

Stanley Milgram's Obedience Studies: An Ethical and Methodological Assessment

Nestar Russell

DANS **PHILOSOPHIA SCIENTIÆ** 2024/2 28-2 , PAGES 107 À 129

ÉDITIONS **UNIVERSITÉ DE LORRAINE**

ISSN 1281-2463

Date de mise en ligne : 04/06/2024

Article disponible en ligne à l'adresse

<https://shs.cairn.info/journal-philosophia-scientiae-2024-2-page-107?lang=en>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour Université de Lorraine.

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur cairn.info/copyright.

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

Stanley Milgram's Obedience Studies: An Ethical and Methodological Assessment

Nestar Russell

University of Calgary (Canada)

Résumé : Avec l'ouverture des archives personnelles de Milgram, à partir du milieu des années 1990, une « seconde vague » de littérature sur les Études sur l'Obéissance s'est développée. Une partie de cette littérature suggère de manière convaincante que les expérimentations de Milgram sont si problématiques sur le plan éthique et méthodologique qu'elles ne mériteraient pas l'énorme attention qu'elles ont reçue et continuent de recevoir. À l'autre extrémité du spectre, certains chercheurs soutiennent qu'il y a encore beaucoup à apprendre de ces expériences. Confronté à des opinions aussi divergentes, voire même contradictoires, qui doit-on croire ? Après avoir examiné cette littérature, cet article aborde deux questions : les expérimentations sur l'obéissance sont-elles éthiquement contestables ? Et restent-elles méthodologiquement valides ? Cet article conclut que bien que les expérimentations sur l'obéissance soient, pour de nombreuses raisons, contraires à l'éthique, elles restent néanmoins méthodologiquement valides. L'article conclut également que c'est précisément en raison du caractère éthiquement scabreux de ces expérimentations qu'elles restent si pertinentes pour comprendre le monde réel, bien au-delà des murs des laboratoires.

Abstract: With the opening of Milgram's personal archive, from the mid-1990s there has been a "second wave" of literature on the Obedience Studies. Some of this literature convincingly suggests Milgram's research is so ethically and methodologically problematic that it does not deserve the enormous attention it receives. At the other end of the spectrum, some scholars argue there is still much to be learnt from these experiments. With such divergent (contradictory?) views, what is one to believe? After reviewing this literature, this article addresses two questions: are the Obedience Studies unethical and do they remain methodologically valid? This article concludes that although the Obedience Studies are, for many reasons, highly unethical, they remain methodologically valid. It also concludes that it is because of Milgram's

unethical practices that his research remains so relevant to the world beyond his laboratory walls.

One could undertake an experiment in which a man were [sic] in fact being shocked. This could be done within ethical grounds. A very large amount of money, say \$100 would be offered [to] volunteers. He would be told the exact nature of the experiment, that he might possibly receive shocks of up to 450 volts, and that they were extremely painful, and that he could do anything he wished to try to have the experiment stopped. My guess is that it would make little or no difference in the overall experimental results.

Stanley Milgram (SMP [Stanley Milgram Papers], Box 46, Folder 176, undated)

1 Introduction

In the early 1960s social psychologist Stanley Milgram ran the Obedience Studies, a set of experiments claiming to demonstrate most ordinary people would willingly follow an authority figure's instructions to ostensibly harm an innocent person. In academic circles and beyond, this research made Milgram both revered and, through accusations of subject abuse, reviled.

Less debatable is Milgram's scholarly influence: the Obedience Studies have inspired a literature so vast, one scholar described its arrival coming in two main waves [Kaposi 2017], much of it polemical and some of it contradictory. In terms of this research's ethical status and methodological strength, what can be concluded: who and what is to be believed? After presenting Milgram's basic findings, this article addresses two questions: are the Obedience Studies unethical and are they methodologically valid?

2 Obedience studies: An overview

Milgram's most well-known experiment is probably the *New Baseline* condition, as shown in his documentary *Obedience* [Milgram 1965a]. In it a subject volunteered to partake in a study purporting to determine if punishment improved learning. The "experimenter" (an actor) ensured, through a rigged draw, the subject—paid \$4.50—became a "teacher" who asked the "learner"

(another actor) a string of memory-type questions. The “experimenter”, in the company of the subject, then entered a small room and restrained the “learner” by strapping them into an electric chair. The concerned but compliant “learner” mentioned he had a mild heart condition and enquired into the intensity of the shocks. The “experimenter” curtly informed him the shocks were painful but not dangerous, then left with the subject and entered an adjoining room. The subject was seated in front of the 30-switch shock generator, which had switches ranging from 15 to 450-volts, increasing in 15-volt increments. The “experimenter” informed the subject that for each of the “learner’s” incorrect answers, they were to administer an electric shock. And for each incorrect answer received, punishment was to increase by 15-volts. As the shocks increased, the verbal designations located below the shock switches intensified: “SLIGHT SHOCK”, “VERY STRONG SHOCK”, “DANGER: SEVERE SHOCK” and “XXX” [Milgram 1974, 28]. When subjects resisted their instructions, the experimenter urged them along with four main prods:

- PROD 1: Please continue, or, Please go on.
- PROD 2: The experiment requires that you continue.
- PROD 3: It is absolutely essential that you continue.
- PROD 4: You have no other choice, you must go on.

If needed, the experimenter deployed two special prods: “Although the shocks may be painful, there is no permanent tissue damage, so please go on” and “Whether the learner likes it or not, you must go on until he has learned all the word pairs correctly. So please go on” [Milgram 1974, 21–22].

As the “learner” was “shocked”, the subject heard his intensifying reactions (standardised tape recordings). For example, at the 150 and 195-volt switches the “learner” mentioned, with minor variations, “My heart’s bothering me” [Milgram 1974, 56]. At 300 volts the “learner” refused to cooperate and thereafter responded with agonised screams. After the 345-volt switch, the “learner” went silent and the “experimenter” then instructed the subject to consider all unanswered questions as incorrect and to inflict further intensifying shocks. On a subject inflicting three consecutive 450-volt shocks, the experiment ended, and they were considered fully “obedient”. If a subject refused to inflict every shock asked of them, they were considered “disobedient”. The experiment was not exploring the effects of punishment on learning, but whether subjects agreed to follow the experimenter’s seemingly harmful instructions. The *New Baseline* condition generated a 65% completion rate. Milgram believed his subjects’ decision to complete the experiment was an illustration of moral failure and, for him, most performed in “a shockingly immoral way” [Milgram 1964, 849].

Milgram described his post-experimental treatment of subjects in a section of his first journal article titled “Interview and dehoax”:

[...] procedures were undertaken to assure that the subject would leave the laboratory in a state of well being. A friendly reconciliation was arranged between the subject and the victim, and an effort was made to reduce any tensions that arose [...]. [Milgram 1963, 374]

To determine why most subjects completed his experiment, Milgram undertook a score of slight baseline variations which, through a process of trial-and-error, he hoped would inferentially lead him to a theory of obedience to authority.

Milgram's research attracted what proved to be some resilient ethical and methodological critiques. Baumrind wrote the most influential ethical critique [Baumrind 1964]. She was appalled by Milgram's treatment of subjects: the deception was so intense (highly realistic?) that, as his article put it, some were converted into "twitching, stuttering wreck[s] [...] rapidly approaching a point of nervous collapse" [Milgram 1963, 377]. Orne & Holland wrote the most enduring methodological critique: Milgram's subjects were unlikely to have been fooled by the deception (insufficiently realistic?) [Orne & Holland 1968, 282]. What follows will summarise this (contradictory?) research, along with the more influential critical literature it inspired.

3 Ethical critics

When Milgram published his first findings, his procedure's explicit reliance on deception attracted a variety of first wave ethics-based critiques. They argued Milgram:

- failed to provide subjects with informed consent about the harm-related risk/s they potentially faced.
- denied subjects' right to withdraw.
- recklessly exposed subjects to a variety of psychological harms.
- failed to provide subjects with a post-experimental debriefing capable of restoring their wellbeing [see Russell 2018, 112].

On Milgram's archive opening around the mid-1990s, a second wave of researchers revealed he engaged in the following unethical practices.

3.1 Most subjects were not dehoaxed

Building on Blass's initial observation [2004, 72], Perry confirmed that although Milgram provided subjects in conditions 19 to 24 with an honest (full) debriefing, the first 70% of subjects in conditions 1 to 18 received a dishonest (incomplete) debriefing [Perry 2013, 92]. So, after the experiment, what were

most subjects told? Most subjects were apparently informed that the shock machine was calibrated for use on “mice and small rats” and that “the verbal designations on this machine [...] are for these small animals” [Russell 2018, 113]. Thus, the learner was not in as much pain as their reactions—and the shock machine’s verbal designations—may have suggested.¹ Consequently, the first 600 subjects left the laboratory thinking they inflicted real shocks [Perry 2013, 84–86]. Milgram’s justification for replacing “one untruth for another” [2013, 85], was his fear the subject pool in New Haven would become contaminated: word would spread that the shocks were fake.

Milgram, however, failed to anticipate that these subjects, having just participated in a deception-based experiment, might remain skeptical of his dishonest debriefing. Indeed, one such subject later checked a newspaper obituary column to confirm if he had been “a contributing factor in the death of the so-called ‘learner’ [...]” [Russell 2018, 113].

Perry is critical of Milgram’s subheading “Interview and dehoax” [Milgram 1963, 374], because the latter word “implies truth-telling”, which thereby promoted a false impression that after the experiment “all had been revealed” [Perry 2013, 85–86]. On publishing his book, Milgram must have realised this error because the word “dehoax” thereafter disappeared from his writing and a technically more accurate, yet still ambiguous (misleading?) phrasing emerges:

A careful postexperimental treatment was administered to all subjects. The exact content of the session varied from condition to condition and with increasing experience on our part. At the very least every subject was told that the victim had not received dangerous electric shocks. [Milgram 1974, 24]

Subtly, the phrase “postexperimental treatment” avoids explicitly stating, but still allows for the possibility, that most subjects were not dehoaxed. Gibson is correct: for years scholars took Milgram’s statements on the post-experimental debrief to mean “participants were informed that no electric shocks were administered” [Gibson 2019, 64]. Furthermore, there is “little doubt” that during his lifetime Milgram never actively corrected—thus encouraged—the false impression that he fully dehoaxed all subjects. [2019, 63]. Also key in Milgram’s above quote is his inclusion of the word “dangerous”: all subjects were at least told that they never administered harmful shocks [2019, 64]. The word “dangerous” (again) allows for the possibility, but avoids explicitly stating, that most subjects were not told the experiments were fake.

1. At least occasionally, the experimenter—without Milgram’s knowledge—failed to administer this debriefing [Perry 2013, 94–96].

3.2 Milgram buried condition 24

Milgram’s final baseline variation was termed the *Relationship* condition. Here subjects brought an acquaintance to the laboratory, with one becoming “teacher” and the other “learner”. Learners were covertly informed that the experiment was exploring if their acquaintance would follow an authority figure’s harmful orders. Learners were quickly trained in how to react to the “shocks”. Only 15% of subjects completed this experiment. Milgram recognised this variation’s significance: the results were “as powerful a demonstration of disobedience than [sic] can be found” [SMP, Box 70, Folder 289, as cited in Perry 2013, 202]. The most ethically disconcerting feature of this experiment was that three of the twenty teacher-learner pairs brought a family member as their acquaintance [Russell 2018, 115]. Although all three pairs refused to complete the experiment, had the teachers inflicted every shock, the damage to these relationships could have been irreparable. Further, Milgram knew this, telling one “obedient” subject during the debriefing his friend:

[...] wasn’t really getting the shock, we just set this up this way to see [...] whether you would be happy to give him the shocks [...]. So [...] let’s tell him that [...] you knew you weren’t giving him the shocks [...] alright? [Russell 2018, 115]

In 1962 Milgram wrote an article [Milgram 1965] that took a few years to appear in print [Russell 2018, 101], which presented various baseline variations. Milgram also mentioned “FURTHER EXPERIMENTS”, one of which concerned a “personal relationship [...]” [Milgram 1965, 71]. Limited by space, he explained these experiments “will have to be described elsewhere [...]” [Milgram 1965, 71]. Two years later Baumrind published her critique of Milgram’s research [Baumrind 1964]. If Baumrind was appalled at Milgram’s first and relatively benign *Remote* condition, one can only imagine her reaction had she known of the *Relationship* condition where, in one case, a father was coerced into shocking his own son [Russell 2018, 205].

In his book, Milgram elaborated on his “FURTHER EXPERIMENTS” [Milgram 1974], with one exception: the *Relationship* condition was missing. Why did Milgram exclude this “powerful [...] demonstration of disobedience”? Milgram likely buried the variation in fear of ethical critics like Baumrind.

3.3 No medical pre-screening: Risks of physical harm

Early during data collection, subject Herbert Winer—a Yale professor—confronted Milgram: this experiment was dangerous because he thought it imposed “unwarranted strain on people who had had no previous medical screening of any kind [...]” [Blass 2004, 117]. Other subjects raised the same concern [Russell 2018, 117]. Despite the repeated warnings, Milgram never

introduced any medical screening, perhaps because he thought such concerns were exaggerated. They were not. In fact, one subject informed Milgram: “since taking part in the experiment I have suffered a mild heart attack—the one thing my doctor tells me that I must avoid is any form of tension [...]” [2018, 117]. With no medical safety-net in place, the experiments could have ended in tragedy. Then again, it transpires that Milgram—perhaps fearing such an outcome—had put a safety-net in place, albeit legal and not medical.

3.4 General release form

After undertaking his experiments, Milgram was publicly adamant: despite the intense stress some subjects experienced, his research was harmless [Milgram 1964, 849]. That said, just before all subjects started the experiment, they were encouraged to sign “a general release form [...]” which stated, “In participating in this experimental research of my own free will, I release Yale University and its employees from any legal claims arising from my participation” [Milgram 1974, 64].

Milgram's book does not state the purpose of the release, implying it was probably another prop. However, Milgram's “EXPERIMENTER'S INSTRUCTIONS” manual (held in his archive) highlights the calculated efforts Milgram went to best ensure every subject inattentively signed this document, thus hinting at something more underhanded. The Experimenter's Instructions states that after securing the learner, the experimenter was to say to the subject:

If you'll sit down there [at the shock machine], we'll begin [...]. Oh, the first thing I should do is to pay you. It slipped my mind. (GET CHECK AND RECEIPT) I guess that, rather than unstrap the learner from his chair, I'll wait and pay him later. Now, I'll have to get your name and signature on this receipt for our records. Please read and sign the standard clearance we must have from all participants in our research, although this project itself is not dangerous. [Russell 2018, 174]

It appears Milgram wanted the experimenter to try and distract subjects over the issue of payment in exchange for their signature on the “standard clearance”. If the experiment was “not dangerous,” subjects, of course, had nothing to fear in signing the document. But there was, it transpires, much more to this standard clearance than Milgram let on. As Milgram more candidly revealed in one archival document dated 16 October 1961: “The release [...] was not used for experimental purposes, but to protect us against legal claims” [Russell 2018, 173]. If Milgram genuinely believed his research was harmless, why did he need legal protection?

Before embarking on data collection, Milgram could not be certain his experiments would prove harmless. Consequently, there remained a risk that

some subjects, in some way, might be harmed and, as a result, complaints might be lodged. Therefore, Milgram, his staff, and his employer Yale University, all required some means of protection to discourage any litigious action. If Milgram knew before embarking on the official experiments that some subjects might be harmed (and the above document implies he did), his completion of the Obedience Studies was highly unethical.

3.5 Post-Baumrind (1964) prevarication and lies

When Baumrind published her ethical critique, it is believed to have initiated “the most intense debate on research ethics in the history of psychology [...]” [McGaha & Korn 1995, 147]. Milgram expected some ethical critiques of his research [Perry 2013, 270], but likely thought such voices would be tempered by how impressively he had captured immoral behaviour in his laboratory. Baumrind, however, saw no merit in the Obedience Studies; even implying the only immoral behaviour observable to her was Milgram’s pursuit of professional advantage.

Of course, wide support for Baumrind could have encouraged the American Psychological Association to label the Obedience research unethical. If so, Milgram’s future career in the academy would be in peril. Probably motivated by self-preservation, Milgram thereafter descended into a hazy cloud of evasive lies and prevarication.

For example, he responded by arguing:

At the outset, Baumrind confuses the unanticipated outcome of an experiment with its basic procedure. [...] The extreme tension induced in some subjects was unexpected. [Milgram 1964, 848]

Several years earlier, in December 1960, when Milgram’s undergraduate class ran the first pilot, it is possible that he found the intense stress experienced by some subjects unexpected. But when he ran the first official experiments the following year, he knew some subjects would—as stated in his research proposal dated 25 January 1961—experience “extreme tension [...]” [Russell 2018, 157]. Because Milgram expected some subjects to experience intense levels of stress *before* running his official experiments, here he lied to Baumrind. Nicholson is correct: under pressure from Baumrind, Milgram “decided to lie his way through the criticism” [Nicholson 2011, 744–745].

After the publication of Baumrind’s critique, Milgram’s tone changed: there was no more mention of “twitching, stuttering wreck[s]”, with him thereafter describing subject stress as “momentary excitement [...]” [Milgram 1964, 849], [Milgram 1974, 194]. This was “a most astonishing about-face” [Patten 1977, 356]. Milgram also accused Baumrind’s article of being “deficient in information” that, in the spirit of transparency, “could have been obtained easily” from him [Milgram 1964, 848]. Milgram’s apparent

openness is, however, undermined by his burying the incomparably unethical *Relationship* condition. If Baumrind's view of Milgram's research was deficient in information, it was because over time he both passively and actively withheld it from her—and everyone else—thereby inhibiting a comprehensive and accurate ethical assessment. Then Milgram released his widely viewed documentary *Obedience* [1965a], which only presented subjects receiving full and honest debriefs, thus bolstering and powerfully perpetuating this false impression [Perry 2015, 630–631]. There is merit to Perry's conclusion that Milgram's documentary “was a publicity coup” that succeeds less as science and more as “a triumph of propaganda” [Perry 2015, 635]. Milgram's spiraling mistruths may support Donald Warwick's warning that deception in research will only “[...] reinforce a cavalier attitude toward truth [...]” [Warwick 1975, 40].

For many of the above reasons, it can be argued that the Obedience Studies were, as the second-wave archival researchers demonstrate, more unethical than previously imagined. However, this label does not render the research internally invalid. So, are the Obedience Studies methodologically valid?

4 The methodological critics

The most resilient methodological criticism directed at Milgram's research is the claim that “obedient” subjects unlikely believed the learner was receiving dangerous shocks. Also, after Milgram's archive opened, another significant criticism gained ground: the experiments lacked standardisation. The following section presents both critiques.

4.1 Construct validity

Milgram's experimental procedure forced subjects to make a choice: side with the learner and stop inflicting shocks or side with the experimenter's goal of collecting data and inflict more shocks. Within subjects, this choice attempted to generate a “conflict” of conscience [Milgram 1963, 378]. Milgram assumed his attempts to deceive subjects into believing they were harming the learner was “critical” [Milgram 1972, 139] to the internal validity of the experiments because if he failed, there would have been no conflict of conscience.

Orne & Holland's most resilient criticism against the Obedience Studies was their claim that Milgram's attempts at deception likely failed because subjects would have known, if only vaguely, that the experimenter (or Yale University) would never have exposed the learner to such danger. Therefore, subjects could safely presume, despite evidence to the contrary, no harm would come from inflicting every shock and that doing so would be “all right” [Orne & Holland 1968, 287]. Many “obedient” subjects' post-experimental comments support this view: “the way I figured it, you're not going to cause yourselves

trouble by actually giving serious physical damage to a body” [Russell & Gregory 2021, 69].

Despite Milgram’s rebuttal (that 80.1% of subjects later said they believed the shocks were “fully” or “probably” real) [1972, 141], subjects’ belief that the experiments were probably harmless lingered, see [Brannigan 2020, 49–51], [Eckman 1977, 94], [Gibson 2019, 38–39], [Harre 1979, 105], [Nicholson 2011, 749], [Perry 2013, 173, 258]. As de Swaan argued:

Nobody in their right mind would ever accept the idea that someone, anyone, would be electrocuted in the presence of certified researchers in the psychology lab on the campus of Yale University [...]. [De Swaan 2015, 28]

Data from recent studies even supports Orne & Holland’s original claim, see [Hollander & Turowetz 2017], [Perry, Brannigan *et al.* 2020, 90].

4.2 Standardisation

Darley observed that Milgram’s book transcripts indicate the experimenter regularly strayed from his instructions to only use four main and two special prods [Darley 1995, 130]. He concluded the experimenter said “*whatever was necessary to get the teacher to continue giving the shocks*” [1995, 131, emphasis original]. Implied here is a methodological criticism: Milgram’s study lacked standardisation. Perry and Gibson’s independent archival analysis confirmed Darley’s observation. In summarising this research, Gibson notes, in conflict with Milgram’s official account, the experimenter’s

[p]rods are used multiple times (Perry 2012), out of sequence, or sometimes not at all (Gibson 2013b), and in nonstandard form (Gibson 2013a, 2017). [Gibson 2019, 67]

In fact, during one trial the experimenter attempted to reassure an apprehensive subject by *leaving the laboratory* to ostensibly consult with the learner over his willingness to continue [2019, 103–104]. Gibson also discovered the experimenter deployed five additional prods [2019, 107], none of which are mentioned in Milgram’s official publications. Perry observed that as the research program progressed, the experimenter strayed further from his script. More specifically, during the early variations, he “scrupulously terminated the experiment after he had delivered the fourth prod” [Perry 2013, 133]. However, nearing the end of data collection—more specifically during Condition 20 *Women as Subject*—the experimenter, for example, mercilessly insisted one subject continue 26 times [Perry 2013, 134]. This condition obtained a 65% completion rate, but to conclude women completed equally to men in Condition 5 *New Baseline* is misleading [Brannigan 2020, 177–178].

All these lines of evidence support the view that, particularly nearing the end of data collection, a lack of standardisation pervaded Milgram’s research,

with the experimenter saying, as Darley originally observed, whatever would encourage the subject to inflict more shocks.

5 Should Milgram's research be discarded?

The collective sum of the above ethical and methodological literature has led some scholars toward a critical terminus. Perry, for example, “cast[s] doubt on Milgram's reliability as a narrator of the obedience research [...]” [Perry 2013, 90], concluding “the closer I looked at the inner workings of the experiment, the more contrived and unconvincing the results seemed” [Perry 2013, 383]. As her research journey ended, Perry discarded her copy of Milgram's book, perhaps implying others act similarly [2013, 312–313]. She is not alone: others found Milgram's research so ethically and methodologically problematic, they titled their special edited journal issue “Unplugging the Milgram Machine” [Brannigan, Nicholson *et al.* 2015]. What follows assesses this damning ethical and methodological terminus.

6 Assessing the critical terminus

To assess the validity of the critical terminus, one must briefly review Milgram's research journey. When inventing the Obedience Studies, Milgram had two main objectives. His *first objective* was to construct a well-designed—highly realistic—baseline procedure. This was because, as mentioned, he thought if subjects did not believe they were inflicting real shocks, they would not have encountered a moral dilemma—no tension—over completing the experiment. On the journey to achieving this first objective, in November 1960 Milgram's students ran the first obedience pilot, which obtained a 60% completion rate [Milgram 1973, 64]. Although this result “astonished” Milgram [Blass 2004, 68], because of the students' disjointed procedure, fake-looking shock generator, and weak acting skills [Russell 2018, 64], Milgram suspected some subjects completed because they did not believe the learner was being harmed. Cautious about his students' “not very well controlled” pilot [Blass 2004, 68], Milgram thereafter refined the basic procedure, built a new shock machine capable of fooling electrical engineers, and hired carefully selected actors—especially the experimenter—who many official subjects found to be thoroughly convincing, see [Blass 2004, 75], [Russell 2018, 68, 76]. He added other touches that bolstered the shock generator's verisimilitude, like ensuring that all subjects received a sample 45-volt shock before starting the experiment.

Milgram's *second objective* was to construct a first official experiment that would “maximize obedience [...]” [Russell 2011, 158]. Milgram's reasoning

here was that an experiment involving harmful orders that produced a low completion rate was predictable, so to capture academia's attention, he believed his first official experiment needed to "create the strongest obedience situation [...]" [Russell 2018, 56]. Milgram's main strategy to achieving this objective was to repeatedly add to the emerging procedure a variety of what he termed Binding Factors (BFs): powerful bonds that push people into doing things they would prefer not to do [Milgram 1974, 148]. Examples of BFs included:

- *The \$4.50 payment*: financial remuneration likely made subjects feel contractually obligated to do as they were told.
- *The Experimenter's prods*: within the context of a hierarchical chain of command, a superordinate authority figure's attempts to impose their will on the subordinate subject likely pressured the latter into inflicting more shocks.
- *The shock generator's escalating shocks*: as determined by the Foot-in-the-Door phenomenon [Freedman & Fraser 1966], persons are more likely to agree to a significant request (inflicting intense shocks) if it is preceded by a comparatively insignificant request (inflicting light shocks).

The more BFs Milgram added to the basic procedure, the cumulatively more coercive it became.

But as Milgram's emerging procedure became both more realistic *and* more coercive, the more stress it generated in subjects. And the more stressed subjects became, the greater their potential resistance to completing. If subjects became more resistant to completing, this outcome threatened to detract from Milgram's second objective: maximising the completion rate. Milgram's solution to this problem was to railroad subjects with even more BFs and/or slightly reduce subject stress over the prospect of hurting another person by inserting, into the emerging procedure, a variety of what he termed "Strain Resolving Mechanisms" (SRMs). SRMs are techniques that can reduce the tensions people normally experience when inflicting harm on another person [Milgram 1974, 153–164]. SRMs do not "resolve" all stress, they typically reduce it enough—often by injecting greater ambiguity over the reality of harm-infliction—to the point that an anxious subject became susceptible to being pushed along by one or more of the many BFs into continuing (thus, SRMs aided Milgram in keeping outright defiance at bay).

There were many SRMs involved in the experiment, including, for example, the shock generator: relative to physically striking the learner, this impersonal harm-inflicting device: "creates a sharp discontinuity between the ease required to depress one of its [...] switches and the strength of impact on the victim" [Milgram 1974, 157]. In fact, Russell argues the shock

generator was “the most powerful variable in the Obedience studies [...]” [Russell 2018, 244].²

Another important example of a SRM traced back to an observation Milgram made about the first student-run pilot study. That is, Milgram noticed some stressed subjects would, on receiving an incorrect answer, look away from the pained learner—whom they could see through a translucent screen—and then inflict another shock. Milgram therefore wondered if removing this anxiety-inducing visual connection—by placing a solid wall between the teacher and learner—would increase future pilot study completion rates beyond 60%. In July 1961 Milgram tested his hunch during the final condition of his second pilot series, when he ran the so-called *Truly Remote Pilot*. In this experiment subjects were assured the learner was receiving shocks even though, due to the wall separating them, they could not be seen or heard throughout. The results demonstrated that “[...] virtually all subjects, once commanded by the experimenter, went blithely to the end of the board [...]” [Milgram 1965, 61]. With nearly 100% of subjects inflicting every shock, Milgram achieved his preconceived second objective of maximising the completion rate; an outcome that signaled his readiness to run the first official baseline.

Soon afterwards, however, Milgram realised that using the *Truly Remote Pilot* procedure as his official baseline would, on likely obtaining a near 100% completion rate, deprive “us of an adequate basis for scaling obedient tendencies [...]” [1965, 61]. Therefore, to ensure his official baseline experiment obtained a slightly lower completion rate, Milgram decided to introduce a force “that would strengthen the subject’s resistance to the experimenter’s commands [...]” [1965, 61]. With this goal in mind, during the first official baseline Milgram decided to slightly *increase* subject strain—thereby decreasing ambiguity—by including some victim feedback: on “receiving” the 300 and 315-volt shocks, the learner was instructed to kick the wall separating him from the subject.

Nonetheless, this kind of tinkering with the basic procedure—his repeated attempts to relieve subjects of strain, but then suddenly increasing it when desired—illustrates that Milgram had gradually learnt how to gain greater control over how subjects were likely to behave during the first official experiment.³

Having achieved his second objective, three days after the second pilot series, on 7 August 1961 Milgram ran his first official *Remote* (baseline)

2. Try, for example, contemplating the running of a hypothetical *No Shock Generator* condition that includes a non-electrical form of intensifying “punishment” (perhaps a fist or baton) capable of—as the shock generator seemed to—rendering the learner at least unconscious [Russell 2018, 239].

3. Another related learning (or tinkering) technique Milgram utilised to obtain his preconceived high first official baseline completion rate was what can be termed *blocking off all the usual escape routes* [Russell 2018, 64–65].

condition, which generated a 65% completion rate. Soon afterwards, the Obedience research program was transferred to a different laboratory, so Milgram decided to start over and run a new baseline experiment. The *Remote* condition had so easily obtained a high completion rate, Milgram became confident he could run a more disturbing (thus eye-catching) baseline that would still likely obtain a reasonably high completion rate. As described earlier, Milgram then ran the *New Baseline* (or “Cardiac”) condition, which also generated a 65% completion rate. Thereafter, the *New Baseline* became the standard procedure which all subsequent variations were modelled on. With the intention of developing a theory of obedience, Milgram thereafter disengaged with his tinkering and proceeded to undertake a score of *New Baseline* variations, with data collection ending many months later on 27 May 1962 [Russell 2018, 91].

Earlier, in December 1961, Milgram submitted his first journal article presenting the procedure and results from the *Remote* (first) baseline experiment. In this article, titled “Behavioral Study of Obedience”, Milgram presented several factors—but no theory—that might explain the 65% completion rate. It was for this reason the article was rejected twice, with one reviewer—Edward E. Jones—arguing the experiment was, at best, a “triumph of social engineering” [Parker 2000, 112, cited in Russell 2018, 260]. Considering Milgram’s meddling to obtain, by the first official experiment, a preconceived high completion rate, there is clearly much merit to Jones’ assessment. After several amendments, in October 1963 this article was published. The *Remote* experiment’s high completion rate, particularly after the publication of Baumrind’s critique, attracted much scholarly attention, thereby ensuring his article had its intended effect. It transpires that Milgram’s research journey has implications on the topics of ethics and methodology.

6.1 Ethics and methodology are related issues

What Milgram’s pilot research journey suggests is that as his procedure became both more realistic *and* coercive, the greater the likelihood of subjects—more effectively convinced they might be inflicting excruciating shocks on an innocent person—would exhibit signs of intense stress. So, from this, what becomes obvious is that the present article’s two main questions—are the Obedience Studies unethical and do they remain methodologically valid—are intimately related. That is, the more intense “obedient” subjects’ stress reactions, the more potentially harmful and thus unethical Milgram’s experiment became. But as the experiment became more unethical, the more subjects’ stressed reactions indicated Milgram’s attempts at deception had worked, and thus methodologically the better their design. More succinctly,

unethically heightened stress was the key indicator of robust methodological design.⁴ As Patten observed:

[...] if the subjects placidly followed directions without exhibiting discomfort we would have very good grounds for supposing that the dupe was not successful, and so for inferring that the experiments were ill-designed. [Patten 1977, 355]

To avoid this outcome, Milgram rarely allowed ethical concerns over subject welfare interfere with his first objective: to design a procedure capable of deceiving most subjects the shocks were real. But it also transpires Milgram's belief that most subjects needed to be completely fooled by the deception for his experiments to remain methodologically valid was wrong.

6.2 A response to the methodological criticisms

Russell & Gregory have criticised Milgram's original premise that it was of crucial importance that most subjects believed they were harming the learner [Russell & Gregory 2021]. Instead, they argue that the basic procedure only needed to be *realistic enough to generate uncertainty*. To clarify, it is true that for a wide variety of reasons, some—perhaps many—subjects suspected the experiment was a ruse [Perry 2013, 155–165]. Because of these suspicions, these subjects often decided to inflict every shock asked of them because they knew no harm would come of it.

However, because Milgram separated the subject from the learner with a wall and used reasonably skilled actors, it was impossible for subjects to know *for certain* if harm-infliction was fake. These (among other) innovations therefore introduced a strong dose of ambiguity into the basic procedure: teachers, particularly because they could not see the learner, could not be sure of the shocks' reality. Consequently, “many subjects” were left in, as Perry notes, “a state of uncertainty and stress” [Perry 2013, 162]. As one subject put it: “[Hmm] [...] At first I thought maybe it wasn't him yelling but I'm, kinda convinced maybe it is him” [Gibson 2019, 167]. This uncertainty meant any decision by a skeptical subject to complete the experiment required they take a major risk: their suspicion the experiment was a ruse could be wrong and, if so, the learner would be seriously harmed. So, a key question emerges: was it morally correct for subjects to prioritise their mere suspicion the experiment was fake and inflict every shock because everything would, as Orne & Holland put it, probably be “all right”?

There is, it transpires, a morally correct and inherently safe resolution to the dilemma to stop or continue inflicting shocks: if a subject was unsure if

4. Thus, Gibson is right: “Milgram cannot [...] have his cake and eat it: his experiments can either be ethically sound or they can be of profound social importance; they cannot be both” [Gibson 2019, 28].

the experiment was real, the safest (moral) choice was to err on the side of caution and refuse to inflict more shocks [see Coutts 1977, 520, cited in Darley 1995, 133]. Doing so eliminated the risk of being wrong and protected the well-being of a fellow human being. This cautious problem-solving approach, although not common, was exhibited among some subjects across numerous conditions. For example, one suspicious subject later said,

When I decided that I wouldn't go along with any more shocks, my feeling was "plant or not [...] I was not going to take a chance that our learner would get hurt". [Russell & Gregory 2021, 78]

In fact, in the *Relationship* condition one subject explicitly stated he did not believe the shocks were real: "I don't believe this!" [2021, 78]. The experimenter then, (again) diverging from his standardised instructions, cunningly tried to corner the subject into inflicting more shocks: "Well if you don't believe that he's getting the shocks, why don't you just continue[...] and we'll finish it?" [Russell & Gregory 2021, 78]. The subject responded: "[...] I can't take that chance" [Russell & Gregory 2021, 78]. So why could this subject not take that chance? Although he suspected the shocks were probably fake, he could not—largely due to the wall separating him from his friend—be certain. The subject's uncertainty in the ambiguous situation confronting him dictated that he could not afford to "take that chance" because there was still a possibility his hunch might be wrong (a mistake he understood would have devastating consequences for his friend).

Russell & Gregory therefore conclude, in conflict with Milgram, it was not important that most, or even any, subjects *believed* they were hurting the learner. What was important was whether uncertain subjects placed the learner's well-being at risk. When viewed from this perspective, the 35% of subjects who discontinued the *New Baseline* experiment—whether fully deceived or suspicious—were unwilling to place the learner's well-being at risk. Conversely, the 65% of subjects who completed—whether fully deceived or suspicious—were willing to place the learner's well-being at risk. If this argument is valid, it nullifies Orne & Holland's resilient methodological critique; along with the recent studies used to support it, see [Hollander & Turowetz 2017], [Perry, Brannigan *et al.* 2020, 90].

But what about Darley and Perry's criticism that the Obedience Studies lacked standardisation? Although Gibson sees the experimenter's divergence from his set prods as a "reasonable" criticism, he is also somewhat forgiving:

Social scientists who have made a speciality out of the study of scientific practice have long argued that standardisation is more of a rhetorical accomplishment in itself than an achievable methodological goal [...]. [Gibson 2019, 120–121]

Thus, Gibson believes scientific standardisation is not, as it would have been for Milgram, strictly feasible; being more of an aspiration than a

technically achievable destination. Gibson reinforces his wider point by citing Harold Garfinkel's analogy: a map "cannot represent all aspects of that terrain", yet this approximation of the outside world still has enormous utility [Gibson 2019, 121].

What is most important is that despite the Obedience Studies' lapses in standardisation, the basic overview Milgram provided of his own experiments—his rough "map" for others to follow—has, on numerous occasions, been independently replicated. And these replications have typically generated the same surprisingly high completion rates [Blass 2012], [Doris, Niemi *et al.* 2024], [Russell 2018, 126]. It is for this reason that Milgram's basic experimental phenomenon is "remarkably robust" [Gibson 2019, 38] and that "they are not flukes [...]" [Gibson 2019, 68]. It is for the above reasons that the Obedience Studies remain internally valid and from a methodological perspective should continue to be regarded as a study of significance.

There is a final criticism directed at those who have "used the archival materials to effectively hang Milgram out to dry [...]" [Gibson 2019, 67], in that their exposure of Milgram's many unethical indiscretions have likely promoted a "bad apple"-like impression whereby he has become the textbook case of an unscrupulous, deceitful, and dishonest scholar. But missing from this literature is a critique of the arguably tainted leadership barrel that surrounded Milgram, which, hidden from view, condoned and likely encouraged the young scholar's unethical conduct.

6.3 Singular rotten apple or tainted barrel?

Perry notes Milgram's attempts to obtain further rounds of funding from the National Science Foundation (NSF) were rejected because of their "damning" concerns over his unethical treatment of subjects [Perry 2013, 247]. She also points out that after data collection, Milgram, "at Yale's insistence" [Perry 2013, 94], hire a psychiatrist to re-assess the wellbeing of some subjects; which Milgram capitalized on as another opportunity to collect more data [Perry 2013, 238–240, 251–253]. Finally, she highlights that a few years after data collection, Harvard University rejected Milgram's tenure application because the committee believed he was "manipulative" [Perry 2013, 275]. But if one pays greater scrutiny to the actions of leaders from all three of these institutions, it can also be argued that a tainted barrel likely contributed to the rotting of the singular apple.

In his first research proposal to the NSF, dated 25 January 1961, Milgram described his basic procedure and bolstered his research idea's potential by noting that the "Pilot studies [...] yielded unexpected results" that observers apparently found "startling" [Russell 2018, 159]. Such results likely captured the NSF's attention because their core organisational objective is to fund innovative research. To again appeal to the NSF, Milgram then generalised these results to better understanding something of great interest

to a US government agency during the Cold War: the machinations of the “Red Chinese in North Korean POW camps” [Nicholson 2011, 244]. Milgram did not hide the most controversial aspect of his research, informing the NSF that because his experiments relied on deception, they would likely—as they did during the first pilots—generate “extreme tension” in many subjects, adding one person was observed “pulling his hair, gripping his chair, and [...] was extremely uneasy” [Russell 2018, 157]. That said, Milgram reassured his prospective funders: because of this stress, important measures would be undertaken

[...] to insure [*sic*] the subject’s wellbeing before he is discharged from the laboratory. Every effort will be made to set the subject at ease[...]. [Perry 2013, 91]

A few weeks after the NSF’s 13 April 1961 site visit to Milgram’s Yale laboratory [Blass 2004, 71], on 3 May the NSF agreed to fund the Obedience Studies, describing it as “a bold experiment on an important and fundamental social phenomenon” [2004, 72].

Despite having agreed to fund the research, two months later—the beginning of July—Milgram was informed the NSF director Alan T. Waterman had concerns over subject safety and was considering withdrawing the initial funding offer. A few days later, on 6 July 1961, Milgram was told that the NSF committee had overcome their objections and on 13 July Yale University was officially informed that the NSF’s original decision was upheld [Russell 2018, 67]. So why did the NSF change its mind? Before the NSF finalised their decision, they wanted to know “who would be responsible—the National Science Foundation or Yale—for any negative effects on the subjects. The [NSF] lawyer thought that Yale would be legally responsible” [Blass 2004, 71]. Because Milgram would ensure all subjects signed his Yale University “general release form”, it can be argued that measures would be in place to avoid this outcome.⁵

It also transpires the NSF and Yale University were not the only institutions interested in benefit maximisation (benefacting a likely famous research project) whilst evading all potential costs (shouldering responsibility for harming innocent people). That is, when Milgram, a then Yale University faculty member, was completing the Obedience experiments, Gordon Allport, then head of Harvard University’s Department of Social Relations, told a colleague “in a slightly conspiratorial manner: ‘I’m rather glad he’s doing these experiments in New Haven, but we’ll hire him as soon as he finishes’” [Blass

5. Yale University Department of Psychology chair Claude E. Buxton—Milgram’s supervisor—was also aware, early on, that, for unstipulated reasons, the Obedience research might attract negative attention. In a letter dated 14 November 1963, Milgram compliments Buxton for his foresight two years earlier when the latter raised concerns about the first research proposal: “It was, indeed, perceptive of you,” Milgram wrote, “to sense the public relations implication of the study back in December 1960” [SMP, Box 1a, Folder 9].

2004, 131]. Although Allport (officially?) “had a deep moral ambivalence about the ethics of the Obedience experiments”, his first reaction on hearing about them was “pure excitement” [2004, 131]. At Allport’s urging, by the end of 1962 Harvard University indeed hired Milgram as a tenure track assistant professor [2004, 132]. But when academic management allows individual scholars to “make it big” using deceptive research, they—as Allport did—transmit a wider message: it “gives its greatest rewards to those who are ingeniously amoral [...]” [Warwick 1975, 105].

The point being is that Milgram—who, fresh out of graduate school, was only 28 years old when he ran the official experiments—was urged along by the different leadership circles surrounding him. So, when the second wave critics laser focused their attacks only on Milgram, they perhaps inadvertently reinforced the perception that he was a singular bad apple, acting alone. However, because these critics’ largely accurate attacks pay little attention to the surrounding bad leadership barrel, they miss management’s crucial role in tainting the Milgram apple.⁶

7 Conclusion

This article concludes that scholars like Perry, Brannigan, and Nicholson are correct in arguing that Milgram was more unethical in his pursuit of innovative results than was previously imagined. Their groundbreaking research into Milgram’s many misdeeds has provided a sorely needed corrective to the misleading fictions promoted to this day in social psychology textbooks. That said, because this article rejected Orne & Holland’s resilient issue of trust and noted that Milgram’s basic findings—despite lapses in standardisation—have been replicated, it also concludes, in conflict with the above scholars, that Milgram’s research remains methodologically valid.

What his research journey revealed was an unethical researcher who rarely let dangers to subject wellbeing interfere with his quest to undertake a methodologically robust research project. And although Milgram initially planned to provide an unvarnished account of his (mis)treatment of subjects—whom he regularly “[...] observed to sweat, tremble, stutter, bite their lips, groan, and dig their fingernails into their flesh” [Milgram 1963, 375]—after Baumrind published her ethical critique in 1964, his tune changed. Probably concerned about the longevity of his academic career, he presented a softer, occasionally disingenuous, and frequently misleading (whitewashed) version of subject experiences. In doing so, however, Milgram exposed himself to methodological criticisms: if his experiments were not all that stressful for

6. To be fair, Perry mentions the self-interested reason Yale University wanted Milgram to hire a psychiatrist to re-assess some subjects’ wellbeing: they were concerned about the “risks” his experiments posed to their “reputation [...]” [Perry 2013, 236].

subjects, then perhaps they were insufficiently realistic. Milgram would seem to have been caught in a presentational bind of his own making: he could not be honest about how stressful (methodologically realistic) his experiments were without bolstering the arguments of his ethical critics.

Some second wave ethical and methodological critics do not believe Milgram was much of a scientist, with one, as mentioned, instead calling him a *propagandist*. There is certainly much evidence from Milgram's research journey that suggests he was an ambitious showman driven by a quest to make an influential (pseudo-?) scholarly splash. But would these critics be willing to call Milgram what his research journey also shows him to have been: a meticulous goal-directed social engineer who gradually honed his manipulative craft to coerce, persuade and/or tempt ordinary people into (ostensibly) inflicting serious harm on another person?

Although a close examination of Milgram's journey of discovery supports the allegation he was a social engineer, this conclusion also favours potential theoretical connections the second wave critics loathe. For example, it can be argued the most genocidal Nazi managers can also be described as meticulous goal-directed social engineers (and propagandists for that matter), who, like Milgram, gradually learnt through a trial-and-error process of discovery, how to make the seemingly undoable doable [Russell 2018, 2019]. Perhaps this purported Milgram-Holocaust connection deserves greater critical attention.

Acknowledgments

The author would like to thank Emeritus Professor Robert Gregory of Victoria University of Wellington, New Zealand, for his comments. All responsibility rests with the author.

Bibliography

- BAUMRIND, Diana [1964], Some thoughts on ethics of research: After reading Milgram's "Behavioral Study of Obedience", *American Psychologist*, 19(6), 421–423, doi: 10.1037/h0040128.
- BLASS, Thomas [2004], *The Man Who Shocked the World: The life and legacy of Stanley Milgram*, New York: Basic Books.
- [2012], A cross-cultural comparison of studies of obedience using the Milgram paradigm: A review, *Social and Personality Psychology Compass*, 6(2), 196–205, doi: 10.1111/j.1751-9004.2011.00417.x.
- BRANNIGAN, Augustine [2020], *The Use and Misuse of the Experimental Method in Social Psychology: A Critical Examination of Classical Research*, London: Routledge, doi: 10.4324/9781003034803.

- BRANNIGAN, Augustine, NICHOLSON, Ian, *et al.* [2015], Introduction to the special issue: Unplugging the Milgram machine, *Theory & Psychology*, 25(5), 551–563, doi: 10.1177/0959354315604408.
- DARLEY, John M. [1995], Constructive and destructive obedience: A taxonomy of principal-agent relationships, *Journal of Social Issues*, 51(3), 125–154, doi: 10.1111/j.1540-4560.1995.tb01338.x.
- DE SWAAN, Abram [2015], *The Killing Compartments: The Mentality of Mass Murder*, New Haven: Yale University Press.
- DORIS, M., John, NIEMI, Laura, *et al.* [2024], The incredulity hypothesis and the enduring legacy of the obedience experiments, *Philosophia Scientiæ*, 28(2), 53–89.
- ECKMAN, Bruce K. [1977], Stanley Milgram's obedience studies, *Et Cetera*, 34, 88–99.
- FREEDMAN, Jonathan L. & FRASER, Scott C. [1966], Compliance without pressure: The foot-in-the-door technique, *Journal of Personality and Social Psychology*, 4(2), 195–202, doi: 10.1037/h0023552.
- GIBSON, Stephen [2019], *Arguing, Obeying and Defying: A Rhetorical Perspective on Stanley Milgram's Obedience Experiments*, New York: Cambridge University Press, 1st edn., doi: 10.1017/9781108367943.
- HARRE, Rom [1979], *Social Being: A Theory for Social Psychology*, Oxford: Basil Blackwell.
- HOLLANDER, Matthew M. & TUROWETZ, Jason [2017], Normalizing trust: Participants' immediately post-hoc explanations of behaviour in Milgram's "obedience" experiments, *British Journal of Social Psychology*, 56(4), 655–674, doi: 10.1111/bjso.12206.
- KAPOSI, David [2017], The resistance experiments: Morality, authority and obedience in Stanley Milgram's account, *Journal for the Theory of Social Behaviour*, 47(4), 382–401, doi: 10.1111/jtsb.12137.
- MCGAHA, Annette Christy & KORN, James H. [1995], The emergence of interest in the ethics of psychological research with humans, *Ethics & Behavior*, 5(2), 147–159, doi: 10.1207/s15327019eb0502_3.
- MILGRAM, Stanley [1963], Behavioral study of obedience, *The Journal of Abnormal and Social Psychology*, 67(4), 371–378, doi: 10.1037/h0040525.
- [1964], Issues in the study of obedience: A reply to Baumrind, *American Psychologist*, 19(11), 848–852, doi: 10.1037/h0044954.

- [1965], Some conditions of obedience and disobedience to authority, *Human Relations*, 18(1), 57–76, doi: 10.1177/001872676501800105.
- [1965a], *Obedience (Film)*, New York: New York University Film Library.
- [1972], Interpreting obedience: Error and evidence. A reply to Orne and Holland, in: *The Social Psychology of Psychological Research*, edited by A. G. Miller, New York: Free Press, 138–154.
- [1973], The perils of obedience, *Harper's*, December, 62–66; 75–77.
- [1974], *Obedience to Authority: An experimental view*, New York: Harper and Row.
- NICHOLSON, Ian [2011], “Torture at Yale”: Experimental subjects, laboratory torment and the “rehabilitation” of Milgram’s “Obedience to Authority”, *Theory & Psychology*, 21(6), 737–761, doi: 10.1177/0959354311420199.
- ORNE, Martin T. & HOLLAND, Charles Howard [1968], Some conditions of obedience and disobedience to authority. On the ecological validity of laboratory deceptions, *International Journal of Psychiatry*, 6(4), 282–293.
- PATTEN, Steven C. [1977], The case that Milgram makes, *The Philosophical Review*, 86(3), 350–364, doi: 10.2307/2183787.
- PERRY, Gina [2013], *Beyond the Shock Machine: The Untold Story of the Milgram Obedience Experiments*, Melbourne: Scribe.
- [2015], Seeing is believing: The role of the film *Obedience* in shaping perceptions of Milgram’s Obedience to Authority experiments, *Theory & Psychology*, 25(5), 622–638, doi: 10.1177/0959354315604235.
- PERRY, Gina, BRANNIGAN, Augustine, *et al.* [2020], Credibility and incredulity in Milgram’s obedience experiments: A reanalysis of an unpublished test, *Social Psychology Quarterly*, 83(1), 88–106, doi: 10.1177/0190272519861952.
- RUSSELL, Nestar [2011], Milgram’s obedience to authority experiments: Origins and early evolution, *British Journal of Social Psychology*, 50(1), 140–162, doi: 10.1348/014466610x492205.
- [2018], *Understanding Willing Participants. Milgram’s Obedience Experiments and the Holocaust*, vol. 1: Milgram’s Obedience Experiments and the Holocaust, Cham: Palgrave MacMillan, doi: 10.1007/978-3-319-95816-3.
- [2019], *Understanding Willing Participants. Milgram’s Obedience Experiments and the Holocaust*, vol. 2, Cham: Springer, doi: 10.1007/978-3-319-97999-1.

RUSSELL, Nestar & GREGORY, Robert [2021], Are Milgram's obedience studies internally valid? Critique and counter-critique, *Open Journal of Social Sciences*, 09(02), 65–93, doi: 10.4236/jss.2021.92005.

WARWICK, Donald [1975], Deceptive research: Social scientists ought to stop lying, *Psychology Today*, 8(9), 105–106.