Milgram’s Obedience Study: A Contentious Classic Reinterpreted

Richard A. Griggs

Abstract
Given the many older criticisms of Milgram’s obedience study and the more damning recent criticisms based on analyses of materials available in the Milgram archives at Yale, this study has become a contentious classic. Yet, current social psychology textbooks present it as an uncontentious classic, with no coverage of the recent criticisms and little coverage of the older ones. Also, none of the texts present any coverage of the recent reinterpretation of the study’s findings in terms of engaged followership based on participants’ acceptance of the experimenter’s scientific goals. Hence, the present article provides sources for summaries of the criticisms and a summary of the reinterpretation and its supporting empirical research for teachers who want to incorporate coverage into their courses.

Keywords
Milgram’s obedience study, engaged followership, social identity theory

Perhaps, the most famous study in psychology is Milgram’s controversial obedience study (Milgram, 1963, 1964, 1965a, 1965b, 1974). Although this set of experiments is more than 50 years old, the debate about the ethical, methodological, and theoretical issues of these experiments shows no signs of abating (Gibson, 2013b). Much of the recent interest in the Milgram experiments is concerned with new criticisms of both Milgram’s experiments and his reporting of them, such as (a) his misrepresentation in publications of the debriefing process in which roughly 600 participants were not told that the learner was not actually shocked until about a year later (Nicholson, 2011; Perry, 2013), (b) the unreported improvisational, off script extended prodding of participant teachers by the experimenter that varied across experiments and participants within an experiment (Gibson, 2013a, 2013b; Perry, 2013; Russell, 2009), (c) Milgram’s own unpublished analysis that shows that the majority of the participants disobeyed when they thought that the learner was actually being shocked (Perry, 2013), (d) Milgram’s selective reporting of the 23 experiments that he actually conducted (Perry, 2013; Russell, 2014),1 and (e) Milgram’s clear use of selective editing to foreground obedience and downplay resistance in his film Obedience (Millard, 2011, 2014; Perry, 2015). This recent Milgram controversy has led Tavris (2014) to label Milgram’s study a “contentious classic” and to recommend that it be taught as such with discussion of its flaws and shortcomings. Teaching Milgram’s study as a contentious classic, however, would be difficult for teachers because currently, introductory social psychology textbooks present it as an uncontentious classic. Indeed, in a review of the latest editions of 10 introductory social psychology textbooks, Griggs and Whitehead (2015b) found that none of the recent criticisms were covered, even in the texts published in 2015.2 Not even the fact that Milgram had misrepresented the debriefing process in his publications was mentioned in any of the textbooks.

Griggs and Whitehead (2015a), using the same sample of textbooks, similarly found very sparse coverage of older criticisms—external validity and generalization concerns (e.g., Darley, 1995; Orne & Holland, 1968) and methodological problems, such as the presence of demand characteristics (e.g., Mixon, 1972; Orne & Evans, 1965). Griggs and Whitehead also noted that when these criticisms are discussed, it is in a Milgram-friendly manner (e.g., one text used the Milgram study as an example of where poor external validity is not always bad). Ethical issues with Milgram’s study, though, are discussed in almost all social textbooks. However, when they are introduced, it is typically in a general manner without citing

1 Department of Psychology, University of Florida, Gainesville, FL, USA

Corresponding Author:
Richard A. Griggs, 4515 Breakwater Row West, Jacksonville, FL 32225, USA.
Email: rgriggs@ufl.edu
references for specific ethical criticisms (only 3 of the 10 texts cited a specific reference, and it was Baumrind, 1964, with one also citing Baumrind, 1985), and these discussions usually come to a Milgram-friendly conclusion, such as that the knowledge-gained benefit outweighed the possible participant cost. This was expected because past research on social psychology textbooks has found that these texts have moved from coverage of specific ethical criticisms to more general, Milgram-friendly coverage of ethical issues toward the end of the last century (Stam, Lubek, & Radke, 1998). Hence, at present, given the tendentious textbook coverage, teachers who are interested in teaching the Milgram study as a contentious classic would have to research the many old and new criticisms of the study on their own. Griggs and Whitehead (2015a, 2015b), however, provided summaries of the old and new criticisms, respectively, and thus comprise an excellent starting point for such research. Nevertheless, I am concerned that teachers will not incorporate these criticisms in their courses because they are not covered in the textbooks that they are using. Supporting this point, Bartels, Milovich, and Mousssier (2016) found that introductory psychology teachers tended not to cover the many criticisms of the Stanford prison experiment in their courses, paralleling the lack of such coverage in introductory psychology textbooks (Griggs, 2014).

Tavris (2014) also argued that as teachers we need to find new ways of persuading students that, despite their shortcomings, the findings of these older contentious classic studies like Milgram’s study apply to them, because today’s students are just as eager as students decades ago “to reject unflattering or counterintuitive portrayals of humanity.” She did not, however, mention that some social psychology researchers have recently found a new way to explain Milgram’s findings, and it does not involve obedience to authority. Thus, if teachers are going to find ways to persuade students of the importance of Milgram’s findings, they would need to be aware of this reinterpretation of those findings. Briefly, this reinterpretation argues that the experiments were not about obedience to authority but rather engaged followership based on identification with the experimenter and his scientific project (Haslam, Reicher, & Birney, 2014, 2016; Haslam, Reicher, & Millard, 2015; Haslam, Reicher, Millard, & McDonald, 2015; Reicher & Haslam, 2011; Reicher, Haslam, & Miller, 2014; Reicher, Haslam, & Smith, 2012). The details of this new interpretation and the empirical evidence gathered to test it will be presented later.

Tavris is not the only one who does not mention this reinterpretation of Milgram’s findings. Griggs and Whitehead (2015b) found that coverage of this reinterpretation of Milgram’s findings was nonexistent in current introductory social psychology textbooks, even though relevant articles appeared as early as 2011. Given this lack of coverage in the current texts and the fact these texts have a revision cycle of at least 3 years, the purpose of the present article is to provide a brief summary of this recent reinterpretation of Milgram’s findings for teachers who may want to incorporate it into their class presentations and discussions of Milgram’s study. The story of this reinterpretation begins with Burger’s (2009) partial replication of Milgram’s baseline experiment, which I discuss next.

From Replication to Reinterpretation

Burger (2009) conducted a partial replication of Milgram’s new baseline experiment at Santa Clara University. His participants were men and women who responded to newspaper advertisements and flyers that had been distributed locally. Their ages ranged from 20 to 81 years, with a mean age of 42.9 years. Obviously, some changes to ensure the welfare of the participants were necessary to obtain permission to run the study (Burger, 2007). The main procedural change was that once participants pressed the 150-volt switch and started to read the next test item, the experiment was stopped. The 150-volt point was chosen because in Milgram’s original research, once participants went past 150 volts, the vast majority continued to obey up to the highest shock level. In a meta-analysis of data from eight of Milgram’s obedience experiments, Packer (2008) also found that the 150-volt point was the voltage level at which participants were most likely to disobey the experimenter, probably due to the fact that it was at the 150-volt point that the learner began to complain verbally. Thus, in Burger’s study, it was a reasonable assumption that the percentage of participants that go past 150 volts was a good estimate of the percentage that would go to the end of the shock generator. Of course, the experimenter also ended the experiment when a participant refused to continue after hearing all four of the experimenter’s prods.

Almost 67% of the men and about 73% of the women that Burger (2009) studied continued to 150 volts. Although these percentages have to be adjusted down slightly because not every participant in Milgram’s study who went past 150 volts maximally obeyed, these results are very close to Milgram’s finding of 65% obedience for both men and women in the baseline condition. Even with such adjustments, Burger’s findings indicate that people reacted in this laboratory obedience situation today much as they did almost 50 years ago in Milgram’s original study. Burger, however, quizzically later concluded that his participants were not displaying obedience (see Burger, Girgis, & Manning, 2011). Why did he conclude this?

Burger, Girgis, and Manning (2011) pointed out that only the fourth prod, “You have no other choice, you must go on,” truly constitutes an order, and in an analysis of the participants’ comments and responses in Burger’s earlier partial replication study, they found that this prod did not elicit any obedience because not a single participant continued after receiving it. In general, the more a prod resembled an order, the less likely it was followed (see Study 2), leading Burger et al. to surmise that the participants’ response pattern was the opposite to what would be expected if they had been following orders. They concluded that “If the participants were not following orders, then alternate interpretations of Milgram’s findings should be explored . . . . Moreover, the way the research is portrayed to students, scholars, and the public may need to be reassessed” (p. 6).
Burger et al.’s (2011) results beg the question as to whether this pattern of participant responding also occurred in Milgram’s experiments. Fortunately, Gibson (2013a) performed a rhetorical analysis of the archival-recorded interactions between the experimenter and participants in two of Milgram’s experiments, providing us with an answer to this question. It is important to remember that analyses of the audiotapes of the experimental sessions have revealed that the experimenter did not follow the standard prod protocol that Milgram described in his publications about the experiments (Perry, 2013). He went off script and improvised prods. Thus, these improvised prods have to be analyzed with respect to their order-like nature. Consistent with Burger et al.’s conclusion, Gibson’s analysis revealed that the experimenter’s most order-like prods were overwhelmingly resisted by the participants. Gibson (p. 303) concluded that his analysis “points to the intriguing possibility that the studies ultimately may have little to do with obedience as conventionally understood.” Thus, rather than showing that Milgram’s participants were obeying the orders of those in authority, Milgram’s experiments seem to provide evidence of the opposite—namely, that orders from an authority lead to disobedience and that the “obedience” that Milgram observed was due to other factors.

**From Obedience to Identification**

In accordance with Burger et al.’s (2011) conclusion that alternative interpretations of Milgram’s findings should be considered, Alex Haslam and Stephen Reicher have proposed that Milgram’s “obedient” participants were motivated not by orders but by appeals to science and that their behavior needs to be reconceptualized as an act of “engaged followership” with the experimenter and the scientific community and not as a product of blind obedience to authority (e.g., Haslam et al., 2014; Haslam, Reicher, Millard, & McDonald, 2015). In these terms, the level of obedience in each of Milgram’s various experiments is predicated upon the extent of the participants’ acceptance of the experimenter’s scientific goals. Haslam and Reicher argued that this leadership induced participants to identify with the experimenter and become engaged in helping him achieve his scientific goals. They further proposed that the participants in Milgram’s experiments may, under certain conditions, also have opted to identify with the learner and not the experimenter, leading them to “disobedient” behavior. Hence, the perturbing process of deciding which identification to make led to the anxiety and upset witnessed in Milgram’s participants.

Which identification participants tended to make was determined mainly by which identification a particular experimental setting favored. In line with this analysis, Haslam and Reicher noted that one can explain the variance observed for the obedience rate in Milgram’s numerous experiments (from 0% to 100%) by examining how the situational factors in each experiment favor each type of identification (i.e., the relative extent of identification with the experimenter vs. identification with the learner). In fact, Reicher, Haslam, and Smith (2012) have shown that estimations of the levels of identification with the experimenter and with the learner made by both expert social psychologists and nonexpert college students for Milgram’s descriptions of 15 of his experiments are strong significant predictors of the level of obedience found in each of the experiments. In agreement with the engaged-followership explanation, identification with the experimenter was a strong positive predictor of the level of obedience observed, and identification with the learner was a strong negative predictor of the level of obedience observed.

Haslam, Reicher, and Birney (2014) pointed out that there was a confound in Milgram’s study and Burger’s (2009) partial replication between the content of the four prods and the order in which they were presented such that it is unclear whether the observed resistance to the fourth prod was the consequence of it being an order or that it was issued after the other three prods had already been resisted. Possibly, the participants were just tired of being prodded or were already committed to resisting when the fourth prod was given. They argued that the second prod, “The experiment requires that you continue,” is the one that relates most to their engaged-followership proposed explanation because it indicates that continuing is essential to the success of the experiment and hence, science. In a very cleverly designed analogue of Milgram’s basic procedure with 30 steps, each involving progressively more toxic responses, Haslam et al. demonstrated that continuation and completion of an objectionable task was positively predicted by the extent to which prods appealed to scientific goals but not by the extent that the prods were seen as orders. In agreement with Burger et al.’s (2011) finding, the participants were more inclined to disobey an order than to follow it.

Haslam et al. (2015) used immersive digital realism (IDR) to restage and reexamine Milgram’s baseline treatment (see Note 1) and four of his obedience experiments.5 The IDR methodology circumvents the ethical barriers to conducting obedience research using Milgram’s original procedure with volunteer participants by using professional actors who deeply immerse themselves into portraying fictional characters. Because the actors can differentiate their characters’ behavior from their own, ethical issues are avoided. Kathryn Millard, the third author of the study, had decided to use the IDR methodology to recreate Milgram’s experiments as part of a documentary film that she was making on Milgram’s obedience study. The first two authors, Alex Haslam and Stephen Reicher, viewed Millard’s recreation of the Milgram experiments as an opportunity to test their engaged-followership reinterpretation. Hence, they observed the shooting of the experiments and collected and analyzed the data from the filmed experiments to determine whether these data fit their reinterpretation.

Functioning as filmmaker and director, Millard worked with the actors to develop their characters before the filming. In this case, their characters were participants in Milgram’s experiments being restaged and filmed in a faithful reproduction of the original laboratory environment, including the ominous shock generator. Millard, however, only informed the actors...
that their characters would be participants in a social psychology experiment in the film. They were not given any information about the nature or design of the Milgram experiments. Following the digital filming of each restaged experiment, the actors were thoroughly debriefed and provided with complete information about the study and its aims. Postexperimental interviews were also conducted and used to assess the participants’ relative identification with the experimenter and the learner in each experiment. For ethical reasons, Millard conducted follow-up interviews with the actors about their experiences a year later.

Validating the use of the IDR, a strong positive correlation was found between the maximum level of shock administered in these restaged experiments and the mean maximum shock administered in the corresponding original Milgram experiments. Consistent with the engaged-followership explanation, relative identification with the experimenter versus the learner as assessed in the interviews was also a good predictor of the maximum shock that participants administered in each experiment. In addition, as Burger et al. (2011) had found, there was near universal refusal by participants to continue after being given Milgram’s fourth prod (“You have no other choice, you must continue”).

Interestingly, Haslam and Reicher’s engaged-followership proposal not only provides an explanation of both Milgram’s various experimental findings and Burger’s replication results but also provides an explanation of the discrepancy between the extremely stressful and aversive experimental experience of the participants and the positive feelings toward the experiments that they expressed in their questionnaire responses about their participation. In an analysis of the Yale archival questionnaire data from Milgram’s study, Haslam, Reicher, Millard, and McDonald (2015) showed that the participants were engaged with the science of the experiments and that they saw science—especially science at prestigious Yale University—as a “social good,” and being associated with this made them feel good. It is critical, as Haslam et al. pointed out, to realize that the participants’ questionnaire responses were made quite some time after their participation, so the stressful experimental situation that they experienced almost a year earlier was in the past, and the debriefing report that accompanied the questionnaire reminded them only of the scientific goals of the study. In sum, led by the report’s reminder of the study’s lofty scientific purpose, the participants felt that they had contributed to scientific progress, and this gave meaning to their participation, transforming the unpleasant, stressful experimental experience into something to feel good about when they completed the postexperimental questionnaire.

Epilogue

Why would social psychology textbooks present such tendentious coverage of Milgram’s obedience studies, providing little, if any, coverage of the myriad criticisms of the study and the recent reinterpretation of the study’s findings? To explain their findings of an analysis of social psychology textbooks from 1965 to 1995, Stam, Lubek, and Radke (1998) provide a tenable answer to this question. According to Stam et al., social psychology textbooks “serve a knowledge-conserving function for the discipline . . . there is a great deal of temporal consistency, a shared core of material and authors to be discussed, and the adoption of a homogeneous knowledge-conserving perspective” (p. 156). They propose that as part of this knowledge-conserving function, a “standard view” of the obedience experiments in social psychology textbooks has developed (see also Griggs & Whitehead, 2015b). According to Stam et al. (pp. 162–163, italics in original),

The obedience research is no longer a case study of the importance of obedience to authority but an important promoter of the importance and necessity of experimental social psychological research.

The visibility of the research has become a token: by its critics, a token of the vulnerability of the discipline; by proponents, a token of its strengths. Within the discipline, Milgram is valorized for his contributions but the recurring appearance of discussions of methodology and ethics indicate that in order to valorize Milgram’s studies social psychologists must continually engage in damage control. It is this combined valorization/defensiveness that we take to be the standard view of the obedience experiments.

If indeed this is the case, then coverage of Milgram’s obedience study as a contentious classic that has been reinterpreted as a demonstration of engaged followership and not obedience is unlikely to materialize in future editions of introductory social psychology textbooks. The next set of editions of introductory social psychology textbooks will tell the story. Hopefully, it’s not the standard one.6

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Because one of Milgram’s unreported experiments (the last one in the study, the relationship experiment) may be the most ethically questionable experiment that he conducted, I briefly summarize it here for teachers who may want to incorporate it into their class discussions of the ethical issues posed by Milgram’s study. This experiment was a second one that Milgram conducted in Bridgeport, CT, and was conducted in the same office building but was procedurally different. Instead of participants as teachers and a confederate as the learner, 20 pairs of males who were related in some way (relatives, friends, or neighbors) served as participants, with one serving as teacher and the other as learner. After the learner was strapped in and the teacher and experimenter left the room, Milgram entered the room with the learner and explained privately to him about the experimental ruse and coached him on how to vocalize in response to the supposed shocks. Rochat and
Blass (2014, p. 457) pointed out another difference between these two Bridgeport experiments. In the reported experiment with unrelated participants, the learner’s protests were aimed at the experimenter. In the unreported one with related participants, the learner’s protests were aimed at the teacher. Was obedience, in the end, the钱币 you wanted? Did Blass teachers obey and inflict pain on a relative, friend, or neighbor? No, the vast majority did not. Milgram found an 85% rate of disobedience, one of the highest rates in all of his experiments. Although Milgram may not have reported this experiment because of this high rate of disobedience, it seems more likely that it was because the experiment itself would be difficult to defend ethically given that the teacher was asked to shock a friend or relative, especially given the ethical firestorm that had already erupted after Milgram’s first publication on the obedience study. Milgram (1963). More details about this experiment can be found in Perry (2013; see Chapter 6) and Rochat and Modigliani (1997).

2. According to Griggs and Whitehead (2015b), this set of 10 introductory social psychology textbooks essentially comprise the population of American introductory social psychology textbooks if briefer versions of two of these texts and Aronson’s briefer, more trade-like The Social Animal (2012) are excluded.

3. Milgram actually conducted two baseline experiments. In the original baseline experiment (the voice-feedback experiment), the teacher and learner are in separate rooms, but the teacher can hear the learner’s escalated screams and protests and his refusal to continue at higher voltage levels and the learner does not respond after 330 volts. Milgram used the results of this experiment as the baseline for two more experiments. Because he had to change the laboratory in which he was conducting the obedience study after these experiments, Milgram felt it necessary to conduct a replication experiment of the baseline experiment to determine whether the new laboratory, which was much more modest and in the basement of the building, had any impact on his findings (Milgram, 1974, p. 55). The second baseline experiment was the same as the first but with one small change to the script—the learner mentions a slight heart condition before the experiment and again during his protests. Milgram (1974) thought that this heart problem might provide additional justification for disobeying (p. 55). It did not. He replicated his earlier 62.5% finding—65% administered the maximum shock. He then used this experiment (the new baseline experiment) for the remaining 18 experiments that he conducted. Burger’s (2009) experiment was a partial replication of Milgram’s new baseline experiment (J. M. Burger, personal communication, April 28, 2016).

4. Haslam et al. (2014) examined Milgram’s experimental notes that are available in the Milgram archives at Yale and found that he had actually questioned the terminology that he was using to describe the behavior of the participants in his obedience experiments. Milgram was not sure whether their behavior constituted cooperation with authority or obedience to authority and pondered the difference between the two. His answer was: “Cooperation implies a certain willingness to perform the action or help out, a certain internal desire to assist, while obedience implies an action that is totally in response to a command, with no motivational support from inner sources” (Milgram Archives, Yale University Library, Box 46). Obviously, Milgram decided to describe the participants’ behavior as obedience, but his thoughts about it as cooperation are congruent with Haslam and Reicher’s reinterpretation in terms of engaged followership to help the experimenter achieve his scientific goals.

5. This study was a collaborative effort across film and psychology. The immersive digital realism performance methodology used to restage Milgram’s experiments was devised by the third author, Kathryn Millard, a noted filmmaker, essayist, and academic. Millard (2012) had initially used this methodology to restage Gamson’s famous sociological research into encounters with an unjust authority (Gamson, Fireman, & Ryuina, 1982). The filmed reenactments of Milgram’s experiments are featured in her new documentary film on Milgram’s obedience study, Shock Room. The first and second authors of the study, Alex Haslam and Stephen Reicher, appear in the film and describe their engaged-followership reinterpretation of Milgram’s obedience findings. The film makes clear that when confronted with the fourth prod (the only outright order of the four prods), participants disobey, demonstrating our ability to choose freely. For more information about Shock Room, go to http://shockroomfilm.com.

6. New editions of 2 of the 10 introductory psychology textbooks were recently published—Aronson, Wilson, Akert, and Sommers (2016) and Gilovich, Keltner, Chen, and Nisbett (2016). Unsurprisingly, neither text mentioned Haslam and Reicher’s engaged-followership reinterpretation of Milgram’s findings. One of the two texts, however, did include one Milgram-friendly sentence on the recent criticisms. This sentence points out that recent criticisms have focused on disturbing allegations that Milgram misrepresented his debriefing process in his publications and that many subjects left the experiment without being told that the learner was not actually shocked. These criticisms, however, are not allegations but rather assertions by more than one researcher supported by factual evidence in the Milgram archives at Yale. This fact and the details of the criticisms are not provided. Hence, the standard view persists in these new editions of 2 of the 10 introductory social psychology texts examined by Griggs and Whitehead (2015a, b).

References


