

Teaching Contentious Classics

September 30, 2014



Last year marked the 50th anniversary of Stanley Milgram’s experiments on obedience to authority, an event that inspired a conference, many reflective papers, and a popular book of vitriolic criticism. Gina Perry’s *Behind the Shock Machine: The Untold Story of the Notorious Milgram Psychology Experiments* aimed to discredit Milgram’s research methods, his findings, his ethics, and the man himself. Reviewing the book and discussing it with Milgram scholars got me thinking yet again about the eternal dilemma for instructors and textbook writers: How much time should we devote to teaching the classics versus making way for new (and often yet-unreplicated) research, and how should we teach them? What should we do when the classics had flaws that would not permit their replication today, but yielded findings that still tell a good story?

In every generation, certain studies get planted in our books and lectures, and they tend to become rooted there. Over time it gets harder to decide how much to prune — let alone decide if it’s time to uproot them. We stop looking at the original studies closely, let alone critically; they just sit in our courses like grand historical monuments.

However, it’s good to reexamine them for two important reasons: One is for our own sakes, to refresh our memories and rethink their contributions; the other is for our students’ sakes. Students today are as eager to reject unflattering or counterintuitive portrayals of humanity as students were decades ago. Teaching the classics, therefore, means finding new ways of persuading students that these findings do apply to them, despite the errors or limitations of the original studies. I’d therefore like to examine the stories behind three classic studies in psychology: Muzafer Sherif’s Robbers Cave study, Milgram’s obedience experiments, and Harry Harlow’s studies of wire and cloth mother monkeys.

Robbers Cave

To begin, I got down my graduate school bible, *Basic Studies in Social Psychology*, and reread Sherif’s “Superordinate Goals in the Reduction of Intergroup Conflict,” which had been published in a sociology journal in 1958. Between 1949 and 1954, Sherif and his colleagues used a Boy Scout camp called

Robbers Cave to test their hypotheses about the origin and reduction of intergroup hostility and prejudice. They randomly assigned 11- and 12-year-old boys to one of two groups, the Eagles or the Rattlers, and set up a series of competitive activities that quickly generated “us–them” thinking and mutual animosity. Later, to undo the warfare they had thus created, they set up situations in which the boys had to work together to achieve a common goal.



Muzafer Sherif's classic Robbers Cave studies focused on 11- and 12-year-old Boy Scout campers who were assigned to separate groups and placed in a series of competitive situations that fostered an “us–them” mentality.

As I reread the study, I realized the data were more limited than I had remembered and not as statistically “scientific” as would be required today. Most of the conclusions, Sherif wrote, “were reached on the basis of observational data”— confirmed by “sociometric choices and stereotype ratings.” He said, “Observations made after several superordinate goals were introduced showed a sharp decrease in name-calling and derogation of the outgroup common ... in the contact situations without superordinate goals.” (By the way, there are unexpected discoveries in going back to read original studies. The “name-calling” in the Robbers Cave experiment is so charmingly outdated: In 1948, boys “derogated” each other with names like “stinkers” and “smart alecks.”)

Sherif did provide some numbers and percentages and a few chi squares, but this was a field study, with all of the uncontrollable variables that field studies can generate. Was everything hunky dory for the Eagles and Rattlers by the end of the study? The numbers of boys favorable toward their outgroup improved, but the majority of boys in each group apparently maintained their hostility toward each other.

Yet Robbers Cave was and remains important for its central ideas: At the time, most psychologists did not understand — and most laypeople don't understand even today — that simply putting two competing, hostile groups together in the same room to, say, watch a movie, won't reduce their antagonism; that competitive situations generate hostility and stereotyping of the outgroup; and that competition and hostility can be reversed, at least modestly, through cooperation in pursuit of shared goals. That's the story of Robbers Cave: It was true then, and it's true now. It was bold and innovative of Sherif to try to test these important hypotheses in a realistic situation outside the lab, and I don't see any need for teachers to raise concerns about his methods with students in introductory psychology classes.

In fact, Robbers Cave generated a long line of experimental and field studies replicating the importance

of superordinate goals. When Elliot Aronson went into the newly desegregated but hostile classrooms in Austin, Texas, where African American, Mexican American, and Anglo American children were at war with each other, Sherif was part of his mental set, strongly influencing his design of the jigsaw classroom. But Elliot did it right, using an experimental intervention and a control group. What a great coda to the Robbers Cave story — a direct link from Eagles and Rattlers, a made-up antipathy, to interethnic warfare in our schools, which is all too real and persisting.

The Shock Box

Teaching the lessons of Stanley Milgram's experiments, of course, is a lot more complicated than teaching Sherif's work. Phoebe Ellsworth (University of Michigan) recently told me that she surveyed her upper-level social psych course to find out how students had first heard about Milgram. Roughly half of them were first introduced to him through his work demonstrating obedience to authority and the power of the situation. The other half first heard about him as an example of how unethical social psychologists could be. The students who had been told about the "evil Milgram," she said, mostly weren't even told about the point of his experiment. "Pathetic," she wrote to me. I agree.



Today, many psychological scientists note ethical lapses in Stanley Milgram's "shock box" experiments on obedience, but his research continues to influence our understanding of the power that circumstances have over behavior.

Again, students need to keep the cultural context of the times in mind. In 1961, when Adolf Eichmann was claiming at his trial that he was "only following orders" in the murder of Jews during the Holocaust, Milgram began his effort to determine how many Americans would obey an authority figure when directly ordered to harm another human being.

Participants came to the Yale University lab thinking they were part of an experiment on the effects of punishment on learning, and were instructed to administer increasing levels of shock to a "learner." The learner was a confederate of Milgram who did not receive any shocks, but played his part convincingly: As the study continued, he shouted in pain and pleaded to be released, according to an arranged script. To almost everyone's surprise at the time, some two-thirds of the participant "teachers" administered what they thought were the highest levels of shock, even though many were experiencing difficulty or distress doing so.

Milgram's experiment produced a firestorm of protest about the possible psychological harm inflicted on the unwitting participants, and as a result, it could never be done today in its extreme version.

Some people hated the method and others the message, and still do — which is why debate about it continues. In her book, Gina Perry, an Australian psychologist and journalist, interviewed everyone she could find who was connected to the original study, along with Milgram's critics and defenders. She pored through the archives of Milgram's voluminous unpublished papers. Her goal was to argue that the experiments were flawed and unethical, in order to counteract what she considers Milgram's "bleak view of human nature."

Reinvestigations almost invariably yield some useful discoveries. Perry found violations of the research protocol: Over time, the man playing the experimenter began to drift off script, urging reluctant "teachers" to keep going longer than they were supposed to. To my own dismay, I learned that Milgram committed what researchers, even then, would have considered a serious breach of ethics: He did not fully debrief subjects at the end of each experiment. "Teachers" got to meet the "learner" face-to-face after the experiment, so they could be assured that the "shocks" had not harmed him. But they were not told that all those escalating levels of shocks were completely fake because Milgram feared the word would get out and invalidate future participants' behavior. It was almost a year before subjects were mailed a full explanation. Some never got it; some never understood the purpose of the whole experiment. That's inexcusable. Such important revelations add complexity to the Milgram story, though again, not, I think, for introductory psychology students, who will use these details to dismiss the larger lesson.

For critics like Perry, these flaws are reason enough to kick Milgram off his pedestal and out of our textbooks. I disagree. I think we must continue giving his experiments the prominent position we do, and for the same reason we originally did. When I first read about Milgram's experiments in grad school, I remember thinking, "Very clever, but what do they contribute? Wasn't Nazi Germany evidence enough of obedience to authority?" But that was Milgram's point: In the early 1960s, Americans — and American psychologists — deeply believed in national character. Germans obeyed Hitler, it was widely assumed, because obedience was in the German psyche: Look at all those high scores on the Authoritarian scale. People believed it could never happen here.

Elliot Aronson tells the following story in his memoir, *Not by Chance Alone*. When he was at Harvard University in 1960, his first year out of graduate school, he gave a guest lecture in Gordon Allport's class. Allport, the grand old man of social psychology, introduced him as a "master of mendacity" because of the dramatic experiments on cognitive dissonance that were already bringing Elliot fame. Elliot was mildly insulted, naturally, and in talking with Allport afterward he defended the use of "deception" in high-impact experiments as "not lying, but theater." Allport replied: "Why do you guys go through all that rigmarole? Why don't you just ask the participants what they would do?" This, from Gordon Allport!

Elliot tried to explain that most people cannot predict or account for their own behavior with any degree of accuracy, but Allport was unpersuaded. A month or two later, Elliot went to Yale to give a colloquium and met Milgram for the first time. Milgram described the experiment he was planning and laid out its basic design. Elliot said, "Wow, I'll bet a sizable number of people dole out more intense shocks than they themselves would ever have predicted." Even he, however, never dreamed that two-thirds of them would go all the way.

For me, reading Perry's criticisms made clear why the Milgram experiments deserve their prominence.

“Deep down, something about Milgram makes us uneasy,” she writes. Indeed something does: his evidence that situations have power over our behavior. This is a difficult message, and most students — indeed, most people — have trouble accepting it. “I would never have pulled those levers!” they cry. “I would have told that experimenter what a ... stinker ... he is!” Perry insists that people’s personalities and histories influence their actions. But Milgram never disputed that fact; his own research revealed that many participants resisted. “There is a tendency to think that everything a person does is due to the feelings or ideas within the person,” Milgram wrote. “However, scientists know that actions depend equally on the situation in which a man finds himself.”

Notice the “equally” in that sentence; many critics, like Perry, don’t.

One of the original subjects in the study, called Bill, tried to explain to Perry why the experiments were so valuable and why he did not regret participating. He hadn’t thought about the experiment for 20 years, he said, until he began dating a psychology professor. She, thrilled to have met a living link to the experiment, asked him to speak to her class.

“Well,” Bill tells Perry, “you would have thought Adolf Hitler walked in the room. I never really thought about it that way, you know?” Bill told the students, who were silently sitting in judgment on him, “It’s very easy to sit back and say, ‘I’d never do this or that’ or ‘Nobody could ever get me to do anything like that.’ Well, guess what? Yes, they can.”

That, of course, is the moral of the story. But what interests me here is the wall of hostility Bill says he felt from the students. That means the students, like Gina Perry, weren’t getting it. They were reading about the experiment, seeing the films, and still not understanding that they themselves might have been Bill. As long as students regard the obedient participants as being the equivalent of Adolf Hitler, the Milgram experiments — which have yielded approximately the same findings whenever and wherever they have been replicated, whether in “softer” versions or cyberversions or TV versions — are as important as ever.

Perhaps one way to broaden acceptance of the Milgram message is to show how it generated research on the psychology of the minority who resisted. The obedience studies might shock or depress students who think they provide a “bleak view of human nature,” but these experiments of majority behavior also launched research into the conditions under which a brave minority becomes more likely to dissent, blow the whistle, disobey, and otherwise resist tyranny. That is, Milgram’s work spurred investigation into the fuller human story: the bleak and the inspiring, the conformist and the rebel.

Harlow’s Monkeys

I turn now to Harry Harlow’s classic experiments, conducted throughout the 1950s and 1960s, on the importance of contact comfort. Harlow took infant rhesus monkeys away from their mothers and raised them with a “wire mother,” a forbidding construction of wires with a milk bottle connected to it, and a “cloth mother,” a similar construction but one covered in foam rubber and terry cloth. At the time, it was widely believed (by psychologists, if not mothers) that babies become attached to their mothers simply because mothers provide food. But Harlow’s baby monkeys ran to the terry-cloth mother whenever they were frightened or startled, and clinging to it calmed them down. They went to the wire mother only for milk, and immediately abandoned it after they had finished feeding.



Although his experiments on primates are today considered cruel, Harry Harlow showed the importance of contact comfort at a time when many experts doubted the developmental significance of physical affection. Photo credit: UW-Madison Archives (#S01464)

Every introductory class and textbook tells this story, with heartbreaking photos of infant monkeys clinging to their wire and cloth mothers when a scary moving toy was put into their cage. Wasn't this discovery something "we all knew" — in this case, that infants need contact comfort even more than they need food if they are to flourish? Didn't we have enough data from psychoanalysts René Spitz and John Bowlby, who famously observed abandoned infants warehoused in orphanages?

Apparently not. As journalist Deborah Blum describes in her biography, *Love at Goon Park: Harry Harlow and the Science of Affection*, most American psychologists at the time were under the influence of either behaviorism or psychoanalysis, two apparently opposite philosophies that nonetheless shared a key belief: that the origin of a baby's attachment to the mother was through food. Behaviorists believed that healthy child development required positive reinforcement: Baby is hungry; hunger drive is satisfied; baby becomes conditioned to associate mother with food; mother and breast are equated. Interestingly, that was the Freudian view as well: No mother need be present, only a breast. "Love has its origin," Freud wrote, "in attachment to the satisfied need for nourishment." Why would cuddling be necessary? For the eminent behaviorist John Watson, cuddling was coddling.

But whereas Milgram's findings need constant reiteration in every generation, Harlow's research no longer surprises us. One might say that its very success has made teaching it unnecessary: No one would argue against Harlow's findings, as many students always want to do with Milgram's. Adult humans could choose to walk out of Milgram's experiment at some point, and a third of them did. But the monkeys were captives, tortured by their isolation. In recent decades, psychologists have learned that the word "torture" is not an exaggeration to describe the experience of isolation for any primate. And to torture infants is horrible. But the fact that so many people think it is horrible now — and so many didn't then — is an extraordinary story for teachers to tell. How has it happened that we have extended the moral circle to include other primates?

In 1973, as a young editor at *Psychology Today*, I interviewed Harlow. I walked through his lab with our photographer, Rod Kamitsuka, and looked aghast at a roomful of monkeys cowering in their individual

cages, electrodes on their heads. When Rod took a picture of one, it became wildly excited and fearful, careening around its tiny cage trying to escape. Rod and I were horrified, but Harlow was amused by us. “I study monkeys,” Harlow said, “because they generalize better to people than rats do, and I have a basic fondness for people.” I asked him what he thought of his critics who said that taking infants from their mothers was cruel and that the results did not justify the cruelty. He replied: “I think I am a soft-hearted person, but I never developed a fondness for monkeys. Monkeys do not develop affection for people. And I find it impossible to love an animal that doesn’t love back.” Today, that sounds like lame moral reasoning: The fact that animals don’t love us is no justification for torturing them.

When I revisited Harlow’s work, however, I was reminded of how many pioneering discoveries he made, most of them lost in the telling of the main story of contact comfort. He also demonstrated that monkeys use tools, solve problems, and learn and explore because they are curious or interested in something beyond just food or other rewards. He demonstrated the importance of contact with peers, which can overcome even the detrimental effects of maternal deprivation. Harlow created a nuclear family unit for some of the monkeys and found that under those conditions, rhesus males became doting fathers — something they don’t do in the wild.

Harlow was hardly the first to demonstrate the power of “mother love,” the necessity of contact comfort, and the devastation that ensues when an infant is untouched, unloved, neglected. Was experimenting with monkeys, by raising them in isolation with only wire and cloth mothers and causing them anguish that no observer could fail to see, essential to make the same point that Bowlby and Spitz had made? I don’t know. What Harlow did, like Milgram, was to make his case dramatic, compelling, and scientifically incontrovertible. The evidence was based not on anecdote or observation, however persuasive, but on empirical, replicated data. As Blum shows, that’s what it took to begin to undermine a scientific worldview in which the need for touch and cuddling — physical expressions of mother love — had been so deeply ignored.

When I was first thinking about this topic, I was prepared to argue for jettisoning Harlow, given that his findings no longer surprise nor serve to persuade students to change a deeply held belief. Perhaps that judgment reflects my ineradicable memory of seeing those helpless, suffering baby monkeys. But in revisiting his work, I changed my mind. We should keep him; we should discuss his discoveries, while expanding our story of what they mean. Harlow’s work is a great chapter in the story of psychology: It shows not only how we thought about mothers, but also how we thought about monkeys. It shows how dominant psychological perspectives influence our lives — in his day, behaviorism or psychoanalysis; in our day, genetics and brain — and seep into the questions we ask and the studies we conduct. The take-home message for students is not “Look how much smarter, kinder, and more ethical we are today than those guys were,” but rather (1) Where would we be without these classics? What do they teach us about humanity that made them classics? (2) What is happening in today’s culture that affects the questions scientists are asking now — and the answers they get? Where might our own mistakes and biases lie — we, with all our institutional review boards and informed consents. Where are our failings of ethics and methods? The classics are living history, and we are not at the end of history by any means.

This article is based on talk that Carol Tavris delivered at the 26th Annual Convention of the Association for Psychological Science. A version of it appeared in the July 18 issue of the *Times Literary Supplement*.

References and Further Reading

Aronson, E. (2010). *Not By Chance Alone: My Life as a Social Psychologist*. New York, NY: Basic Books.

Blum, D. (2011). *Love at Goon Park: Harry Harlow and the Science of Affection*. New York, NY: Basic Books.

Perry, G. (2013). *Behind the Shock Machine: The Untold Story of the Notorious Milgram Psychology Experiments*. New York, NY: The New Press.

Sherif, M., Harvey, O. J., White B. J., Hood, W. R., & Sherif, C. W. (1988). *The Robbers Cave experiment: Intergroup Conflict and Cooperation*. [Orig. pub. as *Intergroup Conflict and Group Relations*]. Middletown, CT: Wesleyan University Press.