As a student, how do you conceive of the work you do in school? What role does it play in your world, and is what you have to say important? These are the sorts of questions, I imagine, that Anna was grappling with when she wrote her paper. When you read it in this anthology, you’ll be reading words that have been worked over again and again. I’m not sure exactly how many drafts Anna produced, but there were at least ten, and probably more like fifteen. You’ll easily be able to pick out her topic and her argument, but what might be more difficult to see, after all the revisions, are the questions that caused Anna to choose this topic in the first place.

I remember a day in class, early in the quarter, probably before the first assignments had been handed in. We were having a discussion about what makes a good argument, and one of the ideas I put forward was that originality is central to a good essay. This is a vague idea, but it’s one I take for granted. If what you have to say is self-evident, or if any number of bright people can easily come to the same conclusion, then what’s the point in saying it? As soon as I had made this point, Anna frowned. After about half a minute, she shot up her hand, and asked, “does that mean we should sacrifice saying something right for saying something that’s controversial or different or,” and here she really skewered me, “just plain weird?”

Of course it doesn’t, but it can be hard for anyone to see how their writing can say something both insightful and right. What Anna was hitting at, I think, is that fundamental need we all feel to know that our work is righteous. When you write for a class, you’re often just writing for a few select people: your fellow students, perhaps, and certainly your professor. PWR essays aren’t presidential addresses or Supreme Court decisions. For the most part, they won’t be read by thousands of people. At best, I think, student essays are written in a subjunctive mode. You’re best served by writing as though you were able to reach an audience of millions, as though your words were able to change the world. There’s a quiet dignity to this, to writing with purpose while developing your craft before you unleash yourself upon a broader audience, and this is what Anna does particularly well, I think.

She chose a topic that she’d heard about in an intro psychology class and she researched the bejesus out of it. Anna’s the sort of student who, when I mentioned that Hannah Arendt’s *Eichmann in Jerusalem* was a book that would be worth reading for her essay, bought and read it in a matter of days. Most importantly, though, she didn’t take on some niggling issue, but one of the most important and problematic psychological studies ever performed. Some students would be overwhelmed by a topic such as the Milgram experiments, but what pulled Anna through was what I like best about her work, a desire to engage the world upon first principles, to cut to what matters most.

—Stephen Osadetz
In Defense of Human Agency
Protection of Self-Image in the Milgram Obedience Experiments

Anna Grummon

Picture the following scenario: Gerald Braverman, a thirty-nine-year-old social worker from New Haven, CT, is going to participate in a psychology study at Yale. Braverman, who, with a furrowed brow and graying hair, looks older than his years, read of the experiment in the local newspaper. A psychologist at Yale, Stanley Milgram, is recruiting participants for this study on learning and memory, and promises them $4.00 for participating. Gerald drives to the Interaction Laboratory on the university campus, enters the lab, and quietly observes the impressive electronic equipment lining the rooms, the heavy curtains draping the windows, and the stern researchers striding through the halls. He is greeted by Mr. Henry Williams, a stony experimenter donning a gray technician’s coat. The atmosphere is one of efficiency, of experiments being conducted by prestigious researchers, and Braverman feels repressed, his eyes darting to take in the scene, his speech attempting to affect composure.

Williams introduces Braverman to another participant in the study, Mr. John Wallace. The man appears to be in his late forties; he wears thick glasses and a striped tie, and smiles warmly at Braverman as they shake hands. Williams begins his instructions to the participants. He is intense, his voice confident, his speech articulate and harsh. “Psychologists have developed several theories to explain how people learn various types of material. Some of the better-known theories are treated in this book.” He points to an impressive looking volume, *Teaching and Learning Processes*, which rests on the table behind him. “But actually, we know very little about the effect of punishment on learning, because almost no truly scientific studies have been made of it in human beings. We want to find out just what effect different people have on each other as teachers and learners, and also what effect punishment will have on learning in this situation. In this study we are bringing together a number of adults of different occupations and ages, and we’re asking some of them to be teachers and some to be learners” (Milgram, *Obedience*, 34-6). The experimenter explains that the learner will attempt to memorize a series of word-pairings; the teacher will both help him with this task and punish him with an electric shock for incorrect answers. The teacher will increase the voltage of the shock with each incorrect response.

Williams continues, “Please draw a slip of paper to determine which duty you are to perform.” He pulls out two slips of paper from his coat pocket and offers them to the gentlemen seated before him. Braverman reaches for a slip. What he doesn’t know is that the entire experimental set up is a fake, a cover story to hide the true intent of the study. The two other men are both trained actors, and the lab equipment is merely a set of stage props, carefully crafted by Milgram to look and feel real. Wallace will not actually receive
any shocks—but he has recorded a tape of himself feigning agony, which will be played at specified intervals in the experiment. The drawing has been rigged so that Braverman will be the teacher and Wallace the learner.

Following the drawing, the experimenter leads the pair into a nearby sparsely furnished room with a large, eerie electric apparatus set up on a table near the wall. He straps Wallace into a chair in front of the machine and explains that these straps are to prevent excessive movement during the shocks. He applies a paste to Wallace’s arm, explaining that this will help avoid blisters and burns from the shocks. Wallace, looking nervous, asks if the shocks will hurt him. “While the shocks can be extremely painful,” Williams curtly replies, “they cause no permanent tissue damage.” Braverman’s eyes sweep the room several times, taking in the scene, before he is led into an adjacent room by the experimenter.

The next room is separated from the first by an opaque, partially soundproof wall; Braverman can no longer see the other participant, but can speak to him through a microphone system. Upon entrance to the room, Braverman is introduced to a shock generator, an impressive construction, complete with 30 different levers, each with a corresponding voltage and a qualitative label ranging from “slight shock” to “danger: severe shock” (figure 2). The last two levels simply read: XXX. Braverman receives instructions on how to use the machine, and receives a mild sample shock so that he understands the pain the learner will experience. Throughout, the experimenter speaks rapidly and with a chilling austerity.

Braverman agrees to deliver the painful shocks to the stranger in the next room. And so the experiment begins. At first, Wallace answers all the questions correctly, and his teacher continues through the procedures almost mindlessly. When Wallace gives his first incorrect answer, Braverman quickly delivers the first shock, and when he continues to deliver wrong responses, Braverman thoughtlessly increases the voltage. Then he administers the 75V shock and, through the semi-soundproof wall, hears his victim grunt. The cycle continues: more incorrect answers followed by increasingly painful shocks. Soon, Wallace is crying out in agony: “I can’t stand the pain! Let me out of here!” ...
Experimenter! Get me out of here! I’ve had enough!” Braverman hesitates, asking if he should continue. He begins to laugh nervously, his hands now clenched, now rubbing his forehead. The experimenter responds immediately. “The experiment requires that you continue,” he prods, repeating as necessary when his participant stalls.

What will Braverman do—break off the experiment, saving his fellow participant from this torture, or continue to obey orders despite the tormented pleas? According to Milgram, if Braverman is anything like the majority of Americans, he’ll obey the experimenter’s orders. In the early 1960s, Milgram recruited forty men like Braverman to participate the study described above. A startling 62.5% of his participants administered what they thought were painful and potentially lethal shocks to a complete stranger, simply because they had been ordered to do so.

From the moment Milgram first published his results, writers in all arenas—from the popular press to professional and academic journals—gave considerable attention to the studies (Ferzagllich, 108). Many of these writers were horrified and disturbed by the intense suffering the participants were willing to inflict when they were ordered to do so by an authority figure. Some even saw Milgram’s experiments as evidence of mankind’s inherently evil nature. However, many scholars have published convincing accounts rejecting this notion and instead arguing that participants in Milgram’s studies were obedient not because they were sadistic, but rather because they were influenced by a number of aspects of their particular setting. Factors such as the participant’s politeness and a desire to uphold his initial promise to participate in the study acted as “binding forces” that locked the participant in the situation, making it difficult to disobey (Milgram, Obedience, 7). Scholars supporting this view argue that the setting itself drove the participants to obedience; this phenomenon is typically referred to by social psychologists as the “power of the situation.” This stance has dominated the discussion of Milgram’s experiments for decades; indeed, this study is often considered the classic example of how profoundly the environment can influence behavior. What’s more, this notion is a crucial assumption of the entire field of social psychology, which is typically defined as the scientific study of how our thoughts, feelings and behaviors are influenced by the presence of real, imagined or implied others—in short, the study of how our social environments shape how we act (Allport, 1985).

Even so, in the case of Milgram’s experiments, this power-of-the-situation explanation is incomplete, lacking a unified theory that explains how the situational elements exert their influence. This gap can be filled when one understands that humans desire to protect their self-images and will put this need above the needs of others. Egoism provides a framework to explain the effects of the various “binding factors” in Milgram’s experimenters. Indeed, we can wholly understand the obedient behavior of Milgram’s participants as a means of protecting their self-images. Put differently, on the most fundamental level, the driving factor behind obedient behavior in the Milgram experiments was the participants’ desire to maintain a positive self-concept.

Viewing Milgram’s participants’ behavior as egoistic, rather than simply the result of situational factors, has frightening implications for how we are to interpret their behavior, and consequently for our understanding of human nature. The power-of-the-situation argument is an important component of any explanation of the experiments. It asks us to direct careful attention to Milgram’s experimental design and to take seriously its potential influence on subjects’ behavior. Yet committing to an explanation based solely
on situational factors yields difficult implications; if we strip agency and choice from the participants in Milgram’s experiments, must we similarly remove responsibility from all who commit violent or destructive acts under the orders of an authority? What are we to make of the horrific instances of violence that taint the world’s history—of Abu Ghraib, Nazi Germany, and My Lai? It is here that the traditional power-of-the-situation explanation distorts the reality of human behavior. When we place all of the blame for participants’ obedience on the situation, we ignore the fact that humans have free will. But with a framework of egoism, we can restore responsibility to the participants: we see that the situation has power only because it interacts with a fundamental human obsession with self-image protection. If we view Milgram’s experiments though this lens, it is clear that man’s inherent tendency toward evil is still at stake, and that we must consider the participants’ complicity in the horrifying events of experiment.

The Experimental Design

“The experiments of Stanley Milgram on obedience to authority have achieved a truly remarkable visibility, one that is rare in the social sciences.”

The scene above detailed the first of twenty-four variations of the obedience studies Milgram ran in the 1960s. Each variation differed in certain aspects—the distance between the teacher and the experimenter, for example—but all followed a general pattern. The experimental design described here relates the general pattern, and is drawn from Milgram’s initial (1963) publication of the study. (Descriptions of other variations, as well as a more detailed account of the first variation, can be found in Milgram’s [1974] monograph *Obedience to Authority*).

In the experiment, a male participant between the ages of 20 and 50 entered a Yale laboratory, believing he was participating in a study of learning and memory. Participants varied in occupation (ranging from postal clerks to engineers to salesmen) and educational background (some had not completed high school and others had doctoral degrees). As in the scene above, this was merely a cover study invented by Milgram to hide the real interest of the study: obedience to authority. In each trial, the subject was “randomly assigned” to be the teacher (the assignment was rigged); the other man in the room, an actor and accomplice of Milgram’s, was made the learner. Following assignment, the subject and the learner were led into separate rooms for the remainder of the experiment (*figure 3*); the rooms, they were
told, were separated by a “partially soundproof” wall. The subject was then introduced to a shock-generator and instructed to deliver shocks of increasing voltage each time the learner answered a question incorrectly.

As in the opening scenario, when the testing began, the learner gave a set of pre-determined responses to the questions, beginning with mostly correct answers, but gradually lapsing into a long streak of incorrect responses, prompting the subject to give him increasingly painful shocks. He remained silent until 75V, when he groaned. As the voltage increased, the learner began to verbally complain of the pain, eventually demanding that he be let out of the experiment. By 300V, he stopped responding altogether. If at any point the learner’s ostensible suffering made the subject pause or ask to discontinue, the experimenter prodded him, urging, “please continue” or, if necessary, “you have no other choice, you must go on” (Milgram, Obedience, 37). Like the learner, the experimenter was also an actor, trained by Milgram to be impassive and stern (Milgram, Obedience, 16), to loom threateningly over the participant and verbally pressure the participants to continue their task.

Milgram’s Findings

“Obedience rather than disobedience can be understood as the expected outcome.”
—John M. Darley

“Such an experimental result … ought to give each of us pause as no other single bit of research has.” –Philip Zimbardo

In reporting his results, Milgram focused on three issues. First—and foremost—he was interested in the maximum shock each subject administered before he actively refused to continue and exited the experiment. He deemed that the subjects were “obedient” when they administered shocks through the highest voltage level; those who broke off at any point before the highest level were “defiant.” He reported 26 of his 40 subjects (65%) were “obedient;” 14 were labeled as “defiant.” As a second point of interest, Milgram argued that the subjects were convinced of the reality of the experimental set-up—that is, the subjects were deceived by his cover story. Finally, Milgram noted the extreme emotional distress the participants endured. He described the subjects’ “bizarre nervous laughter, uncontrollable seizures, sweating, trembling, stutters, lip biting, and groaning”—in short, the subjects exhibited a “degree of tension…rarely seen in sociopsychological laboratory studies” (375). According to Milgram, this visible emotional stress reveals the tension the participants felt between their desire to be obedient and their wish to avoid harming innocent people.

To scholars and lay people alike, this obedience was disturbing, and entirely unpredictable. For example, prior to the experiment, numerous groups—including college students, psychiatrists, middle-class adults of various occupations, and graduate students and faculty in the behavioral sciences—cast predictions about the subjects’ behavior. Together, these groups predicted that about one subject in one thousand would administer the highest-level shock (Milgram, Obedience, 31). These predictions are understandable—to obey, participants had to glaringly breach what Milgram argued was a fundamental ethical proscription; as Milgram writes, “of all moral principles, the one
that comes closest to being universally accepted is this: one should not inflict suffer-
ing on a helpless person who is neither harmful nor threatening to oneself” (Obedience, 13). But what are we to make of this surprising tendency to abandon morals? Some have argued that the Milgram experiments are evidence of a dark conception of hu-
man nature. Milgram himself noted that, “A commonly offered explanation is that those who shocked the victim at the most severe level were monsters, the sadistic fringe of society” (Obedience, 5). Likewise, the popular press, for example, frequently portrayed the obedient subjects as patently wicked, describing Milgram’s experiments with article titles such as “Experiment Found Plenty of Sadists” (23).

This explanation falls under a general human tendency to attribute evil behavior, such as that displayed by Milgram’s obedient subjects, to evil personalities. Humans intuitively believe that an evildoer must contain some element of evil, a certain “qual-
ity of evilness” or “kernel of evil” almost like a physical characteristic (Darley, 202). A headline for an article in Science Digest, for example, asked its readers if Milgram’s study demonstrated that “the seeds of such slavish inhumanity exist in all of us?” (qtd. in Fer-
maglich, 109). But is this popular view correct? If it were, one would expect that we could see some sort of measurable personality differences between the obedient and defiant participants—specifically, that the obedient participants were more malevolent than their disobedient peers. However, when Milgram and Alan Elms compared par-
ticipants’ scores on a series of standardized personality tests, they found no differences between defiant and obedient individuals (282). Likewise, many scholars have argued that evil doers are by no means necessarily evil people (e.g. Staub; Milgram, Obedience). Political theorist Hannah Arendt, for example, raised the question of whether evil acts are simply the result of the tendency of ordinary people to mindlessly obey orders. Ar-
endt reported for The New Yorker on the trial of Adolf Eichmann, a high-ranking Nazi and influential architect of the Holocaust; her articles evolved into a book in which she argues that Eichmann’s evil acts were utterly banal: “The trouble with Eichmann was precisely that so many were like him, and that many were neither perverted nor sadistic, that they were, and still are, terribly and terrifyingly normal” (276). Scholars like Arendt discredit the idea that evildoers—from Nazi criminals to Milgram’s participants—were sadistic and somehow different from common people.

Rather than arguing that a “kernel of evil” within an individual produces horrifying behavior, many of these scholars take a power-of-the-situation stance, arguing that evil deeds such as those of Milgram’s participants are not the volitional products of sadistic individuals, but rather the result of the convergence of social and environmental factors. Scholars have frequently applied this power-of-the-situation argument to Milgram’s ex-
periments. Thomas Blass, for example, notes that, “The findings of the obedience studies have been held up as examples, par excellence, of the controlling power of the situation” (1991, 399), and researchers have identified a profusion of environmental factors be-
lieved to have contributed to the high levels of obedience Milgram achieved. Scholars have argued the importance of various features of the lab setting such as the ambiguous nature of the situation, the prestige of the laboratory, the ostensible legitimacy of the experimenter, and the physical and psychological proximity of the learner to the teacher. Indeed, Milgram demonstrated the importance of such factors by manipulating them—Milgram was able to obtain obedience rates as low as 10% and as high as 100% simply by altering a single variable of the experimental procedure (see Milgram, Obedience, for
descriptions of all 24 variations of the obedience experiments). Simply put, the power-of-the-situation argument goes a long way toward explaining the behavior of Milgram’s participants. Yet this framework lacks a unified theory of precisely why obedience occurred. Why were the features of Milgram’s experiments able to influence participants? What fundamental psychological process or processes underlie the effects of these elements of the experimental environment? I argue that we can answer these important questions if we understand the participants’ behavior as primarily motivated their desire to protect their self-images. Below, I will describe how the general human desire to maintain a positive view of ourselves mediated the influences of each of the most important environmental factors. I will show that, at the most fundamental level, efforts to protect one’s self-concept caused obedience in the Milgram experiments.

Defining the Situation: The Role of Role Models

“Society everywhere is in conspiracy against the manhood of every one of its members. The virtue in most request is conformity.” —Ralph Waldo Emerson

Among the most influential situational factors at play in Milgram’s experiments was the sheer newness of the experimental setting—participants had never been in a similar situation; the experimenter was novel, strange, confusing. Scholars refer to this newness and confusion as the “ambiguity of the experimental situation” (e.g. Lutsky, Nissani). This ambiguity caused participants to initially struggle to define the scenario playing out in front of them, a difficulty that led directly to their obedience (e.g. Nissani, Kelman & Hamilton, Blass, 1991). We engage in this process—of defining the situation in which we find ourselves—virtually all the time, often without even thinking about it. In fact, doing so is necessary for functioning in daily life. In each moment, we must define the scenario in front of us in order to know how to appropriately act (Smith & Kosslyn, 148-151)—no definition, no action. When participants entered the laboratory, they encountered an unfamiliar situation (Lutsky, 59), brimming with a multitude of alien objects and procedures. The situation evolved quickly from a tame experiment on learning and memory to an intensely stressful scenario, augmenting the ambiguity (Lutsky, 59; Milgram, 1965, 375). Because of this ambiguity, when the subjects tried to define the situation, their best course of action was to look to the experimenter as a role model—after all, he seemed like a credible authority and was an expert in this setting. When the participants looked to the experimenter as a guide for appropriate action, they did not find any reason to worry, but rather saw a calm, stern face, a face that looked neither anxious about the learner’s condition nor pleased with the victim’s pain. The experimenter was impassive, but not evil. As Neill Lutsky notes, “Thus, even though subjects have sufficient reason to believe the learner is being harmed by the procedure, the absence of both concern and malevolence on the part of the experimenter undermines a clear definition and ethical evaluation of the situation” (60). Thus subjects, confused about this strange situation, conformed to the experimenter’s definition of the situation and believed the learner was fine. The subjects continued delivering the shocks, not out of malice, but because the ambiguity of the situation led them to believe the experiment was morally sound.
Why did the participants model their behavior after the experimenter, rather than, say, decide for themselves how to act? In general, when individuals find themselves in an ambiguous situation, they tend to match their behavior to that of those around them. This conformity often occurs because of our innate desire to be “right” (Aronson, 37)—humans intrinsically wish to act in a manner we believe is appropriate, just, and correct. When the behavior of others is our only guide to suitable action—as is often the case (Aronson, 28)—individuals will conform, believing that following the example of others increases their likelihood of being right. We tend to conform the most when we perceive the available role model as having expertise or trustworthiness, which is exactly how participants in Milgram’s studies viewed the experimenter.

The desire to be right certainly contributed to obedience in the Milgram experiments. Because the experimental setting was ambiguous, the participants stood the best chance of defining the situation correctly—of “being right”—if they accepted the definition of the expert in the scene: the experimenter. Because the participants (like all humans) had a desire to be right, they obeyed. But wanting to be right was not the most fundamental desire at play. Rather, the participants’ desire to be right stemmed from an underlying force: the need to maintain a positive self-concept. Satisfying our desire to be right necessarily protects and improves our self-image. Intuitively, believing we are right is universally rewarding and inherently beneficial to our self-concepts—we like to see ourselves as acting appropriately and holding correct attitudes and beliefs (Aronson, 37, 138). When we do not have the necessary experience to know how to act appropriately in a given situation—to be “right” in that setting—we conform to the behavior of others to ensure that we do not appear foolish or act improperly; in short, we conform to ensure that our self-concepts won’t be damaged. It was the participants’ desire to maintain a positive self-image that caused them to imitate experimenter, and thus to obey.

Re-defining the Situation:
The Challenges of Cognitive Conservatism

The ambiguity of the experimental setting, which made it difficult for participants to define the situation, was one environmental factor that contributed to obedience. Closely related to the difficulties in forming an initial definition of the laboratory situation is the challenge of redefining the scene once a once has formed his initial judgment. Many have argued that this difficulty led Milgram’s participants to continue to obey the experimenter despite the obvious indications that they were harming the participant (e.g. Nissani; Kelman & Hamilton, 138; Saltzman, 135). Moti Nissani offers support for this claim with an empirical example of the difficulty of redefining the situation. In his study, participants were given an instructional manual and told to review it for several hours. In one section of the manual, the participants encountered a false formula for the volume of a sphere. Upon reaching this section, subjects were asked to use both this formula and a water-displacement method to estimate the volume of a spherical object they had been given. The empirical measurements contradicted wildly with the estimates from the formula, yet the participants stuck with the first thing they learned. Even when the subjects held PhDs in science, one hundred percent continued to use the “observationally absurd formula” throughout the four-hour experiment (Nissani, 1385).
Nissani argues that these results demonstrate that humans find it difficult to discard pre-existing beliefs; that is, the participants failed to reject the false formula because they initially believed in the reliability of the instructional manual and avoided abandoning this belief, even in the face of obvious contradictory evidence.

Like Nissani’s participants, Milgram’s subjects brought to the experiment may beliefs about scientific research, beliefs they later struggled to overturn. When subjects entered Milgram’s experiment, they knew that scientists work under various legal and ethical constraints. They believed the researcher, a scientist at a prestigious institution, was a trustworthy individual. They had no previous experience with inhumane experiments, no reason to believe a study on learning and memory would be sadistic. To disobey the experimenter, participants had to allow an enormous conceptual shift—they had to redefine an ostensibly ethical situation as frighteningly immoral, a presumably trustworthy experimenter as irresponsible and inhumane (Nissani, 1385). As Kelman and Hamilton write, the participants could challenge the experimenter’s orders “only on the claim that the demands were illegitimate—that, contrary to the way in which the situation had hitherto been defined, the authorities were not entitled to make these demands and the citizens, therefore, were not obligated to obey them” (138). But because humans tend to cling to their first interpretations of a situation, redefining the experimental setting proved too difficult a task for the majority of Milgram’s participants. Without this reinterpretation, the inevitable result was obedience.

Participants struggled to re-define the experimental situation because of a psychological tendency known as “cognitive conservatism.” According to this theory of behavior, humans try to preserve already established knowledge, beliefs, attitudes and stereotypes (Aronson, 157-160). Cognitive conservatism is a direct result of our desire to maintain a positive self-concept: preserving our initial assessments allows us to perceive of ourselves as consistent, stable individuals, and believing we possess these qualities clearly benefits our self-esteem (Aronson, 158 & 179). Thus, participants avoided redefining the situation because doing so meant they could no longer view themselves as consistent individuals. Here, again, egoism drove obedience.

**Cognitive Dissonance, Self-Justification and the Sequential Nature of the Task**

![Figure 4. Cognitive dissonance is also depicted in the popular press (Adams).](image)
In addition to our tendency to preserve our beliefs and assessments, another important psychological phenomenon called “cognitive dissonance” likely contributed to participants’ difficulties in redefining the situation and therefore to obedience. Cognitive dissonance is, in essence, an uncomfortable state of tension that occurs when an individual simultaneously holds two conflicting cognitions (see Festinger). These cognitions can be ideas, attitudes, beliefs, opinions or simply memories of an action or event. As the comic above nicely depicts, dissonance is uncomfortable simply because holding two contradictory ideas simultaneously is absurd and humans dislike being absurd (Aronson, 184). Thus, we try to reduce cognitive dissonance when we experience it.

When participants in Milgram’s experiments attempted to reduce dissonance, obedience was the likely result. While it seems that most participants did not have the cognitive resources necessary to make the conceptual shift needed to redefine the experiment as immoral (e.g. Nissani, see above), that a substantial number of participants (though still the minority) were disobedient demonstrates that interpreting the experimental situation was certainly possible. However, even if participants were able to change their understanding of the scene, their desire to reduce cognitive dissonance made them unlikely to act on their new definition. As soon as a participant understood the situation as immoral, he recognized that his complicity in the experiment was also immoral. This created uncomfortable dissonance: the belief “I am a decent, humane individual” was dissonant with the participant’s knowledge that “I am violating a universal moral code by inflicting pain on this harmless stranger." To reduce this dissonance, the participants had two options. First, they could either engage in self-justification, convincing themselves that shocking the victim as a justified and not immoral. One participant, for example, justified his behavior by viewing it in the larger, benevolent context of a psychological experiment, describing his thoughts during the experiment as, “I’m a nice person, I think, hurting somebody, and caught up in what seemed a mad situation... and in the interest of science, one goes through with it” (Milgram, Obedience, 54). The second course Milgram’s participants could take to reduce dissonance was to disobey the experimenter and cease their horrifying actions. While the latter option may seem like the most efficient choice for dissonance reduction, it also required admitting all previous behavior in the experiment was immoral. Because the participants would still have to deal with their complicity in the first portion of the experiment, dissonance could only be partially reduced. As Milgram notes, for the subject to disobey he must say to himself, “Everything I have done to this point is bad, and I now acknowledge it by breaking off” (149). Consequently, participants tended to opt for the former, more effective method of dissonance reduction—deciding to view their obedience as moral—and as a result continued the experiment through the very end.

Yet even before participants had to choose between acknowledging their complicity in an immoral experiment or attempting to reinterpret their behavior as justified, cognitive dissonance was at work preparing participants for continued obedience. Specifically, after each decision the participant made, he engaged in dissonance reduction that ultimately prevented him from viewing his actions as immoral. In general, we experience dissonance almost every time we make a decision (Aronson, 195). This dissonance occurs because decisions are seldom black and white—the unchosen alternative almost always has some positive qualities, causing us to experience dissonance between the cognitions “I did not choose this particular alternative” and “This particular alterna-
tive had X desirable quality.” We tend to reduce this post-decision dissonance through self-justification: we emphasize the positive aspects of the option we have chosen and derogate the alternatives.

The sequential nature of the teacher’s task meant the participants in Milgram’s experiments repeated this self-justification many times: every time the learner gave an incorrect response, the participants had to decide whether to shock him. When the participants chose to administer the shock, they reduced dissonance by devaluing the alternative of not shocking the victim. At first, these self-justifications were very easy to make: even though the participant knew the shocks were painful, the learner did not offer any negative responses to the milder shocks. Moreover, the participant believed the experiment was contributing to useful new scientific knowledge. Participants’ initial justifications set the stage for later obedience: if a participant justified an increase of 15 volts early in the experiment—say, from 45 to 60 volts—shouldn’t he also see a later increase of the same magnitude, such as from 300V to 315V, as reasonable? Once an individual starts down this slippery slope of self-justification, it becomes increasingly difficult for him to draw the line between a moral and immoral voltage level (Aronson, 201). John Darley, describing destructive obedience in general, explains this phenomenon, known as the foot-in-the-door-effect, this way:

The essence of the process involves causing individuals, under pressure, to take small steps along a continuum that ends in evildoing. Each step is so small as to be essentially continuous with the previous ones; after each step, the individual is positioned to take the next one. The individual’s morality follows rather than leads. (208)

By the time a participant gets to the second highest level, the final voltage level hardly seems different than the shock he has just administered. Justification for previous shocks provided rationalization in advance for complying with later, more demanding commands. There is evidence that the foot-in-the-door-effect was indeed at play in Milgram’s studies: if this sort of justification were taking place, it follows that the sooner in the course of the experiment a subject shows notable resistance, the more likely it is that he will eventually discontinue the study. Andre Modigliani and Francois Rochat reanalyzed the data collected from one variation of Milgram’s experiments and found that this hypothesis held.

In addition to devaluing the alternative course of action each time they administered a shock, participants could also reduce dissonance by disparaging the learner. Many psychologists have noted that individuals tend to derogate their victims, even when they only harm the victim under an authority’s orders. For example, in an experimental situation similar to Milgram’s, David Glass induced participants to deliver a series of electric shocks to others. Glass measured the participant’s attitudes toward the victim before and after the shocks were delivered. He found that individuals had more negative attitudes of the victim at the end of the experiment, suggesting that the aggressors derogated the victim to justify their cruelty towards him. Derogating the victim of the shocks maximized his culpability, thereby justifying immoral treatment of him—in essence, allowing the participants believe the learner got what he deserved (Aronson 226-7). Milgram notes that participants often made comments such as “He was so stupid and stubborn he deserved to be shocked” (Obedience 10). Milgram sums up the process well, explaining that once the subjects had harmed the victim, they “found it necessary to view him as an
unworthy individual, whose punishment was made inevitable by his own deficiencies of intellect and character” (10). Because the participants had to decide again and again whether to shock the subject, they would have greatly devalued the learner by the time he even started protesting. This process ultimately led the participants to view their behavior as perfectly moral, thus eliminating dissonance and ensuring obedience.

Attempts to reduce cognitive dissonance played a significant role in shaping participants’ behavior—and this dissonance reduction was, at its core, motivated by the participants’ desire to protect their self-image. When participants decided to shock the victim, they chose to devalue both the victim of their cruelty and the alternative course of action, a process that set the stage for continued obedience. If they had not engaged in this self-justification, they would have had to view their behavior either as immoral (because the victim did not deserve to be shocked), threatening their self-images. So participants justified their actions, and the result was obedience. Likewise, if a participant managed to realize his initial assessment of the experimental situation might be wrong, and internalized a new interpretation that deemed his behavior immoral, his view of himself would suffer. Alternatively, obedience allowed the participant to implicitly reject such situational definitions, and thus to walk out of the experiment with his self-esteem intact. In sum, when participants attempted to protect their self-images—both after they administered each shock, and when they recognized immoral definitions of their behavior—obedience resulted.

The Constraints of Social Norms

“Social norms rigidly constrain how we live, and individuals who transgress the limits can expect trouble. We may think that our personal life-style represents our own free choice, but that belief is often mistaken.” —Geoffrey Rose

As discussed above, many scholars have noted the ambiguity of Milgram’s experimental situation. Yet there were also characteristics of the laboratory setting that were simply unmistakable. Specifically, many researchers have noted how social norms, codes of behavior that are by definition unambiguous, influenced behavior in the obedience experiments. Specifically, social norms constrained the interaction between the experimenter and the participant, ultimately leading to obedience. Breaking off the experiment meant breaching his initial, implicit promise to aid the experimenter, and norms of interpersonal etiquette constrained the participant’s freedom to disobey (Milgram, Obedience 149). In the experiments, participants had to find a socially appropriate way to disobey—simply mustering the will to disobey was insufficient (Collins & Brief 89). Collins and Brief found that American teenagers preferred participants who politely disobeyed to those who defiantly disobeyed (99), suggesting that society expects us to disobey in an appropriate way, even if obedience means breaching universal moral codes. To disobey without facing social rejection, participants needed cognitive resources and social adeptness often beyond their reach, especially considering the extreme stress they experienced. Moreover, the experimenter employed adaptive interpersonal strategies to further hinder the participant’s attempts to disobey (Lutsky 61). For example, when faced with a worried participant, the experimenter in Milgram’s study consistently redirected the focus of the interaction from the condition of the learner to the experimental
procedures (“The experiment requires that you continue,” he might prod). To disobey, subjects needed the cognitive resources to both direct the focus of the interaction back to the learner and to reject the experimenter’s prods—all without treating a man of authority in a socially unacceptable way. When participants lacked the necessary verbal skills to do this, or could not access them due to their emotional stress, the result was obedience (Lutsky 61).

In essence, the participants’ decision to obey societal behavioral codes was intimately linked with their desire to protect their self-concepts. Humans have a universal desire to be liked—to gain social acceptance, to belong to a community. Satisfying this desire is clearly beneficial to our self-concepts: our self-esteem increases when we gain peer approval (Aronson 378-80). But in the Milgram experiments, disobeying all but guaranteed the experimenter would dislike the defiant participant. Indeed, to disobey, participants had to “disrupt the established social order and the smoothness of social interaction and to suffer the consequences of such a disruption” (Kelman & Hamilton 138). Disobedience meant forfeiting the possibility of gaining the experimenter’s approval, and was therefore an unattractive—and unpopular—course of action.

Conclusions

Scholars often use a power-of-the-situation explanation to understand the otherwise unsettling obedience of the participants in Stanley Milgram’s experiments. Yet at a more fundamental level, it was not simply situational factors that led to obedience, but also the subjects’ motivation to protect their self-images. Such efforts were the driving force behind participants’ obedience. Because they desired to be right, participants modeled their behavior after the experimenter, relying on the authority figure as a guide for appropriate action. Later, because they wished to view themselves as individuals with stable beliefs, they could not conceive of their own actions as immoral, and thus lacked the motive to disobey. When participants justified their behavior, they avoided confronting their own immorality and were unable to accurately understand their own behavior. Finally, because they desired social approval, which was sure to bring boosts in self-image, participants entered the laboratory with a propensity for obedience. The participants subscribed to social norms of interpersonal interaction, increasing the likelihood that the experimenter would bestow his approval—and that they would obey his malevolent orders.

A “power-of-the-situation” explanation for participants’ obedience takes responsibility off the shoulders of the participants—it is the situation itself that controls outcomes, not the individuals involved. While this interpretation has been well supported by many scholars, it is frustrating. If we are always to blame the situation, is no one accountable for his or her wrongdoings? Where does the power of the situation end and free will begin? And perhaps most importantly, if situations can be so powerful as to cause ordinary people to commit atrocious acts—acts incompatible with fundamental standards of morality—if normal individuals can become agents of terrible destruction simply because they happened to end up in a particular situation, what can anyone do to avoid the possibility of perpetrating horrifying acts?

A new look at the experiments can resolve these troubling questions. When we reinterpret the behavior of Milgram’s participants as not simply the result of the controlling
power of the situation, but also a result of the human tendency to protect our self-images, we restore to the participants responsibility for their actions, while still giving fair attention the influence of situational factors. Under this interpretation, the experimental situation encourages obedience, but it is ultimately up to the participants to decide whether to obey, or to shake off concerns about their self-images and end the experiment. We, too, can escape the pressures that lead to destructive obedience, but only if we are willing to confront our mistakes and abandon our need for a flawless self-image. This reinterpretation is more than another way of making sense of the participants’ behavior; it is a way by which we can reclaim our own ability to act morally, it is a lesson in our flaws and instruction for improvement. With it, human agency and responsibility alike are recouped. This reinterpretation has an undeniable power: we hold close our beliefs about free will, and are deeply unsettled by the thought of losing our agency; under my argument, no situation can force us to commit destructive acts—we always have a choice.

Milgram saw broad significance in his experimental findings. From the outset, his work was driven by a desire to answer some of the most profound questions about human nature that were begged by the Second World War. Why do people commit such violent acts against innocent others? Do each of us have a kernel of evil lurking within us—are we really so different from Eichmann and Hitler? Would we, too, kill another human being at the order of an authority? Milgram’s experimental findings extend far beyond the walls of his laboratory—they force us to question our most deep-seated assumptions about human nature. My contentions with Milgram’s claims are borne out at this level, as well. In Milgram’s studies, we have seen how the fundamental desire to view oneself positively can lead to glaringly negative results. However, we cannot wholly condemn the participants for attempting to protect their self-images, as there are circumstances in which high self-esteem is desirable. For example, when we view ourselves as sensible and good human beings, we are sometimes less likely to engage in immoral behavior. Researchers have found that students with higher self-esteem are less likely to, for example, cheat to win money (Aronson & Mettee). Self-esteem is also powerful motivator—when an individual believes in himself, he is more likely to work harder and more persistently toward a goal, and such diligence is a necessary pre-requisite for the most important human advancements, and for helping society function at the highest possible level. Such hard work leads to life-saving scientific discoveries, to more competent doctors, teachers, policemen, civil servants, and political leaders—in short, all of society can benefit from an individual’s high self-esteem. But even so, we cannot forget what the Milgram experiments have taught us: attempting to protect one’s self-esteem often leads one to hold a dangerously distorted perception of the world and oneself. Milgram’s participants disparage the innocent victim of their shocks, and came to believe that inflicting severe pain on a harmless man was appropriate. So on one hand, having high self-esteem can lead to many desirable outcomes; but on the other, the things we sometimes must do to maintain this positive self-image can be horrifically destructive and simply inhumane. This fundamental aspect of human nature is a double-edged sword, one that we must each decide how to wield.
Endnotes
2 This attribution falls under the category of a more general tendency to assume that another’s behavior can be explained by their personalities. This phenomenon is known as the “fundamental attribution error.”
3 Arendt did not totally agree with this interpretation. Rather, she insisted that we always have moral choices, even under totalitarian rule. She was careful to point out that even in ‘powerful’ situations like that of Nazi Germany, there are some people who did not comply.
4 See, e.g., Darley, 204; Milgram, Obedience, 6; Zimbardo, 566; Blass, 1991, 399
5 High rates of conformity have been found even when participants knew the appropriate course of action in a given situation without the guidance of others. Asch (1951) offers a powerful demonstration of this tendency. In a famous study, the overwhelming majority of Asch’s participants matched their peers’ incorrect responses on a simple perceptual task. The task was truly unambiguous: when the participants answered in private, they never failed to offer the correct responses. Even so, the participants repeatedly conformed to their peers. If many participants will conform even in unambiguous situations, it is no small wonder high levels of conformity occur in situations that are more difficult to define. Sherif (1937) offers the classic study of this phenomenon. In this study, a point of light in a dark room seems to move while actually remaining stationary. Participants were to estimate the distance the light moved. When the participants responded individually, answers were erratic. When participants responded in groups, on the other hand, participants conformed to one another and answers converged around a single measure. This convergence was not due to a desire to be liked: individual participants continued to report the converged measure even when the other group members were no longer present.

Works Cited


