

Replicating Milgram

Would People Still Obey Today?

Jerry M. Burger
Santa Clara University

The author conducted a partial replication of Stanley Milgram's (1963, 1965, 1974) obedience studies that allowed for useful comparisons with the original investigations while protecting the well-being of participants. Seventy adults participated in a replication of Milgram's Experiment 5 up to the point at which they first heard the learner's verbal protest (150 volts). Because 79% of Milgram's participants who went past this point continued to the end of the shock generator's range, reasonable estimates could be made about what the present participants would have done if allowed to continue. Obedience rates in the 2006 replication were only slightly lower than those Milgram found 45 years earlier. Contrary to expectation, participants who saw a confederate refuse the experimenter's instructions obeyed as often as those who saw no model. Men and women did not differ in their rates of obedience, but there was some evidence that individual differences in empathic concern and desire for control affected participants' responses.

Keywords: obedience, authority, Milgram

Stanley Milgram's (1963, 1965, 1974) obedience studies are arguably the most well-known social psychological research inside or outside the field. References to the studies continue to appear in popular media, including movies and songs (Blass, 2004), and a social psychology textbook that does not include a discussion of the research is almost unthinkable. In truth, Milgram's work is more properly described as a series of demonstrations rather than as an experiment (Burger, 2002), and the absence of a theoretical model at the outset of the research impeded Milgram's efforts to publish the initial reports of his investigations (Blass, 2004). Nonetheless, the haunting images of participants administering electric shocks and the implications of the findings for understanding seemingly inexplicable events such as the Holocaust and Abu Ghraib have kept the research alive for more than four decades (Miller, 2004).

Although Milgram developed many variations of his basic procedure, the study most psychologists are familiar with is Experiment 5 (Milgram, 1974). Briefly, a participant and a confederate were told the study concerned the effects of punishment on learning. Through a rigged drawing, the participant was assigned the role of teacher while the confederate was always the learner. The participant watched as the experimenter strapped the learner to a chair

in an adjacent room and attached electrodes to the learner's arm. The participant's task was to administer a paired-associate learning test to the learner through an intercom system. The learner indicated his responses by pressing buttons connected to answer lights on the participant's side of the wall. Participants sat in front of an imposing shock generator and were instructed to administer an electric shock to the learner for each incorrect answer. Labels above the 30 switches that spanned the front of the machine indicated that the shocks ranged from 15 to 450 volts in 15-volt increments. Participants were instructed to start with the lowest switch and to move one step up the generator for each successive wrong answer.

In actuality, the learner received no shocks. But he gave many wrong answers, which required the participant to administer shocks of increasingly stronger voltage. Following the administration of the 150-volt punishment, the participant heard the learner's cries of protest through the wall. The learner said he wanted out, that he was experiencing excessive pain, and that his heart was bothering him. From that point until the 330-volt switch, the learner yelled in pain and demanded to be released after each shock. After the 300-volt shock, the learner refused to answer (which the experimenter said to treat as a wrong answer). After the 330-volt shock, the learner no longer screamed or protested when receiving a shock, suggesting that he was physically incapable of responding. The major dependent variable was the point in the procedure at which the participant refused to continue. The experimenter, who sat a few feet away, encouraged the participant to continue at each verbal or nonverbal sign of resistance. The study proceeded until the participant expressed resistance to each of four increasingly demanding prods by the experimenter or until the participant had pressed the highest switch on the shock generator three times. The unsettling finding was that 65% of the participants in this version of the experi-

The research reported in this article was featured on ABC News's January 3, 2007, broadcast of *Primetime*.

I thank the following individuals for their assistance with data collection: Matthew Bell, Kenneth Courtney, Sarah Gibson, Brian Oliveira, Neda Raymond, and Tamara Trumbo. Thanks are also extended to the clinical psychologists who screened potential participants and to Thomas Blass, who supplied information about some of the materials used in the study.

Correspondence concerning this article should be addressed to Jerry M. Burger, Department of Psychology, Santa Clara University, Santa Clara, CA 95053-0001. E-mail: jburger@scu.edu

**Jerry M.
Burger**



ment continued to administer shocks all the way to the end of the generator's range.

Ethical Concerns

In addition to their scientific value, the obedience studies generated a great deal of discussion because of the ethical questions they raised (Baumrind, 1964; Fischer, 1968; Kaufmann, 1967; Mixon, 1972). Critics argued that the short-term stress and potential long-term harm to participants could not be justified. In his defense, Milgram (1974) pointed to follow-up questionnaire data indicating that the vast majority of participants not only were glad they had participated in the study but said they had learned something important from their participation and believed that psychologists should conduct more studies of this type in the future. Nonetheless, current standards for the ethical treatment of participants clearly place Milgram's studies out of bounds (Elms, 1995). No study using procedures similar to Milgram's has been published in more than three decades (Blass, 2000).

The 150-Volt Solution

I always anticipate the reaction to one scene in particular when I show my undergraduate students the grainy black-and-white video from the Milgram studies. When the participant presses the 150-volt switch, the learner vehemently protests and demands to be released from the study. This is the critical moment in the procedure. Nearly every participant paused, and most turned to the experimenter to indicate verbally or nonverbally their reluctance to continue. For students seeing the film for the first time, it is the jaw-dropping moment. The man said he wanted out. How could anyone continue? Indeed, when Milgram asked psychiatrists, college students, and middle-class adults to predict their own behavior, the 150-volt

point was by far the most common guess as to how far they would go (Milgram, 1974).

The data make the same point. Of the 14 participants (out of 40) who stopped prior to reaching the 450-volt switch in this version of the experiment, 6 stopped after hearing the protests at 150 volts. One participant had stopped earlier. Only 7 participants who went past 150 volts stopped at all. Another way to say this is that 79% of the people who continued past 150 volts (26 of 33) went all the way to the end of the shock generator's range. In short, the 150-volt switch is something of a point of no return. Nearly four out of five participants who followed the experimenter's instructions at this point continued up the shock generator's range all the way to 450 volts.

This observation suggests a solution to the ethical concerns about replicating Milgram's research. Knowing how people respond up to and including the 150-volt point in the procedure allows one to make a reasonable estimate of what they would do if allowed to continue to the end. Stopping the study within seconds after participants decide what to do at this juncture would also avoid exposing them to the intense stress Milgram's participants often experienced in the subsequent parts of the procedure.

Additional Safeguards

In my replication of Milgram's research, I took several additional steps to ensure the welfare of participants. First, I used a two-step screening process for potential participants to exclude any individual who might have a negative reaction to the experience. Second, participants were told at least three times (twice in writing) that they could withdraw from the study at any time and still receive their \$50 for participation. Third, like Milgram, I had the experimenter administer a sample shock to the participants (with their consent) so they could see that the generator was real and could obtain some idea of what the shock felt like. However, a very mild 15-volt shock was administered rather than the 45-volt shock Milgram gave his participants. Fourth, I allowed virtually no time to elapse between ending the session and informing participants that the learner had received no shocks. Within a few seconds of the study's end, the learner entered the room to reassure the participant that he was fine. Fifth, the experimenter who ran the study also was a clinical psychologist who was instructed to end the study immediately if he saw any signs of excessive stress. In short, I wanted to take every reasonable measure to ensure that the participants were treated in a humane and ethical manner. Of course, the procedures also were approved by the Santa Clara University institutional review board.

Explaining the Effect

Although ethical concerns have prevented psychologists from replicating Milgram's procedures, discussion and debate about how to interpret the findings have never ended (Blass, 2004; Miller, Collins, & Brief, 1995). Nonetheless, most social psychologists appear to agree on one point. The obedience studies are a dramatic demonstration of how

individuals typically underestimate the power of situational forces when explaining another person's behavior. Prior to conducting his research, Milgram (1974) asked Yale students and 40 psychiatrists to predict the outcome of the studies and found nearly universal agreement that virtually no one would continue to the end of the shock generator's range. It should be noted that subsequent investigations found that individuals provided with more details about the procedures still gave estimates of obedience lower than those in the actual findings but not as far off the mark as Milgram's data suggested (Blass, 2000). What caused Milgram's participants to act in such unexpected ways? Below I briefly describe four features of the situation that likely contributed to the high rates of obedience.

Obedience to Authority

Milgram (1974) maintained that the key to obedience had little to do with the authority figure's manner or style. Rather, he argued that people follow an authority figure's commands when that person's authority is seen as legitimate. Moreover, our culture socializes individuals to obey certain authority figures, such as police officers, teachers, and parents. Milgram's experimenter was granted the legitimacy of authority by virtue of his association with the experiment, the university, and perhaps even science. Subsequent discussions have raised questions about the nature of the experimenter's authority. In particular, it seems likely that the perceived expertise of the experimenter contributed to the participants' decision to follow the instructions (Morelli, 1983). The experimenter presumably had knowledge about the procedure and had gone through many previous sessions, and so participants deferred to his judgment. Milgram (1983) acknowledged this confounding of position and expertise in his studies but pointed out that this is often the case in real-life examples of obedience.

Gradual Increase in Demands

Another feature of the situation Milgram created that most likely contributed to the high rates of obedience was the incremental nature of the task (Gilbert, 1981). Participants always started with the lowest voltage switch, a relatively mild 15 volts with no noticeable effect on the learner, and proceeded in 15-volt increments up the shock generator's range. We know from a great deal of subsequent research that this type of gradual increase in the size of demands is an effective tactic for changing attitudes and behavior (Cialdini & Goldstein, 2004). Among the psychological processes set in motion with this procedure are a need for consistency and a self-perception process. The well-demonstrated need to act and appear in a consistent manner would have made it difficult for a participant to refuse to press the 195-volt switch after just pressing the 180-volt switch. Moreover, agreeing to small requests, such as pressing the low-voltage switches, can change the way people think about themselves (Burger, 1999). Participants may have come to see themselves as the kind of persons who follow the experimenter's instructions.

Limited Sources of Information in a Novel Situation

It is a fair assumption that Milgram's participants had never before been in a situation like that in the obedience studies and that they had not given any thought to how they might or should act if they ever found themselves in this kind of setting. It can also be assumed that once participants realized the difficult position they had gotten themselves into, they began an immediate search for information about how they ought to respond. However, relevant information was quite limited. The primary source of information was the experimenter, who presumably knew all about the experiment and had gone through the procedure with many previous participants. Although he obviously heard the learner's cries and complaints, the experimenter acted as if nothing was wrong and continually reassured the participant that the shocks were not dangerous and that continuing with the procedure was appropriate. In the absence of additional information, it was not unreasonable for participants to defer to the experimenter's expertise, at least for a while. Data consistent with this analysis can be found in one variation of the basic procedure in which Milgram (1974) used two experimenters. When the learner first protested at the 150-volt level, one experimenter encouraged the participant to continue, but the other experimenter expressed concern and asked the participant to discontinue. In this situation, obedience virtually disappeared.

Although Milgram's participants typically had only the experimenter's behavior to rely on, individuals making these kinds of decisions often have one other source of information—the behavior of other people. A wealth of studies demonstrates that people often rely on perceived norms when making decisions about their own behavior (Cialdini, Kallgren, & Reno, 1991). It is reasonable to speculate that obedience rates would have dropped considerably if Milgram's participants had been informed before the study that the vast majority of previous participants refused to continue when they heard the learner's protests. One of Milgram's (1974) variations provides support for this interpretation. In Experiment 17, three "teachers" were used, one real participant and two confederates. Teacher 1 read the word pairs, Teacher 2 announced whether the answer was right, and Teacher 3 (the real participant) administered the shock. After hearing the learner's protests following the 150-volt shock, Teacher 1 dramatically refused to participate and moved to the other side of the room. The test continued until Teacher 2 also refused to continue in a dramatic fashion at the 210-volt level. At this point, the experimenter told the real participant to continue the study by himself. Milgram found that 7.5% of the participants refused to go on as soon as Teacher 1 quit, 30% refused to continue immediately after the second teacher quit, and only 10% continued to the end of the study.

Responsibility Not Assigned or Diffused

Absence of responsibility has often been cited by psychologists as a contributing factor to aggressive and abhorrent

behavior (Bandura, 1999). Indeed, when Milgram's participants asked about responsibility, the experimenter specifically stated that he himself was responsible for any harm to the learner. Milgram (1974) reported that many of his participants placed responsibility for their own actions on the experimenter, taking a "just following orders" position in explaining why they continued the shocks. When Milgram arranged the situation so that a confederate delivered the shocks while participants performed "a subsidiary act" in carrying out the study (Experiment 18), only 3 of 40 participants refused to be a part of the study. The participants in this condition presumably attributed responsibility for hurting the learner to the person pressing the switches, which thereby allowed them to continue their roles as assistants.

Hypotheses

Would People Still Obey Today?

A persistent question asked about Milgram's research is whether his findings would be replicated today (Blass, 2000). Some people have argued that individuals these days are more aware of the dangers of blindly following authority than they were in the early 1960s. However, Blass (2004) found no evidence for a change in obedience over time when reviewing studies patterned on Milgram's procedures. I predicted that any differences in obedience between the 1961–1962 participants (Milgram, 1963) and the 2006 participants would be minimal. Although changes in societal attitudes could have an impact on obedience, I argue that the question about changes over time may represent another example of the fundamental attribution error. That is, rather than acknowledging the power of the situational forces set in motion in Milgram's procedure, those who suggest changes in obedience-proneness over time may be too focused on the individual. There is no reason to think that the situational features described earlier would not still be operating 45 years after Milgram's investigations.

Norm Information

If participants in obedience studies seek out information about appropriate responses to their situation, then information about how others behave in the situation should influence their decisions about whether to continue. But how much norm information is sufficient to overcome the forces pushing participants toward obedience? Certainly if informed that 40 out of 40 participants to date had refused to press the switches, few if any participants would continue. But I was more interested in a situation in which some degree of ambiguity about appropriate behavior remained. I speculated that seeing just one other person refuse to shock the learner might provide useful information about what participants are supposed to do in this setting. Because participants are torn between doing what the experimenter tells them and not wanting to hurt the learner, I reasoned that they might eagerly rely on this limited norm information to conclude that refusing the experimenter's instructions is appropriate. Researchers find that people often rely on single examples when drawing inferences, particularly when the example is salient (Gilovich & Savitsky, 2002).

I also wanted to create a situation in which the modeled refusal was less dramatic than Milgram's variation in which two confederates boldly announced their refusal and physically moved away from the shock generator. Moreover, I wanted to examine people's reactions to the modeled refusal before they themselves engaged in the questionable behavior. Before witnessing the peer refusal, Milgram's participants had already pressed the shock switches 14 times, including several times after hearing the learner's demand to be released. As described earlier, these decisions to continue make it increasingly difficult for the participant to stop.

Gender Differences

Milgram relied almost exclusively on male participants in his obedience studies. The one exception was a replication of the basic procedure in which women were used as participants. The women complied fully with the experimenter's commands 65% of the time, a rate identical to that in the comparable condition with men as participants. However, the obedient women reported higher levels of nervousness during the procedures than did the obedient men. Milgram (1974) speculated that the similarity between men's and women's behavior could be attributed to two opposing tendencies. On the one hand, women may have been more empathic toward the learner's suffering, which would have led them to end the procedure. On the other hand, women may have had a more difficult time asserting themselves in the face of the pressure exerted by the experimenter. Although gender differences do surface on occasion in obedience research, Blass (2000) found no evidence of a gender difference in eight out of nine conceptual replications of Milgram's studies he reviewed. Thus, although it seemed important to examine the effect of gender in this situation, I did not anticipate finding a gender difference.

Personality

Although the Milgram studies demonstrate the power of situational variables, it is also the case that some participants went along with the instructions whereas others did not. What can account for this difference? A case can be made that at least some of the variance in obedience can be explained by personality variables. Although few obedience studies have included personality measures, a handful of findings suggest personality may play a role (Blass, 1991). I identified two personality variables that seemed good candidates to affect obedience behavior.

First, I looked at individual differences in dispositional empathy (Davis, 1994). Participants in Milgram's studies were torn between wanting to follow the experimenter's orders and not wanting to harm the learner. It is reasonable to argue that when empathy for the learner's suffering is more powerful than the desire to obey the experimenter, participants are likely to refuse to continue. If that is the case, then individuals with a strong tendency to empathize with the suffering of others should be less likely to obey the experimenter's commands than should those low on this personality trait.

Second, I examined individual differences in the extent to which participants were motivated to control events. Milgram's research demonstrated that standing up to an authority figure is difficult. Moreover, whether participants accepted responsibility for their actions played an important role in their decisions to continue or to stop the experiment. I speculated that individuals who were motivated to exercise control and to make their own decisions would be more likely to disobey the experimenter than would those who were less motivated to feel in control. Although no study to date has examined the effect of desire for control on obedience, one set of studies found that participants with a high desire for control were less likely than those low in desire for control to conform to a perceived norm (Burger, 1987).

Method

Participants

Individuals who responded to advertisements and flyers went through a series of screening procedures. As described below, these procedures resulted in a final sample of 29 men and 41 women. Participants' ages ranged from 20 to 81 years, and the mean age was 42.9 years ($SD = 15.67$, $Mdn = 41$). Information about education and ethnicity are presented in Table 1.

Procedure

Recruitment and screening. Participants were recruited through advertisements in the local newspaper and in an online listing service. In addition, flyers were distributed at libraries, farmer's markets, coffee shops, and community centers. The message in these ads and flyers was patterned after Milgram's recruitment notices. Participants were promised \$50 for two 45-min sessions. Interested individuals were instructed to provide their names and telephone numbers by either calling a phone number established for the study or sending the information to an e-mail address set up for the study.

People responding to the ads or flyers were phoned by a research assistant, who conducted the initial screening procedure. Participants were first asked if they had been to college and, if so, if they had taken any psychology classes. The purpose of these questions was to screen out individuals who might be familiar with Milgram's obedience research. People who had taken more than two psychology classes were excluded from the study. Individuals not excluded at that point were asked the following six questions: "Have you ever been diagnosed with a psychiatric disorder? Are you currently receiving psychotherapy? Are you currently taking any medications for emotional difficulties such as anxiety or depression? Do you have any medical conditions that might be affected by stress? Have you ever had any problems with alcohol or drug use? Have you ever experienced serious trauma, such as child abuse, domestic violence, or combat?" The questions were created by the two clinical psychologists who conducted the second screening procedure. As per the clinicians' judgment, the research assistant excluded anyone who answered yes to any of the questions. Approximately 30% of the individuals who responded to the ads or flyers were excluded during the initial screening.

Those who made it through the first screening were scheduled for a second screening session held on the Santa Clara University campus. Upon arrival, participants were given a series of scales to complete. These included, in order, a demographic sheet asking about age, occupation, education, and ethnicity; the Interpersonal Reactivity Index (Davis, 1983); the Beck Anxiety Inventory (Beck, Epstein, Brown, & Steer, 1988); the Desirability of Control Scale (Burger & Cooper, 1979); and the Beck Depression Inventory (Beck, 1972).

The Interpersonal Reactivity Index is a 28-item self-report inventory designed to measure dispositional empathy. The scale assesses four kinds of empathy, although I was primarily interested in the Empathic Concern subscale, which measures "the tendency to experience feelings of sympathy and compassion for unfortunate others" (Davis, 1994, p. 57). Research has produced evidence for the internal consistency, test-retest reliability, and validity of the scale (Davis, 1994). The Beck Anxiety Inventory is a 21-item self-report scale designed to measure severity of anxiety. Test takers indicate on 4-point scales the extent to which they have experienced each of 21 anxiety symptoms during the previous week. Studies have found evidence for good internal consistency and validity (Beck, Epstein, et al., 1988; Steer & Beck, 1997). The Desirability of Control Scale is a 20-item self-report inventory designed to measure the extent to which test takers are motivated to see themselves in control of the events in their lives. Researchers have found evidence for good internal consistency and test-retest reliability for the scale, and the scale has been used to predict a number of behaviors related to desire for control (Burger, 1992). The Beck Depression Inventory is a 21-item self-report measure designed to assess severity of depression. Test takers indicate on 4-point scales the extent to which they have experienced each of 21 depression symptoms in the previous week. Extensive use of the

Table 1
Education and Ethnicity of Participants

Education and ethnicity	<i>n</i>	%
Education		
High school or less	12	17.1
Some college	16	22.9
Bachelor's degree	28	40.0
Master's degree	14	20.0
Ethnicity		
White Caucasian	38	54.3
Asian	13	18.6
Latin/Hispanic	9	12.9
Indian (Asian)	6	8.6
African American	3	4.3
Did not state	1	1.4

inventory has produced evidence of good reliability and validity (Beck, Steer, & Garbin, 1988).

Upon a participant's completion of the scales, a research assistant escorted the participant to a separate room, where the participant was interviewed by a licensed clinical psychologist. The assistant also provided the clinical psychologist with the completed anxiety and depression inventories. The clinical psychologist was fully aware of the experimental procedures and was instructed to err on the side of caution in identifying anyone who might have a negative reaction to participating in the study. The psychologist used responses on the anxiety and depression inventories along with responses to a semi-structured interview to make this assessment. The interview was structured through use of the Mini International Neuropsychiatric Interview (MINI) procedure (Lecrubier et al., 1997; Sheehan et al., 1997). The MINI procedure briefly touches on 22 categories of psychological disorders, and follow-up questions for each category are asked as needed. The psychologist was given as long as she needed to make a determination. The interviews lasted, on average, about 30 min. Of the 123 people who participated in this second screening process, 47 (38.2%) were excluded from the study by the clinical psychologist. These excluded individuals were given their full participation payment. Because participants were promised confidentiality by the interviewers, the specific reasons for these exclusions are not known.

The remaining 76 participants were scheduled for a second on-campus session, typically held about a week after the clinical interview. Six of these participants were dropped from the study. One did not return for the second session, and five expressed awareness of Milgram's obedience research at some point during the second session. These knowledgeable participants were fully debriefed and paid.

Base condition. Participants were randomly assigned to one of two conditions, but an attempt was made to keep the gender ratios in the conditions approximately equal. Upon arrival, participants assigned to the base condition were escorted by a research assistant into the lab room and introduced to the experimenter. The experimenter was a White Caucasian man in his mid-30s. Approximately one minute later, the research assistant escorted the confederate into the lab room. The confederate was a White Caucasian male in his 50s. Both the experimenter and the confederate were selected in part because of their resemblance to the experimenter and confederate used by Milgram. At this point, the experimenter gave the participant and the confederate each a \$50 bill for their participation and said the money was theirs to keep even if they chose to end their participation at any time.

Using a script taken largely from the one used in Milgram's research, the experimenter explained that the research was concerned with the effects of punishment on learning. He then informed participants that one would play the role of teacher and the other the role of learner and that the roles would be decided through a random drawing. The drawing was rigged so that the participant was always the teacher and the confederate was always the learner. The participant and the confederate were then asked to read and

sign consent forms. The consent form informed participants that they could end their participation at any time and still keep their money and that they might be videotaped during the study. If participants asked whether they were being videotaped, the experimenter told them they should assume they were being taped. In fact, participants were being videotaped by two hidden cameras.

The experimenter then escorted both individuals to the room next door, which contained a chair facing a table. An intercom box and a box with a set of four switches sat on the table. The experimenter asked the confederate to sit in the chair and then secured the confederate's arms to the wooden armrests with nylon straps. The experimenter explained the step was necessary "to prevent excessive movement during the experiment." The experimenter attached an electrode to the confederate's left wrist and added electrode paste, explaining that the paste would "provide a good contact to prevent blisters or burns." The experimenter also pointed out that the electrode was connected to the shock generator in the adjacent room. The participant stood a few feet away throughout this process.

The experimenter then placed a list of sample problems in front of the confederate and gave the confederate instructions patterned after those used by Milgram. The confederate was told to try to remember each of 25 word pairs read by the experimenter (e.g., *strong-arm*) and that during the test the experimenter would read the first word in each pair followed by four other words. The confederate's task was to indicate by pressing one of four buttons which of the option words was paired with the word on the original list. The experimenter also explained that incorrect answers would result in an electric shock and that the test would continue until all 25 word pairs were learned.

At this point, the confederate said, "I think I should say this. When I was in Kaiser Hospital a few years ago, they detected a slight heart condition. Nothing serious. But as long as I'm getting these shocks—how strong are they? How dangerous are they?" The experimenter answered, "While the shocks may be painful, they're not dangerous."

The experimenter and the participant returned to the experiment room, with the experimenter closing the door to the learner's room and the lab room behind him. The participant was seated in front of the shock generator, which was built to match the one used in Milgram's research. The machine was approximately 3 feet long, 15 inches high, and 16 inches deep (91.4 cm × 38.1 cm × 40.6 cm). A label in the upper left-hand corner of the box contained the words "Shock Generator, Type ZLB, Dyson Instrument Company, Waltham, Mass. Output 15 volts – 450 Volts." There were 30 switches across the face of the machine, with accompanying volt indicators ranging from 15 to 450 in 15-volt increments. The table with the shock generator was on the other side of the wall where the confederate sat.

The experimenter gave the participant instructions on how to press the switches, demonstrating how the switch remained in the middle position after being pressed to show which switches had been used. As in Milgram's studies, the experimenter also gave the participant a sample shock if the participant agreed (only two declined). The experimenter

then handed the participant a list of word pairs and reviewed the instructions with him or her. The word pairs and response options were identical to those used in Milgram's studies. The experimenter then turned the page over to reveal the response options. Each of the 25 words that began the word pairs was presented with four numbered response options. The experimenter explained that the participant should read the first word and the four options. The experimenter also said, "If he gives the wrong answer, you say, 'wrong' and then tell him the number of volts you are going to administer. Administer the punishment, read him the correct word pair once, and then go on to the next line." After reviewing the procedure one more time, the experimenter said, "Each time he gives a wrong answer, you move up one switch on the shock generator. It's important that you follow the procedure exactly."

The experimenter then pretended to turn on the intercom system (which was always on). He explained that it was a one-way system such that the confederate could hear what the participant said but could not respond through the intercom. The experimenter sat at a table approximately four feet (1.2 m) behind and a few feet to the right of the participant. The experimenter assisted participants who had difficulty with the procedures (which was often the case with the first few items) but otherwise limited comments to those described below.

The confederate responded to each test item by pressing one of the buttons on the switch box that lit up either 1, 2, 3, or 4 on the answer box in the adjacent room. The confederate gave correct and incorrect responses according to the same predetermined schedule used in Milgram's research. A computer inside the shock generator was attached to a speaker on the confederate's side of the wall. The computer was programmed so that prerecorded sounds emanated from the speaker when certain switches were pressed. The effect was that participants thought they were listening to the confederate's slightly muffled but clear voice through the wall. After participants pressed the 75-volt switch, they heard a small grunt (what Milgram described as "ugh"). Participants heard the confederate grunt after each successive shock, and each grunt was a little louder than the previous one. When the participant pressed the 150-volt switch, he or she heard the confederate yell through the wall, "Ugh. That's all. Get me out of here. I told you I had heart trouble. My heart's starting to bother me now. Get me out of here, please. My heart's starting to bother me. I refuse to go on. Let me out."

The experimenter was instructed to end the experiment when the participant either (a) refused to continue after hearing all the experimenter's prods or (b) read the next item on the test after hearing the confederate's protests following the participant's press of the 150-volt switch. If at any point during the learning task the participant expressed either verbally or nonverbally a reluctance to continue the study, the experimenter gave one of four prods, in this order: "Please continue" or "Please go on"; "The experiment requires that you continue"; "It is absolutely essential that you continue"; and "You have no other choice, you must continue." The prods were given in sequence each time the participant hesi-

tated to continue, beginning with the first prod. Thus, if the participant responded to the first prod with more reluctance to continue, the experimenter gave the second prod, and so forth. If the participant continued to express reluctance after the fourth prod, the experimenter ended the study. No participant was allowed to press any more switches after pressing the 150-volt switch.

Again following Milgram's procedures, the experimenter gave predetermined answers to specific questions by the participant. If the participant asked whether the learner would suffer physical damage from the shocks, the experimenter said, "While the shocks may be painful, there is no permanent tissue damage." If the participant said that the learner did not want to continue, the experimenter said, "Whether the learner likes it or not, you must go on until he has learned all the word pairs correctly." If the participant asked who was responsible for any harm to the learner, the experimenter said, "I am responsible." If the participant asked about the money he or she was given for participation or wanted to give back the money, the experimenter said, "The money is yours to keep regardless." Any of these specific questions also was considered an indication that the participant did not want to continue. To make the answer a response to the participant's question or comment and to keep from sounding repetitive, the experimenter sometimes varied the wording of these statements slightly.

As soon as the experimenter announced that the study was over, he told the participant that the shock generator was not real and that the confederate was not receiving electric shocks. The confederate entered the lab room at that point to assure the participant that he was fine. After the experimenter determined that the participant understood the situation, the participant was escorted to a nearby room, where the principal investigator conducted a thorough debriefing.

Modeled refusal condition. Participants assigned to the modeled refusal condition went through the same procedures as participants in the base condition, with a few exceptions. First, two confederates were used in the modeled refusal condition. One of the confederates was the same man who played the learner in the base condition. The other confederate, also posing as a participant, was of the same gender as the real participant. For the women, this second confederate was a White Caucasian woman in her late 20s. For the men, the confederate was a White Caucasian man in his mid-30s. The drawing was rigged so that the participant was assigned the role of Teacher 2 and the new confederate was assigned the role of Teacher 1. Both teachers watched the learner being strapped into the chair, and both were given a sample electric shock (the confederate went first). Second, the experimenter instructed both teachers to sit next to one another in front of the shock generator, with the confederate on the left and the real participant on the right. The experimenter explained that Teacher 1 would begin the procedure. If the participant asked what Teacher 2 was supposed to do, the experimenter said he would explain that later. Third, the confederate posing as Teacher 1 read the list of word pairs and began administering the test and pressing the switches. The

confederate showed no signs of hesitation until hearing the learner's "ugh!" after pressing the 75-volt switch. At that point, the confederate paused for a few seconds before continuing. After pressing the 90-volt switch and hearing another "ugh!", the confederate glanced at the experimenter and said, "I don't know about this." The experimenter responded with his initial prod, "Please continue." The confederate paused a few seconds, then said, "I don't think I can do this," and pushed his or her chair a few inches back from the table. The experimenter then asked the real participant to continue the test, picking up where the other teacher had left off. The confederate sat silently throughout the rest of the study and avoided making eye contact with the participant.

Results

The percentage of participants who continued the procedure after pressing the 150-volt switch was examined. As shown in Table 2, 70% of the base condition participants continued with the next item on the test and had to be stopped by the experimenter. This rate is slightly lower than the percentage who continued beyond this point in Milgram's comparable condition (82.5%), although the difference fell short of statistical significance, $\chi^2(1, N = 80) = 1.10, p = .29$.¹ Contrary to expectation, the percentage of participants in the modeled refusal condition who continued past the 150-volt point (63.3%) also was not significantly different from the percentage who did so in the base condition, $\chi^2(1, N = 70) = 0.11, p = .74$.

I also looked at when in the procedure the participant received the first prod from the experimenter. This was the point at which the participant first expressed concern about continuing the procedure. Participants were assigned a value from 1 to 12 depending on the last switch they pressed before receiving a prod, with 1 = *pressed no switches*, 2 = *after pressing the 15-volt switch*, and so forth. Participants who received no prods were assigned a value of 12. I compared the two conditions on this measure. Participants in the base condition ($M = 7.65$) received a prod from the experimenter significantly earlier than participants in the modeled refusal condition ($M = 9.56$), $t(68) = 2.75, p = .008, d = .64$. However, it should be noted that participants in the modeled refusal condition did

Table 3
Numbers (and Percentages) of Participants Who Stopped and Who Continued, by Gender

Condition and behavior	Men	Women
Base condition		
Stopped at 150 volts or earlier	6 (33.3)	6 (27.3)
Continued after 150 volts	12 (66.7)	16 (72.7)
Modeled refusal condition		
Stopped at 150 volts or earlier	5 (45.5)	6 (31.6)
Continued after 150 volts	6 (54.5)	13 (68.4)

not have an opportunity to express reluctance until after the 90-volt switch was pressed (a score of 7 on the 12-point scale). Not surprisingly, participants who refused to continue ($M = 6.52$) had lower first-prod scores than those who continued until stopped by the experimenter ($M = 9.43$), $t(68) = 4.20, p = .001, d = .96$.

As shown in Table 3, there was little difference in obedience rates between men and women. Although women were slightly more likely than men to continue in both the base condition and the modeled refusal condition, the differences were not significant in either condition, $\chi^2(1, N = 40) = 0.01, p = .95$ and $\chi^2(1, N = 30) = 1.35, p = .74$, respectively, or when the conditions were combined, $\chi^2(1, N = 70) = 0.25, p = .62$. In addition, the first-prod score for the men ($M = 8.34$) was not significantly different from the first-prod score for the women ($M = 8.56$), $t(68) = 0.29, p = .77$.

I also was interested in the relation between participants' rates of obedience and scores on two personality trait measures—empathic concern and desire for control.² Three participants in the base condition did not complete the personality test items on either test (they failed to notice items on the back side of the test) and thus could not be used in these analyses. I compared the personality scores of those who continued with the procedure with the scores of those who stopped. As shown in Table 4, no significant

Table 2
Numbers (and Percentages) of Participants Who Stopped and Who Continued

Behavior	Base condition	Modeled refusal condition	Milgram's Experiment 5
Stopped at 150 volts or earlier	12 (30.0)	11 (36.7)	7 (17.5)
Continued after 150 volts	28 (70.0)	19 (63.3)	33 (82.5)

¹ The failure to find a significant difference cannot be easily attributed to a lack of power. The sample size of 80 is within the ballpark for Cohen's (1992) recommendation for finding a medium-sized effect and is much larger than the suggested number for a large effect.

² Because the entire Interpersonal Reactivity Scale was administered, scores for the three other subscales were also obtained: Perspective Taking (the tendency to adopt another person's point of view), Personal Distress (the tendency to be uncomfortable seeing others in stress), and Fantasy (the tendency to put oneself in imaginary situations). I compared the subscale scores between those who stopped and those who did not, and I looked at the correlation between the subscale scores and the first-prod scores. I had no theoretical reason to expect that significant effects would emerge in any of these analyses, and indeed they did not.

differences were found for either the empathic concern score, $t(65) = 1.13, p = .26$, or the desire for control score, $t(65) = 0.90, p = .37$. I also compared the average personality score of continuers and stoppers within each of the two conditions. As can be seen in Table 4, the empathic concern scores of continuers and stoppers did not differ in either the base condition, $t(35) = 0.03, p = .97$ or the modeled refusal condition, $t(28) = 1.63, p = .11$. In the base condition, participants who stopped had significantly higher scores on desire for control than did those who continued, $t(35) = 2.06, p = .05, d = .71$. However, in the modeled refusal condition, participants who continued and participants who stopped did not differ on desire for control, $t(28) = 0.88, p = .39$.

I examined the correlations between scores on the two personality scales and participants' first-prod scores. As shown in Table 5, high empathy scores were related to lower first-prod scores in the modeled refusal condition and when scores for the two conditions were combined. Participants with a high desire for control tended to show reluctance earlier than did participants with a low desire for control, but only in the base condition. I also examined the relationship between the obedience variables (whether participants stopped, first-prod scores) and age, ethnicity, and education. No significant effects emerged in any of these analyses.

Finally, I compared the excluded participants with those who participated in the full experiment on all of the measures and information I had for both groups. I found no difference between the two groups in terms of gender, $\chi^2(1, N = 117) = 1.13, p = .29$; education, $\chi^2(3, N = 117) = 3.96, p = .27$; age, $t(115) = 0.29, p = .77$; empathic concern, $t(112) = 1.89, p = .06$; or desire for control, $t(112) = 0.58, p = .56$. There were not enough participants in some of the ethnic groups to make meaningful comparisons between excluded and nonexcluded participants. But among those who gave their ethnicity (all except 7 people), the excluded individuals did not differ from those who participated in the second session in terms of the percentage of ethnic minority members, $\chi^2(1, N = 109) = 0.09, p = .76$.

Table 4
Mean Personality Scale Scores

Condition and behavior	Empathic Concern	Desire for Control
Base condition		
Continued	19.20 (4.64)	98.24 (12.22)
Stopped	19.25 (2.93)	106.92 (11.45)
Modeled resistance condition		
Continued	18.37 (4.19)	103.68 (11.30)
Stopped	20.91 (3.91)	99.64 (13.56)
Conditions combined		
Continued	18.84 (4.42)	100.59 (12.01)
Stopped	20.04 (3.46)	103.43 (12.77)

Note. Standard deviations appear in parentheses.

Table 5
Correlations Between First-Prod Measure and Personality Scale Scores

Personality scale	Base condition	Modeled refusal condition	Conditions combined
Empathic Concern	-.26	-.51**	-.32**
Desire for Control	-.33*	.24	-.12

* $p < .05$. ** $p < .01$.

Discussion

People learning about Milgram's (1963, 1965, 1974) obedience studies often ask whether similar results would be found today. Ethical concerns prevent researchers from providing a definitive answer to that question. But my partial replication of Milgram's procedure suggests that average Americans react to this laboratory situation today much the way they did 45 years ago. Although changes in societal attitudes can affect behavior, my findings indicate that the same situational factors that affected obedience in Milgram's participants still operate today. The similarity between my results and Milgram's is also noteworthy because of a few procedural changes I implemented that should have made it easier for the participants to resist authority. The participants were told explicitly and repeatedly that they could leave the study at any time and still keep their \$50. In addition, participants were aware that the experimenter had given this same assurance to the confederate. Indeed, several of the participants who stopped the procedure after hearing the learner's protests pointed out that the confederate had been promised he could stop when he wanted to.

I cannot say with absolute certainty that the present participants would have continued to the end of the shock generator's range at a rate similar to Milgram's participants. Only a full replication of Milgram's procedure can provide such an unequivocal conclusion. However, numerous studies have demonstrated the effect of incrementally larger requests. That research supports the assumption that most of the participants who continued past the 150-volt point would likely have continued to the 450-volt switch. Consistency needs and self-perception processes make it unlikely that many participants would have suddenly changed their behavior when progressing through each small step.

I anticipated that participants would be less obedient after seeing another person refuse to continue. In novel situations such as Milgram's obedience procedure, individuals most likely search for information about how they are supposed to act. In the base condition, the participants could rely only on the experimenter's behavior to determine whether continuing the procedure was appropriate. However, in the modeled refusal condition, participants saw not only that discontinuing was the option selected by

the one person they witnessed but that refusing to go on did not appear to result in negative consequences. Although a sample of one provides limited norm information, it is not uncommon for people to rely on single examples when drawing inferences. Nonetheless, seeing another person model refusal had no apparent effect on obedience levels in the present study. I interpret this high rate of obedience in the modeled refusal condition as a demonstration of the power of the situational forces leading participants to go along with the experimenter's instructions. One can speculate that a stronger message about normative behavior would have had a noticeable effect on participants' behavior. For example, if participants saw 10 out of 10 other individuals refuse to continue the procedures, the descriptive norm would be clear and the likelihood of going against this norm would be small. Milgram (1974) succeeded in lowering compliance when using only two resisting confederates, although those confederates refused to continue in a dramatic fashion.

I found no evidence for gender differences in obedience. Researchers have speculated that the tendency for women to be more concerned about the learner's plight might be offset by the tendency for women to be less assertive than men when standing up to the experimenter. This may well have been the case in the present study. Then again, the finding may simply reflect the tendency for situational variables to overpower individual differences in this setting. There was some evidence that personality traits were related to participants' reactions to the situation. However, the data were not entirely consistent or easily interpretable. Participants who were high in empathic concern expressed a reluctance to continue the procedure earlier than did those who were low on this trait. But this early reluctance did not translate into a greater likelihood of refusing to continue. This latter finding fails to support the notion that a lack of empathy explains the high obedience rates in Milgram's studies. Rather, the results again are in line with those who point to the power of situational variables to overcome feelings of reluctance in this situation. I also anticipated that a high desire for control would increase the likelihood that participants would act on their own feelings rather than obey the experimenter. However, this effect was found only in the base condition. It is not clear why the presence of the refusing model would undermine this tendency. One possibility is that the base condition may have represented more of a me-versus-him situation that consequently triggered a desire to assert personal control. In sum, although I found evidence that personality traits play a role in participants' responses to the situation, the relationship between personality and obedience remains speculative.

I did my best to replicate Milgram's procedures up to the 150-volt point. However, there were some intentional and some unavoidable differences between Milgram's procedures and mine that should be pointed out. It is difficult to know what effect the screening procedures had on the findings. For ethical reasons, I excluded people with a history of psychological or emotional problems and anyone the clinical psychologist deemed might have a negative

reaction to participating in the study. I also excluded people who had taken more than two college-level psychology classes. In addition, I recruited participants from all adult ages, whereas Milgram limited his participants to those age 50 and younger. I also had what was most likely a more ethnically diverse group of participants than did Milgram. On the other hand, great effort was expended to replicate the key features in Milgram's study (script, shock generator) as well as many minor features (experimenter's lab coat, words in memory task). This attention to detail probably goes beyond most efforts to replicate psychology studies. Although I cannot rule out all possible differences between the present sample and Milgram's that might have affected the findings, all of the data fail to identify any such difference. Participants excluded during the second phase of the screening did not differ from participants who were not excluded on any of the characteristics examined—age, ethnicity, gender, education, personality variables. Moreover, I found no effect for education, age, or ethnicity on participants' behavior in the study. In short, I am as confident as a psychology researcher can ever be that my findings can be legitimately compared with Milgram's.

Milgram's obedience studies have maintained a place in psychology classes and textbooks largely because of their implications for understanding the worst of human behaviors, such as atrocities, massacres, and genocide. Indeed, Milgram frequently drew inferences from his studies to account for the behavior of people who went along with the Holocaust. Although one must be cautious when making the leap from laboratory studies to complex social behaviors such as genocide, understanding the social psychological factors that contribute to people acting in unexpected and unsettling ways is important. Since Milgram's studies, concern for the well-being of participants has limited research on obedience to authority. I hope future investigators will utilize the 150-volt procedure presented here to address the weighty questions that motivated Stanley Milgram nearly half a century ago.³

³ Researchers interested in borrowing the shock generator used in this study should contact the author.

REFERENCES

- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3, 193–209.
- Baumrind, D. (1964). Some thoughts on ethics of research: After reading Milgram's "Behavioral Study of Obedience." *American Psychologist*, 19, 421–423.
- Beck, A. T. (1972). *Depression: Causes and treatments*. Philadelphia: University of Pennsylvania Press.
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56, 893–897.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77–100.
- Blass, T. (1991). Understanding behavior in the Milgram obedience experiment: The role of personality, situations, and their interactions. *Journal of Personality and Social Psychology*, 60, 398–413.
- Blass, T. (2000). The Milgram paradigm after 35 years: Some things we

- now know about obedience to authority. In T. Blass (Ed.), *Obedience to authority: Current perspectives on the Milgram paradigm* (pp. 35–59). Mahwah, NJ: Erlbaum.
- Blass, T. (2004). *The man who shocked the world: The life and legacy of Stanley Milgram*. New York: Basic Books.
- Burger, J. M. (1987). Desire for control and conformity to a perceived norm. *Journal of Personality and Social Psychology*, 53, 355–360.
- Burger, J. M. (1992). *Desire for control: Personality, social and clinical perspectives*. New York: Plenum.
- Burger, J. M. (1999). The foot-in-the-door compliance procedure: A multiple-process analysis and review. *Personality and Social Psychology Review*, 3, 303–325.
- Burger, J. M. (2002). Four decades and counting [Review of the book *Obedience to authority: Current perspectives on the Milgram paradigm*]. *Contemporary Psychology: APA Review of Books*, 47, 665–667.
- Burger, J. M., & Cooper, H. M. (1979). The desirability of control. *Motivation and Emotion*, 3, 381–393.
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591–621.
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 201–234). San Diego, CA: Academic Press.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126.
- Davis, M. H. (1994). *Empathy: A social psychological approach*. Dubuque, IA: Brown and Benchmark.
- Elms, A. C. (1995). Obedience in retrospect. *Journal of Social Issues*, 51, 21–31.
- Fischer, C. T. (1968). Ethical issues in the use of human subjects. *American Psychologist*, 23, 532.
- Gilbert, S. J. (1981). Another look at the Milgram obedience studies: The role of the graduated series of shocks. *Personality and Social Psychology Bulletin*, 7, 690–695.
- Gilovich, T., & Savitsky, K. (2002). Like goes with like: The role of representativeness in erroneous and pseudo-scientific beliefs. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 617–624). New York: Cambridge University Press.
- Kaufmann, H. (1967). The price of obedience and the price of knowledge. *American Psychologist*, 22, 321–322.
- Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., et al. (1997). The Mini International Neuropsychiatric Interview (MINI): A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*, 12, 224–231.
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371–378.
- Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations*, 18, 57–76.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Milgram, S. (1983). Reflections on Morelli's "Dilemma of Obedience." *Metaphilosophy*, 14, 190–194.
- Miller, A. G. (2004). What can the Milgram obedience experiments tell us about the Holocaust? Generalizing from the social psychology laboratory. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 193–239). New York: Guilford Press.
- Miller, A. G., Collins, B. E., & Brief, D. E. (Eds.) (1995). Perspectives on obedience to authority: The legacy of the Milgram experiments [Special issue]. *Journal of Social Issues*, 51(3).
- Mixon, D. (1972). Instead of deception. *Journal for the Theory of Social Behavior*, 2, 145–177.
- Morelli, M. F. (1983). Milgram's dilemma of obedience. *Metaphilosophy*, 14, 183–189.
- Sheehan, D. V., Lucrubier, Y., Sheehan, K. H., Janavs, J., Weiller, E., Keskiner, A., et al. (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*, 12, 232–241.
- Steer, R. A., & Beck, A. T. (1997). Beck Anxiety Inventory. In C. P. Zalaquett & R. J. Wood (Eds.), *Evaluating stress: A book of resources* (pp. 23–40). Lanham, MD: Scarecrow Press.