

In Defense of the Milgram Experiments

Adam Chernew

Abstract

In the early 1960s, social psychologist Stanley Milgram conducted a series of studies at Yale University in which he measured the willingness of subjects to obey an authority figure (the experimenter) who instructed them to administer electrical shocks to a confederate under the guise that the experiment was testing the effects of punishment on learning. Although the electrical shocks were fake, these famous obedience experiments are, to this day, recognized as some of the most controversial psychology experiments of all time. While Milgram's experiments yielded seemingly profound insight about human obedience to authority, many in his field were quick to criticize his work for violating research ethics. Over the past fifty years, not much has changed. The consensus amongst the philosophical community is still that Milgram's obedience experiments were largely unethical, and that his procedure would never be approved by an IRB today. This paper, however, challenges this popular notion. To do so, it reexamines the criticism of some of Milgram's sharpest detractors, namely Diana Baumrind, Steven Patten, and Steve Clarke. In addressing these critiques, I incorporate both arguments that Milgram made in his own defense, as well my own arguments. Ultimately, I show that none of the arguments accusing Milgram of harming his subjects purport definitive evidence that the subjects were actually considerably harmed.

What constitutes an unethical experiment?

In discussing whether the Milgram experiments were unethical, the term "unethical" must first be defined. This is because the exercise of determining the validity of the criticism levied against Milgram would be a useless if there were not a mutual understanding of what constituted an "unethical" experiment in the first place. Although the criticisms of Baumrind, Patten and Clarke are distinct in a number of different ways, they all, directly or indirectly, seem to define unethical research in the same way: that which substantially harms the subject of the experiment. Fortunately, I agree with this definition. Thus, while there are a variety of different standards that could be used to judge the ethics of a psychology experiment, for the purposes of this paper, an "unethical" experiment will be defined simply as one that significantly harms its subjects.

Evidence of perceived harm

In reporting the results of his obedience experiments, Milgram mentions in great detail

the adverse reactions that some his subjects experienced throughout the course of the trial. For example, in the abstract of his paper, Milgram notes that “The procedure created extreme levels of nervous tension in some [subjects]. Profuse sweating, trembling, and stuttering were typical expressions of this emotional disturbance. One unexpected sign of tension – yet to be explained – was the regular occurrence of nervous laughter, which in some [subjects] developed into uncontrollable laughter.” Delving further, Milgram highlights the experience of one subject in particular. He writes:

I observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, who was rapidly approaching a point of nervous collapse. He constantly pulled on his earlobe, and twisted his hands. At one point he pushed his fist into his forehead and muttered: “Oh God, let’s stop it.” And yet he continued to respond to every word of the experimenter, and obeyed to the end.

On the whole, Milgram admits that:

Many subjects showed signs of nervousness in the experimental situation, and especially upon the administration of more powerful shocks. In a large number of cases the degree of tension reached extremes that are rarely seen in psychological laboratory studies. Subjects were observed to sweat, tremble, stutter, bite their lips, groan, and dig their fingernails into their flesh. These were characteristic rather than exceptional responses to the experiment.

Based on these “striking reactions of tension and emotional strain,” Milgram’s critics allege that his obedience experiments considerably harmed his subjects in two distinct ways. First, some argue that Milgram’s obedience experiments simply exposed the subjects to such undue levels of stress and anxiety, that doing so actually constituted harming them. As Patten writes, “In discussions of the morality of the Milgram experiments much of the dispute has centered on the amount of pain or stress experienced by the subjects, so that some of the critics are inclined to say that Milgram exposed his subjects to the risk of real harm.” Furthermore, Baumrind and Clarke suggest that Milgram’s obedience experiments also harmed its subjects by revealing to them negative self-knowledge, thus damaging their self-esteem and dignity. Clearly, there are significant accusations suggesting that Milgram substantially harmed his subjects. This paper will review each accusation in turn.

Illegitimate defenses of milgram

Before reviewing each of these charges, however, it would be remiss of me not to examine some of the more illegitimate arguments made to justify the Milgram

experiments on ethical grounds. While this paper illustrates that there is much to be said in the defense of Milgram, this does not mean that all arguments to this end are reasonable. In fact, because illegitimate arguments in defense of Milgram are prone to “crowd out” legitimate arguments supporting the same conclusion, it is imperative to separate the reasonable from the flawed. Only then can the best case for supporting Milgram’s experiments on ethical grounds be made.

Some Milgram apologists have justified his obedience experiments by suggesting that even if they significantly harmed subjects, doing so was justifiable because the experiments contributed considerably to our knowledge of human obedience. Regardless of whether this claim is actually true (more recent research suggests that it may not be), this utilitarian sentiment is dangerous. The reason is because the “value” of an experiment is inherently subjective and often unclear. As Baumrind explains:

The behavioral psychologist is not in as good a position to objectify his faith in the significance of his work as medical colleagues at points of breakthrough. His experimental situations are not sufficiently accurate models of real-life experience; his sampling techniques are seldom of a scope with which he would like to endow his results.

In other words, because the “benefit” of a psychological experiment cannot be quantified with any certainty, substantial harm to a subject ought not to be justified with this line of reasoning. Yet even if the value of the experiment was unambiguous such that, despite significant harm to the subject, society clearly benefited, would we not still be concerned for the rights of the subject? Do subjects not have the right to remain uninjured while participating in a psychological experiment? Analogously, if we could save five lives by picking one random individual off of the street and harvesting their organs, would we deem this acceptable? We obviously would not. Hence, because we value individual rights, considerable harm to a subject in any experiment, no matter how groundbreaking, cannot be justified on the grounds that society has been made better off. For this reason, utilitarian defenses of the Milgram experiments fall markedly short.

A second illegitimate defense of the Milgram obedience experiments is that even if subjects were substantially harmed by participating, Milgram could not have predicted this harm, and thus his experiments were not unethical. As Milgram himself writes in his response to Baumrind:

The extreme tension induced in some subjects was unexpected...there was every reason to expect, prior to actual experimentation, that subjects would refuse to follow the experimenter’s instructions beyond the point where the victim protested; many colleagues and psychiatrists were questioned on this point, and they virtually all felt this would be the case.

To be fair, there is little reason to question the legitimacy of the expectations that Milgram had going into his experiment. It is reasonable to believe that the majority of Milgram's colleagues severely underestimated the degree of obedience that Milgram could expect to obtain. It is also not the case, as Patten argues, that "it is an essential part of the idea of these experiments being properly put together that tension and stress should result." Had the majority of Milgram's subjects stopped administering shocks after hearing protests from the learner, as was expected, significantly less tension and stress would have resulted.

The reason that Milgram's expectations do not exonerate him from the accusation that his experiments were unethical is because, as Milgram himself notes, "After a reasonable number of subjects had been exposed to the procedures, it became evident that some would go to the end of the shock board, and some would experience stress." As such, the "unintended consequences" argument is only applicable to the first few trials that Milgram ran. After that, Milgram admits that he fully expected the patients to experience severe stress.

Therefore, to reiterate what was stated at the beginning of this paper, I agree with Milgram's detractors that any psychological experiment that injures its subjects in a considerable way ought to be considered unethical. It should be sufficiently clear that no harm is justified by the outcome of research, and it must never be explained away as an "unintentional consequence." Those who defend the Milgram obedience experiments on these grounds are doing Milgram a great disservice, both because these arguments are flimsy in their own right, but more importantly, because there are better arguments to be made in defense of his work. Milgram's obedience experiments were ethical, but not because the "harm" caused to the subjects was outweighed by other considerations. They were ethical because none of the subjects were actually harmed in a substantial way in the first place. This claim may seem dubious given how Milgram described the reactions of some of his subjects, but it is not. In examining each type of accusation of harm individually, it becomes clear that throughout the course of the experiments, the subjects were never exposed to danger, and never ran the risk of injurious effects resulting from their participation.

Harm by stress and anxiety

Milgram's critics allege that his obedience experiments significantly harmed their subjects in several different ways. The most common of these accusations is that Milgram, in forcing his subjects to choose between disobeying an authority figure or administering potentially lethal doses of electrical shock, exposed them to severely harmful levels of stress and anxiety. On its face, this charge does seem to hold weight. As Milgram himself admits, a number of subjects in the experiment demonstrated striking reactions

of tension and emotional strain, such as sweating, trembling and stuttering, not to mention nervous laughter and groaning.⁹ Nevertheless, there are numerous reasons to believe that the subjects were not seriously injured in any way. First, as Milgram notes in his reply to Baumrind, the responses to a follow-up questionnaire sent to all subjects indicated that on the whole, participants felt positively toward the experiment. 84% of the subjects stated that they were glad to have been in the experiment; 15% indicated neutral feelings, and only a mere 1.3% indicated negative feelings. Likewise, 80% of the participants felt that more experiments similar in nature to Milgram's should be carried out, and 74% of the subjects indicated that they had learned something of personal importance as a result of being in the study. Issues of response bias aside, it seems highly unlikely then, that an experiment that harmed the majority of its participants would receive such overwhelmingly positive reviews. Furthermore, as Milgram notes in his reply to Baumrind, an impartial medical examiner also found no reason to believe that the subjects were injured during the course of the experiment. Specifically, after interviewing 40 of the subjects he felt were the most likely to "have suffered consequences from participation," the psychiatrist concluded that "none was found...to show signs of having been harmed by his experience...Each subject seemed to handle his task [in the experiment] in a manner consistent with well-established patterns of behavior. No evidence was found of any traumatic reactions." Again, it seems highly unlikely that if Milgram's experiments had legitimately injured its participants, an impartial medical examiner would have found no signs thereof, particularly among those participants who he deemed most likely to have been harmed.

In addition to "subject-based" evidence, there are other reasons to suspect that Milgram's subjects were not actually harmed in a considerable way during the course of their participation. One such reason is the careful post-experimental treatment administered to all subjects. As Milgram explains in his reply to Baumrind:

At the very least all subjects were told that the victim had not received dangerous electric shocks. Each subject had a friendly reconciliation with the unharmed victim, and an extended discussion with the experimenter. The experiment was explained to the defiant subjects in a way that supported their decision to obey the experimenter. Obedient subjects were assured of the fact that their behavior was entirely normal and that feelings of conflict or tension were shared by other participants.

Of course, a thorough debriefing does nothing to prevent harm during the course of the actual experiment. Yet, to the degree that an effective post-experimental treatment allows the subject to walk away from their experience feeling significantly better than they otherwise would have, it helps to mitigate harm from the overall experience. To be perfectly candid, no one piece of evidence mentioned previously proves that Milgram's subjects were not substantially harmed. Taken together though, they make a pretty strong case. Given that (1) the vast majority of subjects reflected positively on their experience, (2) an impartial psychiatrist found no evidence of harm amongst the most

vulnerable participants, and (3) an elaborate debriefing procedure was put in place to assuage the concerns of the subjects, it seems highly improbable that the participants who experienced stress and anxiety were truly “harmed” by these emotions. To give a simple analogy, during the course of a particularly difficult exam, it would be expected that some students taking the exam would become stressed and/or anxious. This stress and anxiety, in turn, could manifest itself in striking reactions of tension and emotional strain, such as sweating, trembling and groaning (or potentially even nervous laughter). Yet we would not suggest that students are being legitimately harmed by taking the exam, and thus we would not label the exam itself as “unethical.” It stands to reason then that the same standard must be applied to the Milgram experiments, and therefore one cannot legitimately argue that despite all of the evidence to the contrary, Milgram’s subjects were significantly harmed by stress and anxiety.

Patten, to his credit, attempts to poke several holes in this defense of Milgram. For example, while he concedes that none of the subjects may have suffered “for any lengthy period of time”, he points out that “the person who is assaulted in a city street and who recovers from his injuries in a few days does not, by hypothesis, suffer any long term injury. Yet it would be most peculiar to say that he was not harmed.” Although true, this is an irrelevant analogy. Would 84% of people assaulted in a city street say they were glad to have been attacked? Would 80% feel that more assaults should be carried out? Would 74% suggest that they had learned something of personal importance from the attack? I think not. The fact that Milgram’s subjects reflected positively on their experience indicates that they likely did not suffer in the long run or during the course of the experiment. Patten also suggests that just because the vast majority of Milgram’s participants reacted positively to their experience, that fact does not justify the experiment on the whole. He contends that:

The fact that most of Milgram’s subject’s decided they did not mind the way they were manipulated by him...will not thereby show that the experiment was excusable. What is needed first is some preliminary case-by-case reasoning by Milgram: to show, perhaps, that the sort of harm at issue is quite like the kind that can be excused by the victim... and unlike those which are not so defeated.

Here, Patten’s argument practically defeats itself. It is true that positive responses to the experiment alone do not prove that harm was not committed. However, the criteria that Patten lays out in the very next sentence seems to contradict this idea. By overwhelmingly endorsing the experiment, are Milgram’s subjects not “excusing” any “harm” that may have been committed against them? By this standard, Milgram’s experiments pass with flying colors. Finally, Patten tries to delegitimize the results of Milgram’s follow-up survey by questioning the credibility of the respondents. Patten writes:

If the results of the experiments bear the interpretation that Milgram recommends, a large portion of the subjects are...persons said to have

behaved in a shockingly immoral manner. So why should we now trust their judgments about the propriety or impropriety of their treatment in the experiments?

Of all of Patten's arguments, this is the most absurd. To answer his question, the reason that we should trust the responses of the subjects is because behaving in an immoral way, particularly in an environment conducive to doing so, does not reflect one's ability to judge how they have been treated.

The two are completely unrelated; Patten is trying to draw a correlation where there is none. A participant who administered maximum shock ought to still be able to sincerely reflect on the experiment. There is no reason to believe otherwise. Thus, Patten's arguments fail to show that Milgram exposed his subjects to undue levels of stress and anxiety, so much so that it constituted significantly harming them.

Damaged self-esteem and dignity

Unsurprisingly, criticisms regarding the ethics of Milgram's obedience experiments have not been limited to a concern over undue stress levels in participating subjects. Other detractors have gone a step further, arguing that the Milgram experiments harmed their participants in a deeper, more psychological way as well. Specifically, these critics allege that Milgram severely damaged the self-esteem and dignity of his subjects. This allegation is less straightforward than that concerning undue stress, but it must be examined nonetheless with equal scrutiny.

As mentioned previously, Baumrind regards the Milgram obedience experiments as considerably harmful not because of "physical discomfort, inconvenience or experimental deception per se" (like Patten), but because she believes "the emotional disturbance described by Milgram could easily effect an alteration in the subject's self-image. She explains that:

The subject's personal responsibility for his actions is not erased because the experimenter reveals to him the means by which he used to stimulate these actions. The subject realizes that he would have hurt the victim if the current were on. The realization that he also made a fool of himself by accepting the experimental set results in additional loss of self-esteem.

Baumrind adds that unless "a fairly corrective interpersonal experience" is employed, she would "expect a naïve, sensitive subject to remain deeply hurt and anxious for some time, and a sophisticated, cynical subject to become even more alienated and distrustful."

Similarly, Clarke criticizes Milgram on the grounds that social psychologists have an obligation to protect their subjects from negative self-knowledge. Clarke contends that:

Many participants in the Milgram obedience studies found out something unexpected about themselves; that they were more prone to obey authority figures than they might have supposed. While there may sometimes be long-term benefits to individuals to be derived from gaining this information about themselves, such self-discoveries can often be harmful rather than beneficial. Subjects who make unexpected and unwelcome discoveries about themselves can be subjected to lowered self-esteem, and other negative feelings.

Like Patten's accusations, Baumrind and Clarke's criticism admittedly seems legitimate at first. It's certainly plausible that subjects, after reflecting upon the fact that they had been coerced by an authority figure into administering "lethal" dosages of electrical shock, would have serious doubts about their own morality and would likely question just what atrocities they are capable of committing. Thus, on the surface, it does seem reasonable to believe that some subjects could have walked away from the experiment with a damaged sense of self-worth and a loss of dignity. This conclusion, however, is not consistent with the results of the follow-up survey discussed previously, as well as the assessment of the independent medical examiner. If the self-esteem of a significant number of subjects had been damaged in the course of the experiment, the vast majority of participants responding to the post-experiment questionnaire would not have indicated that they were glad to have been in the experiment, learned something of personal importance from it, and thought that similar work should be carried out in the future. If this were the case, the psychiatrist examining subjects a year after their participation would have found signs of traumatic reactions. Therefore, despite what Baumrind and Clarke theorize, it is clear that Milgram's subjects did not suffer losses of dignity as a result of their participation in the obedience studies.

The reasons that most subjects did not leave the experiment feeling deeply disturbed by their behavior are twofold. First, as mentioned earlier, a "careful post-experimental treatment was administered to all subjects." Subjects that administered shocks all of the way until the end were assured after the study that their seemingly deplorable actions were entirely normal, and that their feelings of anxiety were shared by many other study participants. Without a doubt, this debriefing technique made otherwise remorseful and guilty subjects feel, at the very least, a little bit better knowing that their actions were not uniquely immoral. Furthermore, subjects also managed to avoid suffering from a guilty conscious due to a phenomenon Milgram describes in his response to Baumrind. As Milgram writes:

The same mechanisms that allow the subject to perform the act, to obey rather than to defy the experimenter, transcend the moment of

performance and continue to justify his behavior for him. The same viewpoint the subject takes while performing the actions is the viewpoint from which he later sees his behavior, that is, the perspective of “carrying out the task assigned by the person in authority.”

In this way, the subject is able to justify his actions in retrospect, even when he realizes that they would have resulted in the death of a random person. By remembering himself as an agent of an immoral act, rather than its main perpetrator, the subject is able to avoid bearing the moral brunt of his actions and is in a sense shielded from negative self-knowledge. It is for this reason then, along with a thorough debriefing session, that subjects are able to avoid unbearable amounts of guilt, a loss of dignity, and therefore the type of “psychological harm” that Baumrind and Clarke had predicted.

Concluding remarks

Although it remains an unpopular notion in social psychology circles, even today, I sincerely believe that the obedience experiments conducted by Stanley Milgram at Yale University in the early 1960’s were not unethical. Despite the adverse reactions that some of Milgram’s subjects experienced during the course of their participation in his experiments, a closer examination of all relevant information reveals that the study participants were not substantially harmed by undue stress nor by alterations of their self-image. When the responses of the subjects, the findings of a medical expert, and Milgram’s experimental design are taken into consideration, this conclusion becomes apparent, and the criticisms of Milgram’s detractors are shown to be based much more on misguided theory than on fact. When the Milgram obedience experiments are scrutinized from all angles, it becomes clear that, perhaps surprisingly, there is nothing particularly unethical about them. This paper matters then because it is one of the few that truly defends the Milgram experiments on ethical grounds. Hopefully, it can help push the dialogue surrounding the ethics of Milgram’s work in the other direction.