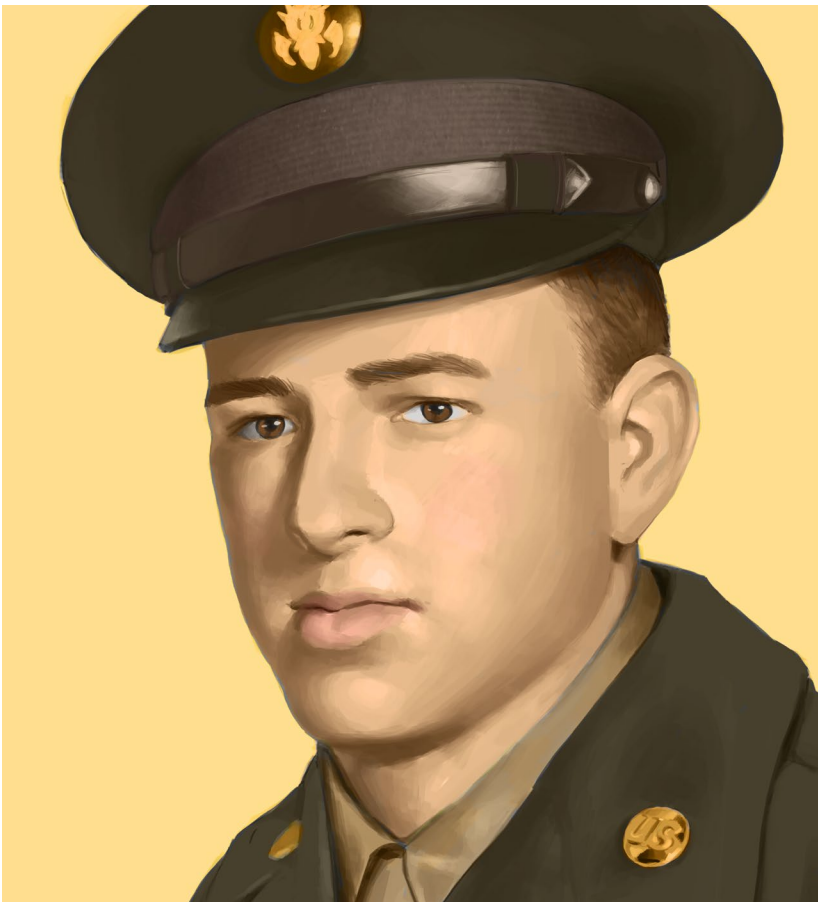
The scientists began testing the effects of a wide variety of chemicals. One of the first drugs they tried was PCP, also known as angel dust. Some soldiers were given PCP and asked to run an obstacle course. Others had the drug surreptitiously put into a drink to test the effects. But the researchers soon discovered that the effects of PCP could be more serious and last longer than they hoped. For example, one soldier had to be hospitalized for 6 weeks because he experienced a very severe paranoid reaction that lasted long after the drug was out of his system.

The drug that showed perhaps the most promise as a psychochemical weapon was called BZ, which can produce a variety of psychological symptoms, including disorientation, agitation, tremor, and stupor. In extreme cases, it can even cause seizures and comatose states. The scientists at Edgewood conducted a major field test to determine whether clouds of BZ could be used effectively against soldiers who were more than 1/4 mile away. This test was code-named Project DORK.[17]

Other soldiers were repeatedly exposed to LSD to examine its effects. One of those soldiers was James Stanley, who arrived at Edgewood in 1958 expecting to test gas masks and protective clothing as part of a study he signed up for because of an advertisement he had seen. Instead, once a week, he was given a glass of water to drink. And what he didn't know at the time was that the water was laced with LSD.

17 A video of the event, titled *Cloud of Confusion*, was produced and might still be on YouTube.

And that's when he began to hallucinate and develop violent reactions, including assaulting 2 guards. And unfortunately, he continued to experience hallucinations, flashbacks, blackouts, and fits of rage even after he left Edgewood. And ultimately those symptoms led to significant personal problems, including his divorce and estrangement from his family.



Stanley was never told that he had been given LSD. In fact, it wasn't until 17 years later that he found out. And ironically, he learned the truth when he received a letter asking him to participate in another army-sponsored research project. Specifically, the letter explained that the army wanted to study the long-term effects of LSD on soldiers who had been exposed to it at Edgewood! Only then did Stanley realize what had really happened to him back in 1958.

That realization led him to file a lawsuit alleging negligence in the administration and monitoring of the project. His case made it all the way to the Supreme Court, but in 1987, the court ruled against him in a close 5-to-4 decision. The majority decision argued that soldiers cannot sue the military for injuries that are suffered as part of their military service.

But after learning about Stanley's case, a Florida congressman named Harry Johnston sponsored a bill to compensate Stanley for his ordeal. And in 1996, nearly 40 years after the Edgewood study, the Palm Beach District Court ruled that the experiment had altered Stanley's personality and had led to the decline of his military career and his first marriage. They awarded him $400,577 in damages.

FRANK OLSON

The experiments at Edgewood Arsenal had some very serious ethical problems, but it turns out that there were other government-sponsored studies going on around the same time that were arguably even worse. Many of these experiments were conducted as part of a top-secret CIA project called MK-Ultra, which started in 1953 at the height of the Cold War and continued for roughly 20 years. The CIA was concerned that the Soviet Union and their allies were developing

techniques to brainwash US prisoners of war in Korea. And they wanted to make sure they didn't get left behind.

It was in this context that the director of the CIA at the time, Allan Dulles, authorized the MK-Ultra project. The goal was to develop ways to control human behavior, and psychedelic drugs like LSD seemed like one of the most promising approaches.

The MK-Ultra project sponsored more than 150 experiments that were conducted all over North America. In one particularly egregious experiment, the CIA hired sex workers to lure clients into safe houses containing one-way mirrors behind which CIA scientists could observe. The sex workers would then give the unsuspecting clients psychedelic drugs like LSD, typically without their knowledge, and the scientists would observe the effect of the drug on their behavior. The victims were unlikely to report the incident to law enforcement or anyone else because they wouldn't want anyone to know that they had hired a sex worker.

Frank Olson was working for the CIA on the MK-Ultra project when, on November 18, 1953, he and a group of fellow CIA employees drove from Fort

Detrick in Frederick, Maryland, to Deep Creek Lake on the far western edge of the state for a retreat. At some point during the retreat, Olson drank from a bottle of Cointreau that had been spiked with LSD.

Soon thereafter, Olson is said to have experienced extreme paranoia and to have

suffered some kind of nervous breakdown. The CIA sent him to New York City to be seen by Dr. Harold Abramson, who wasn't a psychiatrist but an allergist and pediatrician who had been working with the CIA on their drug research. Specifically, Dr. Abramson was very interested in the use of LSD in the treatment of mental illness.[18]

While Olson was in New York, he fell to his death from the 13th floor of the Hotel Statler, across the street from Penn Station. He was staying in the room with a fellow employee when it happened. His boss went to Olson's home in Frederick to break the news to his wife and 3 young children. He told them that Olson had been in an accident and had either fallen or jumped to his death.

Olson's family had no idea that he had been exposed to LSD a few days before his death. And they probably would never have found out if it hadn't been for a journalist named Seymour Hersh, who, in December 1974—more than 20 years after Olson's death—wrote a lengthy article for *The New York Times* that described numerous CIA abuses.

Less than 2 weeks later, President Ford established the President's Commission on CIA Activities within the United States, also known as the Rockefeller Commission, because it was led by then–Vice President Nelson Rockefeller. The commission was charged with investigating activities of the CIA and other intelligence agencies. The US Senate and House quickly followed suit by creating the Church Committee and the Pike Committee, both of which had similar mandates.

Although the CIA destroyed many of the documents related to their more questionable projects, these investigations uncovered the MK-Ultra project and the large number of CIA-sponsored experiments in which people had been exposed to psychedelic drugs without their consent. They also reported the fact that Olson had been exposed to LSD without his knowledge before he died and suggested that

18 In fact, Dr. Abramson edited a book on the topic in 1967.

his death had been a suicide due to an extreme reaction to the LSD. In response, the government gave Olson's family $750,000 as a settlement, and both President Ford and CIA director William Colby issued them formal apologies.

But the story didn't end there.

Olson's wife died in 1993, and the children decided to have their father's remains exhumed to be buried with his wife. They also decided to have another, independent autopsy conducted. The new autopsy discovered a large hematoma on the left side of Olson's head that had not been reported in the original autopsy. It also reported a large injury on his chest. Most of the autopsy team thought that these injuries had occurred in the hotel room before the fall.

Also, the original autopsy had reported cuts and abrasions on Olson's body, but the doctors who conducted the second autopsy saw no such injuries. The doctor who led the second autopsy said that the evidence was "rankly and starkly suggestive of homicide." The family now firmly believes that the CIA had Olson murdered because they viewed him as a serious security risk who could expose what was actually happening with the MK-Ultra project.

Finally, in 2012, nearly 60 years after Olson's death, his sons filed a lawsuit seeking compensatory damages and requesting access to CIA documents related to their father's death that they claimed were being withheld.[19] The case was dismissed, partly because the family had agreed to a settlement back in 1976. But in his decision, the judge wrote this:

> While the court must limit its analysis to the four corners of the complaint, the skeptical reader may wish to know that the public record supports many of the allegations, farfetched as they may sound.

19 If you're interested in learning more about the Olson case, Netflix produced a miniseries about it called *Wormwood* in 2017.

HAROLD BLAUER

Harold Blauer was a tennis pro who gave lessons at the Hudson River Club in Manhattan. He was married with 2 daughters, but in 1952, his marriage fell apart and he fell into a depression. He decided to check himself into Bellevue Hospital and was subsequently transferred to the New York State Psychiatric Institute, where he began therapy.

Blauer remained in the institute for about a month before he fell into a deep coma and died. His ex-wife was told that he had suffered an atypical reaction to a drug.

It turns out that the institute had a secret agreement with the US Army Chemical Corps to test mescaline-based hallucinogenic drugs on patients. The goal of the research was to "provide a firmer basis for the utilization of psycho-chemical agents both for offensive use as sabotage weapons and for protection against them."

As part of this research, Blauer received injections of different mescaline derivatives on 5 different days in December 1952 and January 1953. He was very apprehensive about the injections and once even lied about having a cold to avoid getting the shot but was given the injection anyway. Afterward, he explicitly told the doctors that he didn't want to get any more shots, but they threatened to send him to one of the less pleasant mental asylums if he withdrew. So he continued.

On January 8, he got his fifth injection. And this time, he got a dose that was 16 times larger than the first injection. The last few entries in the study notes are chilling:

```
11:12 Increasing
      restlessness.
      Intermittently
      generalized rigidity.

11:17 No longer talking.
      Lapsing into coma.
      Still restless.

11:30 Becoming cyanotic.
      Respiration rapid
      and stertorous.

11:45 Quiet. Deep coma.
```

Thirty minutes later, Blauer was pronounced dead.

The report to the medical examiner implied that the drug was given for therapeutic reasons and failed to mention the experiment or its purpose. The medical examiner was also asked to keep his report confidential.

The family didn't find out what had really happened until the Rockefeller Commission, the Church Committee, and the Pike Committee raised public awareness about these kinds of unethical experiments in 1975. The army finally admitted the truth about what had happened, and 13 years later, Blauer's estate was awarded $702,000 in damages.

ETHICAL QUESTIONS AND CONCERNS

The cases of James Stanley, Frank Olson, and Harold Blauer obviously raise a host of ethical questions and concerns.

First, they highlight the critical need to obtain fully informed consent from all research participants before they get

involved in the research. This is a core component of exhibiting appropriate respect for persons—one of the *Belmont* principles—and it was obviously violated in all of these cases.

- Stanley and the thousands of other soldiers who were studied at Edgewood Arsenal were often never told what drugs they were being exposed to, nor were they told what kinds of symptoms to expect or the potential risks. Many of the soldiers who volunteered were under the impression that they were going to be testing out gas masks or clothing, not being exposed to psychoactive drugs. If they had truly known what they were getting into, many of them likely would not have volunteered in the first place. Also, it's almost impossible to be sure that soldiers like Stanley are participating in a study of their own free will, as they're trained to obey authority without question or hesitation.
- Likewise, most of the people who were drugged as part of the CIA's MK-Ultra project never gave their consent. In fact, many of them were completely unaware that they were being drugged. Olson certainly didn't expect to get a dose of LSD when he drank from that bottle of Cointreau at Deep Creek Lake.
- Although Blauer agreed to be treated for his depression, he never gave his consent to be given large doses of mescaline derivatives simply to observe the effects. In fact, he explicitly asked for the injections to stop but was threatened to obtain his cooperation. That's a pretty obvious example of coercion, which is antithetical to the principle of respect for persons.

In addition, all of these cases very clearly violate the *Belmont* principle of beneficence. Remember, according to modern-day standards, psychological studies should always strive to maximize

benefits and minimize risks. And at the very least, they should do no harm. But all the studies in this lesson clearly did significant harm to some of their participants. Soldiers at Edgewood experienced hallucinations, paranoid delusions, and fits of rage, and both Olson and Blauer died as a result of the studies they were in.

Furthermore, none of these studies followed rigorous scientific methods, presumably because they were overseen by military personnel or government officials rather than by highly trained scientists. As a result, it's hard to draw strong scientific inferences from any of these studies. So in addition to exposing participants to significant risks, the scientific benefits of the studies were also questionable.

Finally, consider the *Belmont* principle of justice. Was the research fair, or did it unfairly expose a subset of people to risk? The studies at Edgewood pretty clearly violated the justice principle because they targeted soldiers—and typically lower-ranking soldiers, a vulnerable population in the military who don't enjoy the privileges that higher-ranking officers do. The same concern applies to experiments on psychiatric patients like Blauer who were seeking help for their illness and instead were subjected to untested and ultimately very harmful procedures.

SUGGESTED READINGS

Albarelli, *A Terrible Mistake*.

Khatchadourian, "Operation Delirium."

Moreno, *Undue Risk*.

Regis, *The Biology of Doom*.

Table of Contents

LESSON 5

ASSIGNING GENDER AND SPYING ON SEX

This lesson features a few notorious studies that investigated sexual behavior and gender identity. When studying sex and sexual behavior—some of the most private and sensitive aspects of being human—scientists need to be extra careful to protect the rights of participants and behave as ethically as possible. Unfortunately, not all scientists have done that.

THE TEAROOM TRADE STUDY

The Tearoom Trade study was conducted by Laud Humphreys for his PhD dissertation at Washington University in St. Louis in the 1960s. This study generated substantial controversy after it was published, and the ethics of the study continue to be debated to this day. But many people also see Humphreys's research as groundbreaking and view Humphreys as a hero.

One reason the study was so controversial was that it investigated a very sensitive subject—specifically, the practice of men going to public restrooms to have sex with other men. Humphreys actually visited public restrooms, observed men engaging in sex, and wrote up his findings. In addition, he conducted this study in the 1960s, when such behavior was not only highly stigmatized, but was actually illegal and could lead to arrest and even imprisonment.

Humphreys wanted to understand who these men were who engaged in these activities and why they did so despite the substantial risks. He therefore spent a large amount of time in these "tearooms," as they were often called, and made careful observations about what went on. He also interviewed many of the men who participated and learned as much as he could about them, including their careers and personal lives outside the tearooms. What he found was startling.

One surprising finding was the sheer volume of sex that took place in the most active of

these restrooms. For example, on one particularly busy day, Humphreys observed 20 sex acts in a single hour. He also regularly saw men waiting in line to participate.

A second surprising finding was that although the participants all obviously wanted sex, they did not want intimacy. In fact, they didn't want any social interaction at all. Humphreys found that in most of these encounters, nothing was said.[20] Humphreys inferred that the men who engage in this kind of activity want it to be fast, impersonal, and anonymous.

But by far the most surprising finding from the study was about the type of men who frequented the tearooms. They came from all walks of life and all social classes. They included businessmen, gas station attendants, physicians, salesmen, and even a priest.

Only 14% of them were openly gay. Another 24% were homosexual but in the closet. And 38% did not consider themselves homosexual or even bisexual. Instead, they identified as heterosexual and viewed the tearooms simply as a quick and easy way to get sexual satisfaction.

In fact, more than 50% of the men were married to women! Their wives typically had no idea about this aspect of their husbands' lives, and the men tried their best to keep it that way. By all outward appearances, these men were entirely heterosexual.

The Tearoom Trade study therefore demonstrated a significant inconsistency between the public and private lives of most of these men. Many of them were actually quite conservative in public life, both politically and socially. Some were active members of churches and other religious organizations. Many publicly denounced homosexuality as sinful and were staunch opponents of gay rights.

20 The single word "thanks" was occasionally spoken at the end of an encounter, but even that only happened in a minority of interactions.

It would therefore be natural to view these men as hypocrites. But that's not the way Humphreys saw them. Based on his interactions with them, he believed that many of them were deeply ashamed of their behavior in the tearooms and tried to make up for it by trying to appear extra righteous in their public life.

Humphreys's also concluded that the behavior he observed in these tearooms did not pose any threat to people in the local community. Men who entered these restrooms simply to use the facilities were never approached for sex or harassed in any way. And the men who did participate were often among the most law-abiding citizens in the entire community in public life. Based on Humphreys's findings, many police departments began to overlook what was happening in the tearooms and devote more of their resources to stopping crimes that they viewed as more dangerous.

The Tearoom Trade study also had a profound influence on the field of sociology. In particular, the recognition that public and private behavior can be radically different significantly influenced theories of human behavior in the field. The work also led sociologists to adopt methodologies that tried to capture private behavior while recognizing that it might be very different from public behavior. Humphreys's research also inspired a generation of sociologists to study sexual behavior using similar so-called ethnographic methods.

But as influential as Humphreys and his research have been, his methods have come under significant criticism for several reasons. First, many of the men had no idea that their behavior was being observed as part of a scientific study, so they obviously never gave their consent. And that seems like a clear violation of *The Belmont Report*'s respect for persons principle.

You might wonder why these men let Humphreys observe their behavior in the first place, especially given that many of them were very concerned about being exposed. They did so because Humphreys misled them about his true intent.

Specifically, he pretended to be a voyeur who would serve as a lookout in exchange for the opportunity to observe the proceedings. He would watch the door and provide a warning signal if the police were in the vicinity or if someone was approaching the restroom. He found that this was standard practice in many tearooms, so by adopting this role, he could avoid suspicion.

But Humphreys wanted to do more than just observe the behavior in the tearooms. He wanted to learn more about the men who participated. And that led him to adopt a strategy that most people consider the most unethical aspect of the study.

He wrote down the license plates of the cars that the tearoom participants drove and, with the help of some friends in the police force, figured out who the men were and where they lived. He then went to their homes and interviewed them under the pretense of conducting survey research. He asked them about their family and their career, among many other topics. And that's when he discovered that many of the men were married and strongly identified as heterosexual.

Because Humphreys was worried that the men he interviewed would recognize him, he waited about a year before conducting the interviews and changed his appearance. None of the men reported recognizing him.

These tactics were roundly criticized. Some of the faculty members in Humphreys's department sought to have his doctoral degree rescinded. The study also led to outrage in the popular press.

On the other hand, some people note the importance of the study and defend its ethics. They point out that the tearooms were public spaces and that observing public behavior does not typically require getting a person's consent. And that's true to this

day.[21] But observing sexual behavior in a public restroom that is being monitored by a lookout seems pretty different from watching people in a mall or at a football game.

Humphreys's defenders also point out that he tried very hard to respect the privacy of the men he studied. He scrupulously protected their identities and by all accounts did his best to make sure that nothing he did would lead to them being harmed in any way. And as far as we know, none of the tearoom participants did experience any negative effects as a result of the study.

Humphreys's writing suggests that he strongly identified with many of his research subjects and related well with them. In fact, Humphreys himself came out as a homosexual years after the study was conducted.

THE JOHN/JOAN CASE

Bruce Reimer was born on August 22, 1965, in Winnipeg, Manitoba, along with his identical twin brother, Brian. The boys were both perfectly healthy.

But the following spring, their parents noticed that their foreskins were closing and that this was causing problems with urinating. This is a fairly common condition called phimosis, which usually resolves by age 3 even without treatment. But in 1966,

21 Current regulations overseeing research on human beings do not require scientists to obtain consent from people whose behavior is being observed in a public space.

circumcision was a common treatment for phimosis, and that's what the Reimers' doctor recommended.

So on April 27, 1966, Bruce and Brian were scheduled for circumcisions. Bruce went first, but unfortunately, the procedure did not go as planned. The urologist used a technique called electrocauterization that seals blood vessels while incisions are made. But Bruce's penis was essentially burned off during the procedure and could not be repaired. The doctors decided not to perform the procedure on his brother, whose phimosis resolved on its own without treatment.

Bruce was seen by numerous specialists, but none offered much encouragement.[22] Bruce's parents had given up hope that their son would ever be able to have children or a normal sex life. But about 8 months later, they saw an interview on TV that changed their lives, and Bruce's future, forever.

The interview was with Dr. John Money, a psychologist at Johns Hopkins who worked with people who had been born with ambiguous sexual organs—for example, girls who were born without a vaginal opening or boys with a scrotum that was divided like labia and an extremely small penis. These intersex people were sometimes raised as girls despite having male DNA and were sometimes raised as boys despite having female DNA. And in studying these cases, Money concluded that it didn't really matter which gender they were raised as. They seemed to do equally well psychologically either way.

Money came to believe that people are born gender-neutral and that they will naturally adopt whichever gender matches their physical attributes and their upbringing. He thought that if they have penises and people treat them like boys, then they will naturally identify as boys. If they have vaginas and people treat them like girls, then they will

22 As one of them put it, "Insofar as the future outlook is concerned, restoration of the penis as a functional organ is out of the question."

identify as girls. And if they're intersex and have ambiguous sex organs, then they will identify with whichever gender other people attribute to them, although their ambiguous sex organs could potentially lead to psychological conflict.

Based on this hypothesis, Dr. Money recommended that intersex babies receive sexual reassignment surgery at an early age and that they be treated as the reassigned gender by everyone at all times. The goal was to remove any doubts or conflicts in the child's mind about his or her gender.

The Reimers decided to seek out Dr. Money's advice about Bruce's situation. They wrote to him and described what had happened to their son. He wrote back soon thereafter. And unlike all the other physicians that they had previously consulted, Dr. Money offered hope. He told them that he was optimistic that he and the doctors at Johns Hopkins would be able to help.

So the Reimers brought the twins to meet with Dr. Money in person. He recommended that Bruce undergo surgery. And because a functional penis could not be constructed, Money suggested that Bruce's genitalia be made female and that he be raised as a girl rather than as a boy. According to Money's theory, the child was still gender-neutral at this age and would readily accept his identity as a girl if his sex organs were female and if those around him treated him like a girl.

Dr. Money undoubtedly believed that these steps were in the child's best interest. He was hoping that once Bruce received the surgery and was raised as a girl, he would become a girl for all intents and purposes. And as a girl, she would not have to face the psychological trauma that might be associated with growing up as a boy without a penis. Money also hoped that after she reached adulthood, she would be able to enjoy a more typical sex life even if she wouldn't be able to bear children.

But the case also provided an extremely rare and very exciting opportunity for Money to prove that his theories about gender identity

were right. After all, it could provide the most convincing evidence to date in favor of Money's theory that gender identity derived more from experience than from DNA.

Furthermore, as a scientific case study, the case had a built-in control condition—namely, the twin brother, Brian. Because he was an identical twin, Brian shared the same DNA as Bruce. So if Brian adopted a male identity while his sibling adopted a female identity, it would very strongly suggest that a person's gender was malleable and was determined by their upbringing rather than by their genes, just as Money's theory claimed.

The timing of the case couldn't have been better for Money. When the Reimer twins first came to Johns Hopkins, his theory was beginning to be questioned in the scientific literature. The Reimer case—better known in the popular press as the John/Joan case—had the potential to vindicate Money and take his scientific career to another level.

And it seems at least possible that these considerations could have clouded his judgment and biased the advice that he gave to the Reimers. At the very least, Money had a vested interest in seeing the Reimer child undergo sex reassignment and be raised as a girl. And that's what he strongly recommended.

After a few months of thinking it over, that's what the Reimers decided to do.

So on July 3, 1967, when Bruce was 22 months old, the surgical team at Hopkins removed his testicles and did their best to construct female genitalia. His parents were told to raise him as a girl and not to tell him or anyone else about the surgery in an attempt to avoid psychological conflicts and emotional trauma.

The Reimers began referring to their child as Brenda rather than Bruce and began using the female pronouns *she* and *her*. And in every other way, they tried their best to raise Brenda as a girl, including

dressing her in feminine clothes and giving her traditionally female toys to play with.

Dr. Money also asked the Reimers to bring the twins back to Johns Hopkins once a year for counseling, which they dutifully did. These visits gave Money an opportunity to interview the children without their parents present and assess the extent to which Brenda identified as a girl. Money also used these private sessions to try to reinforce her female identity.

Some of the things that Money asked about and asked the children to do were ethically questionable, to say the least. He asked both children very detailed questions about their sexual fantasies and whether they fantasized about having sex with men or women. He showed them pictures of naked children and graphic pornographic pictures of adults engaging in sex. He also forced the twins to undress and examine each other's genitals in front of him, sometimes with other observers in the room. He even asked them to simulate sex with each other, with Brian in the male role and Brenda in the female role.

Money used his observations from these sessions as evidence in scientific publications. In 1972, he published a book called *Man & Woman, Boy & Girl: Gender Identity from Conception to Maturity* that included a description of the Reimer case. And according to the book, the case clearly supported Money's theories. Brenda was described as identifying as female, having typical feminine interests, and behaving like most other girls.

Dr. Money's goal was to do everything he could to reinforce gender-specific ideas about sexual anatomy and sexual behavior for Brenda. He wanted to be sure that Brenda thought of herself as a female in every way possible.

Three years later, when the twins were 9 years old, he published another report in which he claimed that Brenda was completely accepted as a girl by those around her and that no one suspected that she had been born a boy.[23] He also published a book called *Sexual Signatures: On Being a Man or a Woman* that described the Reimer case as "dramatic proof that the gender-identity option is open at birth for normal infants."

These publications cemented Money's position as the world expert on gender identity. And the unqualified success of the Reimer case also had a major impact on medical practice. In particular, infant sex reassignment surgery, which had been rare and had been performed almost exclusively at Johns Hopkins, began to be practiced all over the world.

Unfortunately, the truth about Brenda Reimer isn't nearly as rosy as Money described in his papers and books. As her brother Brian put it, "I recognized Brenda as my sister, but she never, ever acted the part." He added:

> There was nothing feminine about Brenda. … She walked like a guy …. She talked about guy things …. We both wanted to play with guys, build forts and have snowball fights and play army.

At school, Brenda was the target of constant ridicule and derision by the other children because of her masculine mannerisms and behavior. Her teachers and other adults who worked with Brenda also noticed that she was different from the other girls.

And even though no one had ever told her that she had been born a boy, Brenda became increasingly uncomfortable with being treated as a girl. She became convinced that on the inside she was really

23 As Dr. Money put it, "No one knows. Nor would they ever conjecture. Her behavior is so normally that of an active little girl, and so clearly different by contrast from the boyish ways of her twin brother, that it offers nothing to stimulate one's conjectures."

a boy and became more and more frustrated with having to comply with feminine social norms. By the time Brenda was 14, she simply refused to continue the act. She stopped wearing feminine clothing and sometimes would urinate standing up.

The situation came to a head when she went to see her endocrinologist and refused to submit to a breast exam. Seeing the obvious struggles that she was going through, the doctor concluded that Brenda should be told the truth. And soon thereafter, her father did just that.

After he explained everything that had happened to her, Brenda was obviously stunned and angry. But most of all, she was relieved. She realized that feeling like a boy wasn't abnormal in any way because, in reality, she *was* a boy.

Almost immediately, she decided to transition back to being male. He adopted the name David and underwent more sex reassignment surgeries, but this time to make him more male rather than more female. And he lived the rest of his life as a man. He even ended up getting married and becoming the father of children that his wife had from previous relationships.

Unfortunately, this tragic story also has a tragic ending. In 2002, Brian died from an overdose of antidepressants. And 2 years later, David died by suicide soon after his wife told him she wanted to separate.

SUGGESTED READINGS

Colapinto, *As Nature Made Him*.

Humphreys, *Tearoom Trade*.

Money and Ehrhardt, *Man & Woman, Boy & Girl*.

Table of Contents

LESSON 6

CURRENT AND FUTURE ETHICAL CHALLENGES

Although regulations have been put in place that would prevent countless shocking studies from the past from being conducted today, there are plenty of ethical dilemmas that still arise in psychological research. This lesson addresses some of those dilemmas and considers what can be done to improve psychological research in the future.

DIGITAL DATA AND PRIVACY

In the digital world we live in, information about basically every aspect of human life is being captured every second of every day from hundreds of millions of human beings worldwide. And the availability of these giant data sets creates exciting opportunities for psychological research.

- These data sets provide an opportunity to observe human behavior "in the wild," meaning you get to see how people behave when they don't think their behavior is being studied. Obviously, people aren't always honest on Facebook and Twitter, but these data sets also include lots of fairly objective data, such as amount of physical activity and GPS locations.

- Another advantage is the enormous size and diversity of the sample. When you analyze data from millions of people from all over the world, the results you obtain are much more likely to be reliable and generalizable.

- Furthermore, this kind of research can be extremely cost-effective. After all, the data has already been collected. You therefore don't have to pay an army of research assistants to recruit participants and administer surveys.

On the other hand, this kind of research also poses new ethical challenges that the field is only beginning to grapple with. In particular, under what circumstances is it appropriate to analyze this kind of data?

Google processes something like 70,000 web search requests every second, providing insight into people's interests and thoughts. That corresponds to more than 2.2 trillion searches every year.

One view is that as long as the data is not personally identifiable, then it's fair game. Current regulations don't actually consider the analysis of such secondary data sets to be human subject research if those data sets don't include personally identifiable information like names, social security numbers, and birth dates. So as long as that kind of information is stripped, each person would just be a single anonymous data point in a vast sea of millions of others and no one could figure out who each person is. Or could they?

When you have a data set that includes lots of different variables, it is actually possible to figure out who individual people are. For example, if you regularly visit the websites for your local school district or city government, then someone could probably figure out the general area where you live. Likewise, your general age and your gender could probably be deciphered from the types of products you buy. And with a few more variables, the possibilities could be narrowed down pretty quickly. And if you throw GPS data into the mix, then the identification becomes almost trivially easy.

Another viewpoint is that researchers should only be able to analyze data that is publicly available. For example, depending on a user's privacy settings, tweets and Facebook posts may be out there for the entire world to see. And many people would argue that such information is therefore also fair game for researchers to analyze.

But one potential concern is that many users don't actually understand the privacy settings

that Twitter, Facebook, and other social media platforms have. Some of those people might have actually preferred to make their account private if they had really understood the implications of being public and had known how to change the settings. But there's no way to know who those people are.

The strictest viewpoint would be to require that scientists obtain explicit informed consent from every individual whose data they plan to include in their analyses. That approach would obviously provide the greatest protection to the research participants, but it would also make it impossible to carry out many of the studies that researchers would like to conduct.

These issues will undoubtedly continue to be debated in the future.

INTENTIONAL RESEARCH MISCONDUCT

Another ethical dilemma in modern-day psychological research involves intentional research misconduct by unethical scientists. And the psychologist who is perhaps most notorious for such misconduct is a Dutch social psychologist named Diederik Stapel.

In 2011, Stapel published a highly publicized article in *Science* that described an experiment providing evidence that people are more racist in messy environments than they are in clean environments. Based on their results, Stapel and his coauthor argued that people use racist stereotyping as a kind of mental cleaning device when they are exposed to chaos in their environment.

Science is arguably the most prestigious scientific journal in the world. It only publishes

a tiny fraction of all the papers that scientists send in for consideration, so the journal's reviewers have to be convinced that a paper makes a truly groundbreaking contribution for it to be published. And Stapel's paper passed that bar.

But it turns out that the experiment he described in that paper had never actually been done. He made up data that were consistent with his racism-as-mental-cleaning hypothesis and then wrote the article based on the fabricated data.

Unfortunately, it wasn't the first time Stapel had done this. In fact, 58 of his articles had to be retracted from the scientific literature based on concerns over the validity of his data. And virtually all of these papers were written with other scientists, including Stapel's own graduate students. You might wonder how that could possibly happen. Wouldn't his coauthors see what was going on?

To avoid this, Stapel would design a fictional experiment and then fabricate a data set with results that were consistent with a particular hypothesis. He then contacted social psychologists whose work was related to the fictional experiment and told them that he had a data set that had not yet been analyzed. He asked if they would be interested in collaborating to analyze the data and write it up for publication. And many of these scientists said yes.

This house of cards came tumbling down in 2011, soon after the paper in *Science* was published. Three of Stapel's junior colleagues had become

The case of Diederik Stapel is extreme, but he's not alone. For example, in 2011, the extremely prominent Harvard psychologist Marc Hauser resigned after an investigation found evidence that he had fabricated and falsified data in numerous studies investigating animal cognition.

suspicious about many of his results. All of his experiments seemed to work, and the data were always extremely convincing. That's just not how science works in practice. Lots of plausible, intriguing hypotheses turn out to be wrong. And experimental data is often messy and ambiguous.

So these colleagues began covertly observing him. And that's when they discovered the truth. For example, after loading up his car with questionnaires that he said he was going to administer to students in a nearby town, they saw him drive off and dump the questionnaires into a trash bin. Nevertheless, a few weeks later, another beautiful data set showed up confirming whatever hypothesis the questionnaires were designed to test.

Once these colleagues were convinced, they reported him. Panels were convened to investigate the charges, and the truth came out—and Stapel's story made headlines. He lost his job and any chance at continuing his research career. The official reports confirmed that his colleagues and students were completely unaware of the fraud and cleared them of any wrongdoing. Nevertheless, many of their papers had to be retracted, so the ordeal significantly damaged their careers.[24]

24 Stapel's students were hit particularly hard. Imagine trying to get a job as a scientist when most of the experiments you worked on were never actually performed.

PSYCHOLOGY'S REPLICATION CRISIS

There have been a few recent psychological studies that are shocking not because they were unethical, but because they suggest that many of the published findings in the psychological literature are not as reliable as once thought.

One of these studies was published in August 2015 by the Open Science Collaboration, which consisted of 270 scientists from all over the world who agreed to try to replicate studies that had been published in 3 very respected psychology journals in 2008.[25] The teams managed to rerun 100 different psychological studies, and they did their best to run those studies in the same way they had originally been run, including using the original materials if they were available and consulting with some of the original authors to make sure their study was appropriately designed.

The results were shocking in a few ways. First, the size of the effects observed in the replication attempts were much smaller than the effects that had been reported in the original studies. In fact, on average, the effect sizes were only 1/2 as big as originally reported. Worse yet, only 36% of the replications produced statistically significant results, whereas 97% of the original studies had.

In 2018, another study was published that made a similar attempt to replicate 21 social science studies that had been published in the 2 most prestigious general science journals: *Nature* and *Science*. This study tested about 5 times as many participants as the original studies had in order to make sure that there was sufficient statistical power to observe effects, even if those effects were small. Despite these efforts, the scientists were unable to replicate the

25 Ironically, the study was published in *Science*, the same journal where Diederik Stapel had published his fictional racism study.

published findings from 8 of the 21 studies. And like the Open Science Collaboration project, the size of the effects in the replication studies were, on average, about 1/2 as big as they had been in the original papers.

The Many Labs project has also conducted similar replication studies, but rather than conducting a single replication of many studies, this project has attempted to conduct many replications of a few studies. There have actually been several Many Labs projects, all with this same goal. In Many Labs 1, published in 2014, 36 different research teams around the world reran the same 13 studies. They then combined results from all the teams to produce well-powered replication attempts with more than 6000 participants. Ten of the 13 studies replicated, but 3 did not.

The take-home message is clear: Psychology has a serious replication crisis. Many of the published findings that we assumed were well established actually aren't. The problem is particularly severe in social psychology, but it extends to other parts of the field. In fact, the problem extends to many other scientific fields, including biology, genetics, and medicine.

FLAWS IN CURRENT SCIENTIFIC PRACTICE

The replication crisis suggests that there are some flaws in current scientific practice that need to be addressed.

First, the standard approach to determining whether an empirical result is statistically significant has some problems. Scientific experiments typically test whether changing

one variable, called the independent variable, has a significant effect on an outcome variable, called the dependent variable.

In Stapel's fictional study, the independent variable was the cleanliness of the environment and the dependent variable was a measure of racist behavior—specifically, how far away white participants sat from African Americans. He made up data suggesting that white participants sat significantly farther away from African Americans when the surrounding environment was messy compared with when it was clean. In other words, changing the independent variable of messiness was claimed to have a significant effect on the dependent variable, which was how far away people sat from others.

But what constitutes a significant effect?

The traditional approach is to try to assess the reliability of the effect. After all, no measurement is perfect. If a thermometer says that your body temperature is 98.7° Fahrenheit, you wouldn't be surprised to learn that it was actually 98.6° or 98.8°. And the same thermometer could give a slightly different reading tomorrow due to other factors, such as the temperature of the air or slight differences in the way you took your temperature or even in the thermometer itself.

That's called measurement error, and it's present in every scientific experiment. But that raises a problem.

Suppose you run an experiment and find that changing your independent variable leads to a change in your dependent variable. What caused the change? The hope is that it was your manipulation of the independent variable. For example, maybe changing from a clean to a messy environment really does make white people sit farther away from African Americans.

But another possibility is that the observed effect is just due to measurement error. So how do you tell the difference?

Most studies use an approach called null hypothesis significance testing. First, you look at how big of an effect manipulating the independent variable had on the dependent variable in your experiment. For example, how much farther away did people sit when the environment was messy compared with when it was clean? If they sat a lot farther away, then that would be a big effect size. But if they only sat a little farther away, then that would be a small effect size. Obviously, big effect sizes are much more convincing.

So what's the probability that you would get an effect of the size you observed, assuming that your independent variable isn't related to your dependent variable? If that probability is very low—less than 5%—then you reject the so-called null hypothesis of no effect. And instead, you accept the alternative hypothesis that the independent variable did have an effect on the dependent variable.

But the problem is that scientists all over the world conduct lots of experiments that never get published, because they didn't produce a large enough effect to reject the null hypothesis using that 5% criterion. And if you can't reject the null hypothesis of no effect, then it's hard to convince other scientists that your results provide convincing evidence for or against any scientific hypothesis.

On the other hand, if you run an experiment and the probability of getting the observed effect is less than 5% under the null hypothesis, then you almost always publish it. In that case, you've passed the conventional bar that scientists have implicitly agreed on, so it's natural to write it up and send it off to a scientific journal so that other scientists can read about it.

But using that 5% criterion means that up to one out of 20 studies will produce a significant effect even when the null hypothesis is true and nothing is going on. So if you tested the same hypothesis over and over again in multiple experiments, occasionally you might get results that allow you to reject the null

hypothesis just because of measurement error, even if the null hypothesis is true.

That wouldn't be such a problem if each experiment was only done once. But there are scientists all over the world who are studying the same topics and trying to test very similar hypotheses, so similar experiments get done all the time. And if any of those experiments produce an effect that is large enough to reject the null hypothesis, then it's likely to be published.

Dozens of other labs around the world may have previously run a very similar experiment and failed to find significant effects, but those experiments were likely never published. So the scientific literature would only include the one study that actually did produce a significant effect.

The result is a significant bias in the scientific literature, with a significant percentage of published results that aren't actually reliable. And that's exactly what the projects that have attempted to replicate existing studies have found.

So null hypothesis significance testing combined with the practice of rarely publishing null results is one important factor contributing to the replication crisis.

Another factor is the use of questionable research practices by some scientists.

The phrase "publish or perish" refers to the idea that if academics don't regularly publish books and papers, then they will inevitably perish professionally. In particular, there's a lot of truth to this idea for scientists at major research universities. The number of papers that you publish, and the number of times those papers are cited, plays a major role in determining whether you'll be able to get a faculty job, whether you'll get a grant to support your research, and what kind of raise you'll get at the end of the year.

There is therefore substantial pressure to publish scientific articles—particularly articles that make a splash. And when that pressure becomes too severe, some scientists cut corners. While some scientists, such as

Stapel, flat-out make up data, there are much subtler practices that can bias results and artificially increase the chances of rejecting the null hypothesis.

For example, you could analyze your data in many different ways and choose the method that produces the largest effect size. You could reanalyze your data set after every participant you test and stop immediately if you happen to achieve statistical significance.

You could also collect data on a whole bunch of different variables and then test if the relationship between any pair of variables passes the 5% threshold, and then you could just report that relationship as if it were your original hypothesis. If you perform enough tests, there's a good chance that one of them will pass the 5% threshold even if the null hypothesis is actually true.

In 2011, Joseph Simmons and colleagues at the University of Pennsylvania published a paper in which they tested how these kinds of questionable research practices can affect scientific results. First, they had a computer generate random data in which they knew that the null hypothesis was true, so there definitely was no real effect. Nevertheless, when they adopted the questionable research practices just described, they were able to produce a result that passed the 5% threshold more than 60% of the time!

Fortunately, there are ways that scientists can overcome these problems and ensure that future published results are solid and replicable.

Perhaps the most promising approach is for studies to be reviewed before any data is collected. Here's the idea: When scientists have an idea for an experiment, they write up their idea before they do any data collection. Then, they send the description of their proposed experiment to a scientific journal to be considered for publication. Other scientists who are experts in the field then review the proposal and evaluate whether the proposed study is solid. If the reviewers decide that it is, then the scientists would conduct the study,

and the scientific journal would agree to publish the paper whether the results turn out to be statistically significant or not. The only condition would be that the scientists have to conduct the study exactly as they described in their proposal.

This model has several advantages. First, it likely leads to better-designed studies because design problems that reviewers point out can be fixed before the study is ever conducted. It also eliminates the problem that only statistically significant results are published and therefore reduces publication bias. Finally, it would significantly reduce the opportunity for scientists to engage in questionable research practices because they have to commit to what they're going to do before they ever conduct the study.

Although bad science does get done, it often gets exposed and corrected. Scientific ethics are discussed, evaluated, implemented, and refined.

SUGGESTED READINGS

Bhattacharjee, "The Mind of a Con Man."

Borgman, *Big Data, Little Data, No Data*.

Chambers, *The Seven Deadly Sins of Psychology*.

Stephens-Davidowitz, *Everybody Lies*.

MULTIPLE-CHOICE QUIZ

1. What was the goal of the Facebook emotional contagion study?

- **a.** To determine if using Facebook improves a person's mood
- **b.** To determine if using Facebook makes a person's mood worse
- **c.** To determine if emotions can spread via social media
- **d.** To determine if Facebook can be used to treat emotional problems

2. What was the goal of the Tuskegee syphilis study?

- **a.** To study the long-term effects of untreated syphilis on the body
- **b.** To determine whether penicillin is an effective treatment for syphilis
- **c.** To determine how syphilis spreads
- **d.** To determine why syphilis was more common in Tuskegee than in other parts of Alabama

Table of Contents

3. Requiring that psychologists obtain informed consent is most related to which of the *Belmont* principles?

a. Respect for persons

b. Beneficence

c. Justice

d. Kindness

4. Which of the following ethical considerations is most related to the *Belmont* principle of beneficence?

a. Recruit participants fairly.

b. Do not coerce people to participate.

c. Respect the rights of vulnerable populations.

d. Do no harm.

5. Who wrote *The Belmont Report*?

a. Peter Buxtun

b. Ted Kennedy

c. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research

d. The Public Health Service

6. What was the main conclusion of the Milgram obedience study?

- **a.** People are inherently evil.
- **b.** People are inherently good.
- **c.** Ordinary people will typically not obey orders that they find unethical.
- **d.** Ordinary people will typically obey orders even if they find them unethical.

7. Which of the following statements about the Stanford Prison Experiment is false?

- **a.** The guards began to act sadistically.
- **b.** Dr. Zimbardo played the role of prison superintendent.
- **c.** Once it was obvious that the prisoners were suffering, the experiment was stopped immediately.
- **d.** The experiment was conducted in the Stanford Psychology Building rather than in a real prison.

8. Which of the following statements about the Neubauer twin study is true?

- **a.** In 2019, the records from the study all became publicly available.
- **b.** When they were children, the twins did not know that they had a biological sibling.
- **c.** The goal of the study was to determine if separating twins at birth was harmful.
- **d.** The study provided conclusive evidence that personality depends more on environment than on genetics.

9. The so-called monster study investigated which of the following?

- **a.** The causes of stuttering
- **b.** The psychological consequences of physical deformity
- **c.** Why ordinary people act in evil ways
- **d.** The role of nature versus nurture in shaping personality

10. On whom was the monster study conducted?

- **a.** People who were disfigured
- **b.** Orphans
- **c.** Twins
- **d.** Adopted children

11. Which of the following men was given psychoactive drugs at Edgewood Arsenal?

- **a.** Harold Blauer
- **b.** Frank Olson
- **c.** James Stanley
- **d.** Ewen Cameron

12. Which of the following organizations did Frank Olson work for?

- **a.** The US Army
- **b.** The Department of Defense
- **c.** The NSA
- **d.** The CIA

13. What was the goal of the MK-Ultra project?

- **a.** To explore the use of psychoactive drugs for use in mind control
- **b.** To test the effectiveness of gas masks
- **c.** To study extrasensory perception
- **d.** To gather scientific evidence for or against telepathy

14. Harold Blauer was experimented on while doing what?

- **a.** Being treated for syphilis
- **b.** Being treated for depression
- **c.** Working for the CIA
- **d.** Serving as a soldier

15. The Tearoom Trade study investigated which of the following?

- **a.** Sex in public restrooms
- **b.** Behavior in restaurants
- **c.** Cultural differences in what is considered polite
- **d.** How people negotiate

16. Which of the following statements about the Tearoom Trade study is true?

- **a.** Most of the participants were openly gay.
- **b.** Most of the participants were aware that their behavior was being studied.
- **c.** The researcher used license plate numbers to track down the participants for interviews.
- **d.** The researcher obtained informed consent before conducting the observations.

17. Which of the following statements about the Bruce/Brenda Reimer case is false?

a. Dr. Money recommended that the child be raised as a girl.

b. The child's parents did their best to raise the child as a girl.

c. Dr. Money believed that children are born gender-neutral.

d. Friends and teachers thought that the child acted like a typical girl.

18. Lesson 6 discussed 3 advantages of analyzing the enormous data sets produced by social media, web search engines, and smartphone apps. Which of the following is not one of the advantages that was addressed?

a. People in these data sets have all already consented

b. The opportunity to study psychological behavior "in the wild"

c. The enormous size and diversity of the sample

d. The cost-effectiveness of analyzing an existing data set

19. How were Diederik Stapel's unethical practices exposed?

a. He told some friends about it and they turned him in.

b. Some of his junior colleagues followed him and observed him throwing out questionnaires that he claimed to be using for data collection.

c. Reviewers identified unethical practices in one of his papers.

d. Other scientists repeatedly failed to replicate his findings and called him out.

20. A number of attempts to replicate published psychological results have found that which of the following is true?

a. Roughly 1/2 of published results don't replicate.

b. Almost all published results are very reliable.

c. Studies that don't replicate were almost always conducted unethically.

d. Scientists are more likely to publish null results than significant results.

Answer Key

1. C, 2. A, 3. A, 4. D, 5. C, 6. D, 7. C, 8. B, 9. A, 10. B, 11. C, 12. D, 13. A, 14. B, 15. A, 16. C, 17. D, 18. A, 19. B, 20. A

BIBLIOGRAPHY

Adjerid, Idris, and Ken Kelley. "Big Data in Psychology: A Framework for Research Advancement." *The American Psychologist* 73, no. 7 (2018): 899–917. doi:10.1037/amp0000190.

Albarelli, H. P. *A Terrible Mistake: The Murder of Frank Olson and the CIA's Secret Cold War Experiments*. Waterville, OR: Trine Day, 2009.

Bhattacharjee, Yudhijit. "The Mind of a Con Man." *The New York Times Magazine*, April 26, 2013. https://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?ref=magazine.

Borgman, Christine L. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: MIT Press, 2015.

Chambers, Chris. *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*. Princeton, NJ: Princeton University Press, 2017.

Colapinto, John. *As Nature Made Him: The Boy Who Was Raised as a Girl*. New York: HarperCollins, 2000.

Humphreys, Laud. *Tearoom Trade: Impersonal Sex in Public Places*. Chicago: Aldine, 1970.

Jones, James H. *Bad Blood: The Tuskegee Syphilis Experiment*. New York: Free Press, 1993.

Khatchadourian, Raffi. "Operation Delirium." *The New Yorker* 88, no. 40 (December 17, 2012).

Kramer, Adam D. I., Jamie E. Guillory, and Jeffrey T. Hancock. "Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks." *Proceedings of the National Academy of Sciences* 111, no. 24 (June 17, 2014): 8788–8790. https://www.pnas.org/content/pnas/111/24/8788.full.pdf.

Milgram, Stanley. *Obedience to Authority*. London: Pinter & Martin, 2010.

Money, John, and Anke A. Ehrhardt. *Man & Woman, Boy & Girl: Gender Identity from Conception to Maturity*. Northvale, NJ: Jason Aronson, 1996.

Moreno, Jonathan D. *Undue Risk: Secret State Experiments on Humans*. London: Routledge, 2016.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. 1979. https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf.

Neubauer, Peter B., and Alexander Neubauer. *Nature's Thumbprint: The New Genetics of Personality*. Reading, MA: Addison-Wesley, 1990.

Oatley, Keith. *Our Minds, Our Selves: A Brief History of Psychology*. Princeton, NJ: Princeton University Press, 2018.

Regis, Edward. *The Biology of Doom: America's Secret Germ Warfare Project*. New York: Henry Holt, 1999.

Reverby, Susan M. *Examining Tuskegee: The Infamous Syphilis Study and Its Legacy*. Chapel Hill: The University of North Carolina Press, 2009.

Schein, Alyse, and Paula Bernstein. *Identical Strangers: A Memoir of Twins Separated and Reunited*. New York: Random House, 2008.

Stephens-Davidowitz, Seth. *Everybody Lies: What the Internet Can Tell Us about Who We Really Are*. London: Bloomsbury, 2017.

Tudor, Mary. "An Experimental Study of the Effect of Evaluative Labeling of Speech Fluency." Master's thesis, State University of Iowa, 1939. https://doi.org/10.17077/etd.9z9lxfgn.

Wright, Lawrence. T*wins: And What They Tell Us about Who We Are*. New York: John Wiley & Sons, 1997.

Zimbardo, Philip. *The Lucifer Effect: Understanding How Good People Turn Evil*. New York: Random House, 2013.

Image Credits

Page 5	Dem10/iStock/Getty Images Plus
Page 7	Tero Vesalainen/iStock/Getty Images Plus
Page 11	CSA Images/Vetta/Getty Images Plus
Page 21	Yoingco34/Fiverr.com
Page 27	catavic/Fiverr.com
Page 28	catavic/Fiverr.com
Page 29	Rattankun Thongbun/iStock/Getty Images Plus
Page 34	JurgaR/iStock/Getty Images Plus
Page 41	maratam/Fiverr.com
Page 47	Sidharth Ojha/www.instagram.com/sidharthojha
Page 49	Sidharth Ojha/www.instagram.com/sidharthojha
Page 52	Sidharth Ojha/www.instagram.com/sidharthojha
Page 57	Meinzahn/iStock/Getty Images Plus
Page 69	Andreus/iStock/Getty Images Plus
Page 73	Melpomenem/iStock/Getty Images Plus