

Are Milgram's Obedience Studies Internally Valid? Critique and Counter-Critique

Nestar Russell¹, Robert Gregory²

¹University of Calgary, Calgary, Canada

²Victoria University of Wellington, Wellington, New Zealand

Email: nestar.russell@ucalgary.ca, bob.gregory@vuw.ac.nz

How to cite this paper: Russell, N., & Gregory, R. (2021). Are Milgram's Obedience Studies Internally Valid? Critique and Counter-Critique. *Open Journal of Social Sciences*, 9, 65-93.

<https://doi.org/10.4236/jss.2021.92005>

Received: October 29, 2020

Accepted: February 5, 2021

Published: February 8, 2021

Copyright © 2020 by author(s) and
Scientific Research Publishing Inc.

This work is licensed under the Creative
Commons Attribution-NonCommercial
International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

This article challenges the most significant methodological criticism directed at Milgram's obedience studies, namely, that they lack internal validity because most obedient subjects probably did not believe that the "learner" was actually receiving dangerous electric shocks (Orne & Holland, 1968). This criticism has been bolstered recently by data that claims to show that this was indeed the case (Perry et al., 2020; Hollander & Turowetz, 2017). We argue instead that while Milgram's experimental paradigm has minor methodological flaws, the resilient issue of believability is actually a red herring, because Milgram's procedure ensured subjects remained uncertain about the reality of the shocks they were ostensibly delivering. This uncertainty forced all subjects into resolving the experiment's inherent moral dilemma. That is, would they prematurely end a potentially real experiment and secure the learner's safety? Or would they continue to inflict "shocks" they believed were perhaps, probably, or even most certainly fake, thus still running the risk of potentially being wrong? We believe the obedience experiments remain, for the most part, internally valid, and that they continue to be externalisable to other moral dilemmas. They help in understanding the perpetration of the Holocaust, contrary to the opposite claim made by some of Milgram's critics.

Keywords

Obedience, Milgram, Internal Validity, Methodology, Believability, Moral Dilemma

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 in New Haven, Connecticut.

—Abram de Swaan (2015: p. 28).

1. Introduction

In the first author's undergraduate class *Genocide and Terrorism*, a set of lectures explore social psychologist Stanley Milgram's "obedience to authority" experiments. In 2018, students were asked in the course evaluation: "In what ways do you think this course could be improved for your learning?" One student responded: "USING ACTUAL THINKERS WHO AREN'T THE LAUGHINGSTOCK OF THEIR DISCIPLINE[.]" This kind of critical response is the emerging reality in university classes that continue to present what are three of the most famous studies in social psychology: the obedience studies, Sherif's robber's cave study, and Zimbardo's Stanford prison experiment. This is because in the last few years all three studies have attracted heated criticisms that have attempted to undermine their internal validity (Brannigan, 2020; Brannigan et al., 2015; Haslam & Reicher, 2012; Perry, 2012, 2018; Le Texier, 2019). Consider, for example, Le Texier (2019) whose archival research questions the accuracy of Zimbardo's presentation of the Stanford prison experiments. In the prestigious *American Psychologist*, Le Texier accuses Zimbardo of stealing a student's research idea, instructing the guards to engage in prisoner abuse, and misleadingly (fraudulently?) presenting the guards' brutal compliance as a natural reaction to the situation. The paper concludes that the Stanford prison experiment was "an incredibly flawed study that should have died an early death" (2019: p. 14).

Any lecturers—and textbook authors for that matter—who continue to present the original versions of these three studies in the absence of the contemporary critical literature will be performing a serious disservice to students. In such cases, some students may, as the above one does, critique the content of their teachers' courses. But with the obedience studies having recently attracted a barrage of criticism (Brannigan, 2013; Brannigan et al., 2015; Griggs & Whitehead, 2015; Gibson, 2013; Nicholson, 2011; Perry, 2012, 2015; Russell, 2009, 2011, 2014a, 2014b, 2018a; Russell & Gregory, 2011, 2015), is it fair to apply Le Texier's potentially ruinous conclusion about the Stanford prison experiment to Milgram's obedience research?

As conveyed by the title of one series of journal articles "Unplugging the Milgram Machine" (Brannigan et al., 2015), it can be argued that some scholars are indeed working toward this end. And as the above student's comment suggests, they may be hitting their mark. Putting aside the obedience study's indisputably unethical nature (Baumrind, 1964, 1985, 2013, 2015; Nicholson, 2011; Perry, 2012; Russell, 2009, 2014b, 2018a), is such a representation fair? Or is there still something of importance that can be learnt from Milgram's obedience studies? As this article will show, although our past research illustrates we are no blinkered admirers of the obedience studies, we are also of the view that—for the most part—the study remains internally valid and, thus, despite the critical content of the latest literature, it should continue to be regarded as a study of significance. After presenting Milgram's basic findings, the aim of this paper is to present, review and critically assess the more important literature challenging

the internal validity of the obedience studies.

2. Milgram's Results and the Deception of Most Subjects

In the early 1960s Milgram published the first of his many obedience study variations. Termed the *Remote* condition (1963), in this experiment a confederate, posing as a potential subject, enters a laboratory where he encounters a scientist wearing a lab coat (actually another confederate, termed "the experimenter"). The ostensible subject is then introduced to a waiting naïve, and actual, subject. The experimenter informs both persons that they have volunteered to take part in an experiment investigating the effects of punishment on learning. They are told one person is required to be the teacher and the other the learner. The selection is rigged to ensure the confederate is always the learner and the subject the teacher. The subject watches as the experimenter "straps" the learner into an "impossible ... to escape" chair and then attaches an electrode to his wrist (Milgram, 1963: p. 373). "In order to improve credibility the experimenter declared, in response to a question by the learner: 'Although the shocks can be extremely painful, they cause no permanent tissue damage'" (Milgram, 1963: p. 373).

The subject is then taken into an adjacent room and placed before a shock generator, a device with 30 switches aligned in 15-volt increments from 15 to 450 volts. The subject is given a sample 45-volt shock (reinforcing the impression that this machine is what it appears to be) and is then instructed by the proximate experimenter to give the learner a shock each time an incorrect answer is proffered. Each incorrect answer warrants for the learner a shock one level higher than its predecessor. Before long, compliant subjects find themselves inflicting shocks of high intensity. When subjects reached the 300 and 315-volt shock switches, the learner would bang suddenly on the wall, thereafter falling silent. The experimenter responded to any of the subjects' concerns surrounding the learner's wellbeing by urging that they continue inflicting more shocks. The learner's silence, however, implied the shocks had perhaps rendered them unconscious. Despite this, the experimenter instructed the subject to treat all subsequent unanswered questions as incorrect and accordingly inflict the next shock. In the absence of any acts of defiance, the experiment was deemed complete upon the subject administering three successive 450-volt shocks. This experiment generated a 65% completion rate ($n = 40$).

Although the *Remote* condition is probably the most widely cited of Milgram's many experimental variations¹, it was the so-called *New Baseline* condition that became the model procedure that all subsequent variations were based on. The *New Baseline* condition differed to the *Remote* condition in that the learner stated before starting that he had a mild heart condition. Also, the learner's "pain" was unambiguously conveyed by way of increasingly intense verbal protests and eventually screams (actually standardized tape recordings).

¹Milgram's (1963) *Remote* condition has been cited over 7000 times.

For example:

120 volts: “Ugh! Hey, *this* really hurts.”

270 volts: (*Agonized scream.*) “Let me out of here. Let me out of here. Let me out of here. Let me out. Do you hear let me out of here.”

300-volts: (*Agonized scream.*) I absolutely refuse to answer any more. Get me out of here. You can’t hold me here. Get me out. Get me out of here.

(Milgram, 1974: pp. 56-57)

After the 345-volt switch onwards, the learner is silent. As in the *Remote* condition, the experimenter instructed the subject to treat all subsequent unanswered questions as incorrect and accordingly inflict a shock at the next level until three successive 450-volt shocks had been administered. Although the *NewBaseline* experiment was likely to have been an even more harrowing subject experience, it also elicited a 65% completion rate ($n = 40$).

After undertaking the *New Baseline*, Milgram undertook more than 20 slight baseline variations. One important variation, as far as this article is concerned, was the *Subject Chooses Shock Level* condition. For incorrect answers, subjects in this variation were instructed to inflict shock-intensities of their choosing (with the experimenter exerting no coercive pressure). The results showed that 97.5% of subjects ($n = 40$) repeatedly chose to inflict low-level shocks (Milgram, 1974: pp. 70-72). Another key variation was the *Relationship* condition. Here subjects were encouraged to inflict intensifying shocks on a learner who was at least an acquaintance, often a friend, and occasionally a family member (prior to starting the experiment Milgram covertly informed the learner of the experimenter’s actual purpose and then instructed them how to react to being “shocked”). The *Relationship* condition generated a 15% completion rate ($n = 20$) (Russell, 2014b). Milgram chose not to publish this variation, and it was first discovered by Rochat & Modigliani (1997) in Milgram’s personal archive held at Yale University.

When the basic experimental procedure is deconstructed, it becomes clear that with each incorrect answer received, Milgram was trying to force subjects into siding with either the experimenter, who wanted them to continue inflicting shocks, or the learner, who wanted them to stop doing so. The subject cannot simultaneously please both parties and must decide (as empirically confirmed through their actions) whether it is more important to inflict shocks to help the experimenter obtain his results, or to stop inflicting intensifying shocks on an innocent person. These mutually exclusive options were designed to generate within subjects a “conflict” of conscience (Milgram, 1963: p. 378), as conveyed by the following subject in his post-experimental interview:

when he [the experimenter] said “continue” I was thinking your side of it too... you’re trying to get some scientific information, and I had to balance in my own mind ... will there be any ... real harm to ... the fella on the other side versus what value it’s going to be to you. (SMP, Box 153, Audio-

tape #2306; see also, [Hollander & Turowetz, 2017: p. 664](#))

The successful deception of most subjects into believing that they were harming the learner was therefore “absolutely critical” ([Miller, 1986: p. 143](#)) to the methodological strength (internal validity) of the experiments. Deception was important because if most subjects did not believe the learner experienced “any real harm” then there would not have been any conflict over whether or not subjects should continue helping the experimenter collect his “scientific information...” Methodologically speaking, “the entire foundation of the obedience research rests on the believability of the victim’s increasingly mounting suffering.” ([Miller, 1986: p. 143](#); see also [Milgram, 1972: p. 139](#))

In his first publication, Milgram was convinced he had successfully deceived most subjects because they typically exhibited signs of “*extreme tension*”: sweating, trembling, stuttering, groaning, biting their lips and digging their nails into their flesh (1963: p. 375). Also, during the post-experimental interviews subjects were asked to indicate on a 14-point scale ranging from *Not at all painful* to *Extremely painful* the learners level of pain, producing a mean response of 13.42 (1963: p. 375). Successfully deceiving most subjects was also of great importance to Milgram because it made the next step in his research possible: generalizing his findings to the outside world (ecological or external validity). In the first page of his 1963 publication Milgram quoted C. P. Snow in regard to the relationship between ingrained obedience to authority and the Nazi regime.

Of course, if critics of the obedience studies could show that Milgram’s basic procedure failed to successfully deceive most subjects, then he could no longer generalize his results beyond the confines of his laboratory walls. Put differently, one cannot generalize to, say, the Holocaust and beyond, from a methodologically weak foundation. And over the past half century or so there has been no shortage of scholars who have challenged the above central methodological foundation on which the obedience studies were built. What follows is a brief review of the early critical literature.

3. Deception and Trust: The Early Critiques

[Orne & Holland \(1968\)](#) authored the first methodological critique exploring whether or not Milgram successfully deceived his subjects into believing the learner received intense shocks. They argued that Milgram’s attempts at deception likely failed because subjects would have known, if only vaguely, that the experimenter (or, especially, Yale University) would not have allowed the learner to be exposed to such danger. Therefore, subjects would have presumed that, despite evidence to the contrary, the final outcome would be “all right” ([Orne & Holland, 1968: p. 287](#)). As one of potentially many subjects who completed the experiment later said, “the way I figured it, you’re not going to cause yourselves trouble by actually giving serious physical damage to a body.” (SMP, Box 153, Audiotape #2430)

Yet if subjects did not believe the learner was being shocked, why did most go

to the trouble of completing the experiment? Orne & Holland argue that most subjects completed the *Remote* condition because they were influenced by what Orne (1962: p. 779) termed “demand characteristics.” Demand characteristics emerge when subjects attempt to detect the meaning and purpose of an experiment and then engage in behaviors they think are likely to please the researcher by confirming the probable hypothesis. In such a scenario, most subjects, realizing the experiment was actually about obedience to authority, and knowing the learner remained unharmed, likely felt obliged to give Milgram the high completion rates he desired. So how did Orne & Holland’s argument fare?

Milgram (1972: p. 141) responded to Orne & Holland by presenting new evidence over the issue of whether or not subjects believed they were harming an innocent person: although a small minority did not believe the shocks were genuine, 56.1% and 24% of subjects, reported in the post-experimental interviews that they “fully” and “probably” believed they were real, respectively. Therefore, Milgram was convinced most subjects (80.1%) were sufficiently deceived by the procedure. In relation to the accusation of demand characteristics having generated the *Remote* condition’s result, it should be remembered that Milgram withheld from subjects his study’s actual objective of determining people’s willingness to obey malevolent authority: the experiment was (apparently) exploring the effects of punishment on learning. So, unlike methodologically weaker studies, like Zimbardo’s, it can be argued that Milgram’s subjects were prevented from gratifying the researcher’s wishes and/or engaging in some pre-conceived role-play (Nussbaum, 2007). Also, Russell (2011: pp. 153-160; 2018a: pp. 61-74) has detailed the great lengths Milgram went to during his winter 1960/1961 pilot studies to increase the believability of his official baseline procedure. Finally, if subjects in the obedience studies knew the shocks were fake, why after completing the *New Baseline* condition did 73% decline an opportunity to experience the apparently fake 450-volt shock they had just inflicted on the learner (Milgram, 1974: p. 57)?

To Orne & Holland, Milgram (1972: p. 140) responded with frustration: their “suggestion that the subjects only feigned sweating, trembling, and stuttering to please the experimenter is pathetically detached from reality, equivalent to the statement that hemophiliacs bleed to keep their physicians busy.” Others, like Eckman (1977: p. 94) agreed: “when one reads the actual transcripts of Milgram’s subjects’ verbal behavior, it is hard to conclude it was all a put-on. There was just too much conflict and stress.”² We therefore agree with Eckman (1977: p. 95): “To invoke the charge ‘demand characteristics’ against Milgram’s work is foolish” (see also Kaposi, 2017: p. 384). In his book *Obedience to Authority: An Experimental View*, an even more confident Milgram generalized his results to not just the Holocaust, but also the American Civil War and the My Lai massacre in Vietnam (1974: pp. 175-178; 186; 183-186). Despite Milgram’s rebuttal of

²Orne & Holland (1968) cite results from Holland’s unpublished PhD thesis as an alternative means of explaining away many of Milgram’s subjects’ displays of stress. To date, however, this thesis remains unpublished.

Orne & Holland, the latter's issue of trust—subjects knew the experiment was fake—lingered (see, for example, [Harré, 1979: p. 105](#)).

Then across a series of publications in the 1970s and 1980s, Don Mixon's research lent support to Orne & Holland's issue of trust. Mixon replicated the obedience studies using role-playing, a methodological technique that enabled him to avoid the potentially unethical use of deception. Before starting his role-play replication, Mixon informed his subjects that the learner was an actor, the shocks were not authentic, and that they were to pretend the experiment was real. In conflict with Orne & Holland, Mixon argues that audiences find the obedience studies convincing because of the subjects' outpouring of emotional tension ([Mixon, 1976: p. 93](#)). The subjects' palpable displays of stress appear to illustrate that they believed the shocks were really harmful. But, according to Mixon, Milgram's subjects were stressed not because they confronted an intense moral dilemma to stop or continue inflicting shocks, but because of their exposure to an *ambiguous situation* in which the information coming from the experimenter and learner was contradictory. The shocks were apparently harmful, but not dangerous; the experimenter was calm, but the learner was screaming in agony. As Mixon observed, "No wonder many subjects showed such stress. What to believe? The right thing to do depends on which actor is believed" ([Mixon, 1989: p. 33](#)). According to Mixon, because "increasingly large chunks of the social and physical world that we live in can be understood only by experts" ([Mixon, 1989: p. 35](#)), most subjects chose to resolve the stressfully ambiguous situation by trusting the authority figure's word that the learner would not be hurt (see also [Baumrind, 1985: p. 171](#)). Here Mixon's argument is similar to that of Orne & Holland, in the sense that both believe that many subjects invested great trust in the perceived expertise of the Yale-based experimenter.

Mixon then argued that his claim that subject stress was generated by confusion can be verified by testing the following assertion: when the consequences of inflicting harm are unambiguous, subjects will not complete the procedure. Mixon provided several lines of evidence that he believed lent weight to this claim. First, Mixon claimed that in Milgram's three least ambiguous variations, where it was (apparently) most clear that if subjects continued up the shock board, the learner would definitely be hurt, all subjects refused to complete. Second, the more ambiguous the learner's fate appeared, the higher the completion rates ([Mixon, 1976: pp. 92-94](#)). Third, in a role-play replication undertaken by Mixon that removed all ambiguity surrounding the learner's fate (the experimenter informed the subject that "The learner's health is irrelevant ... continue as directed"), completions slumped ([Mixon, 1972: p. 164](#)). In sum, Mixon argued it was the purposefully ambiguous situation—not, as Milgram believed, the subjects' eventual resolution of a moral dilemma to engage in wrongdoing—that generated the obedience study's high completion rates.

In support of Mixon, during his pilot studies Milgram injected ambiguity into the emerging basic procedure, and he did this with the intention of maximizing

the rate of obedience (Russell, 2011). For example, he changed the title of the last button on his shock machine from “LETHAL” in the pilots to “XXX” in the official experiments. Likewise, he substituted a translucent screen between the subject and learner in the first pilot study with a fully opaque partition. Each change Milgram introduced created a more ambiguous, and therefore more confusing and stressful situation where, to alleviate that stress, subjects might have chosen to side with the only expert present.

That said, Mixon’s methodological critique is not without its own explanatory problems. First, in Milgram’s *Touch-Proximity* condition, where the subject could directly hear, see and eventually had to touch the learner to shock them, 30% inflicted every shock ($n = 40$). This result conflicts with Mixon’s assertion that when the consequences of inflicting harm are unambiguous, subjects will not complete the obedience studies. Second, Mixon does not explain the basis of his selection of what he believed were Milgram’s three least ambiguous conditions which, according to him, were the *Authority as Victim* ($n = 20$), *Two Authorities (Contradictory Commands)* ($n = 20$), and *Learner Demands to be Shocked* ($n = 20$) conditions (Mixon, 1976: p. 95). In this last condition, for example, the experimenter informed the subject to stop inflicting shocks at 150 volts, but the pained learner demanded that the subject inflict further shocks, because he (the learner) wanted to endure more shocks than his friend who had been a learner in an earlier trial. Why in this variation is it much clearer than in the *New Baseline* that the learner is more likely to be receiving shocks? Put differently, how is the *Learner Demands to be Shocked* variation slightly, let alone significantly, less ambiguous than the *New Baseline*? It seems to us that Mixon selected the above three conditions because they all ended in 0% completion rates, and he then assumed they must have been the least ambiguous of all variations. Third, as Miller (1986: p. 175) observes, a weakness with Mixon’s role-play methodology is that although it enables researchers to gain the subjects’ informed consent and circumvents many of the ethical problems associated with deception, it ironically increases the probability of the results being influenced by Orne & Holland’s (1968) demand characteristics.

Although Mixon’s critique has its weaknesses, the combination of Orne & Holland’s issue of *trust* with Mixon’s point about ordinary people in confusing (*ambiguous*) situations deferring to the authority of specialist *experts*, led to the formation of what we term the *trust-ambiguity-expertise nexus*. As we will show, the *trust-ambiguity-expertise nexus* proved highly persuasive in the obedience study research area and continues to exert a powerful influence on the contemporary literature. It is to this latest research that we now turn.

4. A Critique of the Contemporary Literature

In 2017 Hollander & Turowetz published a journal article titled *Normalizing trust: Participants’ immediately post-hoc explanations of behaviour in Milgram’s “obedience” experiments*. In it, the authors analyzed all of the available recorded

interviews Milgram conducted on subjects after their completion of his various experimental conditions. Hollander & Turowetz were particularly interested in identifying the subjects' post-experimental justifications for obedience and disobedience. Of the 117 available recordings held at Milgram's Yale archive (across the *Voice Feedback*, *Proximity*, *Women Only* and the less complete *Bridgeport* and *Relationship* conditions), 91 different subjects (46 obedient; 45 defiant) explicitly provided one or more justifications for their actions. As far as the present article is concerned, Hollander & Turowetz's (2017: p. 660) most interesting discovery was that although there were various justifications provided for completing, "[t]he most common 'obedient' explanation is *L[earner] not being harmed*, as 33 (of 46 total) 'obedient'-outcome participants used it at least once (72%)." That is, nearly three-quarters of the obedient subjects in their sample "continued because they did not think the situation was dangerous, despite appearances to the contrary" (2017: p. 662).

This evidence supports the Orne-Holland-Mixon *trust-ambiguity-expertise nexus*: during the confusing (ambiguous) situation, 72% of subjects who completed trusted that a professional like the experimenter would not allow the learner to be harmed. As the authors put it:

Although [obedient] participants' numerous post-experimental claims that *L[earner] was not really being harmed* all worked to make sense of the ambiguous situation, they did so in different ways. [...] we analyse thematic differences among the *L[earner] not really being harmed* accounts, which normalize the situation by 1) trusting E[xperimenter]'s judgement, 2) treating L[earner] as overreacting, and 3) doubting the cover story. Though contrasting in theme, these accounts all justify T[eacher]'s continuation in terms of the perception that L[earner] faced no serious danger. (Hollander & Turowetz, 2017: p. 666)

The reason these obedient subjects thought the learner faced no serious danger was because they believed the experimenter to be "sufficiently competent to have kept L[earner] from harm." (2017: p. 663). As one subject put it: "↑there: fore what I am doing (1.3) uh:: is ↑probably alright ..." and, as another subject said, "If it was ↑THAT serious you woulda ↑stopped me." (2017: p. 663).

It must be noted that Hollander & Turowetz's paper is not a methodological but a theoretical critique directed at what is currently the leading contemporary explanation of the obedience studies: *Engaged Followership* theory (Haslam et al., 2014; Haslam et al., 2015; Reicher et al., 2012). Engaged Followership theory suggests that obedient subjects were aware they were inflicting harm on an innocent person but continued to inflict shocks because they came to see merit in and identified with the experimenter's noble "scientific" cause of determining the effects of punishment on learning. But as Hollander & Turowetz (2017: p. 655) note, most of their sample of "obedient" subjects did not explain themselves by identifying with science but instead "justified compliance in several distinct and not entirely consistent ways"—most commonly completing because they did

not think the learner was being harmed. If this is true, Hollander & Turowetz argue, how can a so-called belief in science have been the driving force behind their compliance?

Although their article is solely a theoretical critique of *Engaged Followership* theory, Hollander & Turowetz seem to be unaware that, as a methodological critique of Milgram's original results, their findings inadvertently provide potentially powerful evidence in favour of the *trust-ambiguity-expertise nexus*. As just stated, of all the post-experimental interviews they examined where obedient subjects justified their actions, 72% stated at least once something to the effect that they thought the "*L[earner] was not really being harmed*" (Hollander & Turowetz, 2017: p. 662). Because Hollander & Turowetz seem unaware of the historic Orne, Holland, Mixon versus Milgram methodological debate, it took scholars who are aware of it—like Perry and Brannigan (see Perry, 2012: p. 69-72; Brannigan, 1997: p. 604)—to sense this potential in the formers' findings. Three years after Hollander & Turowetz's publication, Perry et al. (2020: p. 92) note:

In his first journal article about his obedience research, Milgram (1963) stressed the dramaturgical credibility of the experiment. He emphasized that "[w]ith few exceptions subjects were convinced of the reality of the experimental situation" (Milgram, 1963: p. 375). [...] The implication was that the subjects fully believed that what was happening was real, and despite indications that the learner was in increasing pain, 26 out of 40 proceeded to administer the maximum shock.

After Perry et al. reference Milgram's old methodological foes—"Mixon 1977; Orne and Holland 1968" (p. 89)—they then direct the reader to Hollander & Turowetz's key finding that 72% of obedient subjects later "did not believe the learner was actually being harmed", adding, "[t]his finding was astonishing in light of Milgram's... assurances of the credibility of the cover story." (p. 90).

Perry et al., however, do more than highlight the potential importance of Hollander & Turowetz's inadvertent methodological critique of the obedience studies. They also add their own, perhaps even more impressive, archival evidence to the methodological debate. That is, they note that in 1962 Milgram instructed his research assistant Taketo Murata to analyse the post-experimental interview survey data and do so with a specific focus on the issue of subject deception. Murata soon after presented Milgram with his report (Perry et al., 2020: p. 93; see also Hoffman et al., 2015: pp. 678-679). Perry et al.'s contribution to the literature is their analysis of this archival document:

[Milgram] asked Murata to compare the degrees of obedience between those subjects who said they were doubtful that the shocks were painful and those who were certain they were. [...] Murata (1962: 1) wrote, "The following is a condition-by-condition analysis to determine whether shock level reached was affected by the extent to which the subject believed that

the learner was actually receiving shock.” [...] The report’s main results were a comparison of the mean shock levels administered by subjects who had “fully believed” (FB) in the reality of shocks versus those classified as having “not fully believed” (NFB). [...] Murata found that in 18 of 23 experiments, those subjects who fully believed the learner was getting painful shocks gave *lower* levels of shock than subjects who doubted the shocks were real. (Perry et al., 2020: p. 94)

The reverse also applies: “Those who were less successfully convinced by the cover story were more obedient” (Perry et al., 2020: p. 99). Put differently: “obedience increases with skepticism of pain” (2020: p. 98), a conclusion that across 18 out of 23 conditions not only bolsters Hollander & Turowetz’s inadvertent methodological critique, but the *trust-ambiguity-expertise nexus* more generally. Perhaps because Murata’s report undermined the methodological strength of the entire obedience research project, Milgram—later rightfully criticised for his “cherry-picking of findings” (Brannigan et al., 2015: p. 551)—chose to ignore the report’s results. Milgram was probably not persuaded by Murata’s report because he (Milgram) suspected the obedient subjects’ common justification that they did not believe the learner was being harmed was actually a “defense function” designed to save face (Milgram, 1974: p. 172).

What are the consequences of Perry et al.’s argument? Alluding to the contemporary research critical of the obedience studies—of which their article is the latest addition—they argue “the current series of critiques may be serious enough to warrant the reconsideration of the studies in toto” (Perry et al., 2020: p. 103). Perry et al. (2020: p. 90, 105) also caution against the “many current scholarly narratives” that “accept Milgram’s assumption that we are all capable of torture and murder at the behest of an authority figure”, here singling out the first author’s 2018 books “*Understanding Willing Participants: Milgram’s Obedience Experiments and the Holocaust, 2 vols*” (Russell, 2018a, 2018b). On this note, Hollander & Turowetz would support Perry et al. in that they are critical of an antecedent of the above book: Russell & Gregory’s (2011) article *Spinning an Organizational “Web of Obligation”? Moral Choice in Stanley Milgram’s “Obedience” Experiments*. More specifically, according to Hollander & Turowetz (2017: p. 659):

Russell and Gregory (2011) argue that most “obedient” participants knew they were engaged in wrongdoing, but that (after 150 V) it was psychologically easier to continue than to stop. When members of this group later seek to explain themselves, their justifications may be consciously self-serving, akin in this respect to Holocaust perpetrator testimony.

Holland & Turowetz disagree, arguing obedient subjects’ post-experimental accounts should not necessarily be dismissed as evasive or defensive (face-saving) because during the “highly ambiguous situation” (2017: p. 659) it was not clear to many of them that what they had done was wrong. At the end

of their experimental trials, these obedient subjects were more likely left struggling to make sense of the fast-paced and confusing experiments (2017: p. 659). Furthermore, [Hollander & Turowetz \(2017: pp. 659-660\)](#) believe the obedience study's post-experimental interviews stand:

in marked contrast to paradigmatic Holocaust testimony, in which perpetrators—often in situations of confinement and trial, knowing some type of punishment was likely—were interrogated and cross-examined by their captors [...] We therefore think it reasonable to take participants' self-justifications seriously, while remaining aware of the possibility that not all are being perfectly candid.

5. Our Response to the Methodological Critics

What follows is our two-part response to the contemporary research that is critical of the internal validity of the obedience studies. First, we will assess the merits of the Orne-Holland-Mixon *trust-ambiguity-expertise nexus* and its influence on the contemporary critics' views of the obedience studies. Second, we argue that the contemporary scholars—specifically Hollander & Turowetz, Perry et al., but also *Engaged Followership theory's* Haslam & Reicher—have too readily accepted Milgram's obedient subjects' post-experimental survey responses and other justifications as accurate. In this section we explain why we think all scholars should treat sceptically the accuracy of most of what the obedient subjects said after the experiments.

5.1. Assessing the Merits of the Trust-Ambiguity-Expertise Nexus

In their article, [Perry et al. \(2020: p. 100\)](#) note:

It was Milgram whose research raised the issue of obedience and its responsiveness to perceptions of harm. *If his premise was correct*, then our work has raised an ironic linkage between belief and obedience that was unexplored by Milgram and the obedience literature. (Emphasis added).

Based on the Murata report that Perry et al. present in support of their argument, this conclusion seems fair and reasonable. However, below we will argue Milgram's premise (that it was critically important most subjects believed they were harming the learner) is flawed. What we are suggesting here is that—*pace* Milgram—it was not important that most, or even any, subjects believed they were hurting the learner. The assertion that Milgram's premise is flawed obviously places us in direct conflict with the *trust-ambiguity-expertise nexus*. To demonstrate the validity of our claim, we must briefly return to Mixon's research, whose work built on Orne & Holland's issue of trust, completing the formation of the *trust-ambiguity-expertise nexus*.

As we have seen, Mixon argued Milgram's inherently ambiguous baseline procedure ensured that subjects could not be certain they were inflicting real shocks on the learner. This ambiguity was accentuated by the inability of sub-

jects to see the learner because of the partition separating them (a partition, it transpires, Milgram purposefully introduced during the pilot studies). Mixon argues that this ambiguity inherent in the basic procedure weakens the methodological strength of the obedience studies because certainty regarding the infliction of harm would have caused subjects to disobey.

It can also be argued, however, that Milgram's deliberate creation of ambiguity was a necessary ingredient in the construction of the subjects' moral dilemma either to stop or continue inflicting shocks. To clarify, if a subject suspected that the learner was not being shocked and, for whatever reason, continued doing as they were told, their decision required them to take a major risk: their suspicion could be wrong and, if so, the learner would be seriously harmed. Inherent ambiguity in the procedure therefore created the possibility of the subject being wrong. We argue that this risk of being wrong is an essential methodological ingredient in the creation of the moral dilemma that Milgram tested. Would the subjects continue to place the learner's welfare at risk because they suspected he was not being harmed? Or would they choose the safe option and eliminate any possibility of being wrong by refusing to inflict further "shocks"? From this perspective, the ambiguity and uncertainty inherent in Milgram's basic experimental procedure is not a methodological weakness. On the contrary, the ambiguity and uncertainty Milgram purposefully injected into the basic procedure introduced the possibility that subjects could be wrong with potentially devastating consequences for the learner (Russell, 2014b: pp. 204-206; Russell, 2018a: pp. 123-126). If all subjects were absolutely certain that the learner was not being shocked, there would have been no chance of being wrong. And if there was no possibility of being wrong, then subjects would not have had to resolve a moral dilemma, which is what the experiment ultimately tested. Thus, the ambiguity inherent in Milgram's basic experimental procedure was not a methodological weakness, but instead was a necessary component in the construction of the moral dilemma that Milgram imposed on the subjects.

It's probable that Mixon would dispute that the subjects' dilemma had a moral dimension for the same reason outlined by Jerry Burger, a researcher who partially replicated the obedience studies over a decade ago (Burger, 2009). In an interview, Burger (cited in Perry, 2012: p. 359) said of the obedience studies:

When you're in that situation, wondering, should I continue or should I not, there are reasons to do both. What you do have is an expert in the room who knows all about this study, and presumably has been through this many times before with many participants, and he's telling you, [t]here's nothing wrong. The reasonable, rational thing to do is to listen to the guy who's the expert when you're not sure what to do.

Burger obviously senses great merit in the *trust-ambiguity-expertise nexus*. But because a subject's decision to trust an expert during an ambiguous situation necessitated they take a potentially devastating risk on an innocent person's well-being, we argue that doing as one was told was neither a reasonable nor a

rational solution to the problem at hand. The most reasonable and rational response when subjects were “not sure what to do” was instead to err on the side of caution and refuse to inflict the apparently more intense shocks (see Coutts, 1977: p. 520, cited in Darley, 1995: p. 133). In other words, they were invoking what is now widely known in scientific research on the effects of climate change as the *precautionary principle* (see O’riordan & Jordan, 1995). Doing so eliminated the risk of being wrong and protected the well-being of a fellow human being. Although not common, this cautious type of problem-solving was exhibited among a minority of subjects across numerous experimental conditions. For example, one actually suspicious and uncooperative 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.’” (SMP, Box 44, Divider (no label, #1106). Another subject in a different condition noted he “wasn’t sure he [the learner] was getting the shocks,” but when “he started to complain vigorously ... I refused to go on” (SMP, Box 44, Divider “8”, #1818). In fact, one subject *during* the *Relationship* condition explicitly stated he was certain his friend was not being shocked, but he still refused to trust the experimenter:

Teacher: “I don’t believe this! I mean, go ahead.”

Experimenter: “You don’t believe what?” [...]

Teacher: “I don’t believe you were giving him the shock.”

Experimenter: “Then why, why won’t you continue?”

Teacher: “Well I, I just don’t want to take a chance, I mean I, I”

Experimenter: “Well if you don’t believe that he’s getting the shocks, why don’t you just continue with the test and we’ll finish it?”

Teacher: “Well I, I can’t, because I can’t take that chance.” (SMP, Box 153, Audiotape #2439)

Clearly, this subject did not believe his friend was being shocked but because of the opaque partition, the fact is *he could not be certain*. The subject’s uncertainty in the ambiguous situation confronting him dictated that he could not afford to “take that chance” and trust the expert in charge because there was still a possibility his hunch might be wrong. And this subject was obviously well aware of the consequences that such a mistake would have for his friend. These kinds of responses all suggest that as far as these subjects knew, the experimenter could have been a rogue “mad scientist” (Perry, 2012: p. 135), someone not to be trusted. Mixon’s suggestion that obedient subjects put their trust in experts collapses when it is compared with how these disobedient subjects resolved the ambiguous situations confronting them.

Because in the *Relationship* condition the victim was at least an acquaintance and not, as was the case in every other variation, a stranger, this more cautious type of problem-solving was, as reflected in this variation’s low 15% completion rate, much more common. This is probably because subjects in this condition could not as easily afford to take the risk that the experiment was “probably” a

ruse, as they could when a stranger's well-being was on the line. Indeed, the *trust-ambiguity-expertise nexus* struggles to explain why nearly all subjects in this unique variation refused to complete the experiment. For example, the *Relationship* condition included just as much ambiguity as the *New Baseline* condition but rather mysteriously generated a 50% lower completion rate (15 versus 65%, respectively). So, despite the *Relationship* condition's inherent ambiguity, 85% of subjects refused to place their trust in the experimenter's "expert" status (with 80% stopping the experiment before the relatively low 195-volt switch; see SMP, Box 46, Folder 163, Titled: "Obedience Notebook 1961-1970"). Intriguingly, the closer the relationship within this variation, the more risk-averse subjects were: three subjects were related to their learner, with none completing (Russell, 2014b: p. 198).

Therefore, we not only dispute the purported importance of the *trust-ambiguity-expertise nexus*, we also believe it to be a methodological red herring. Although this conclusion naturally renders us critical of Orne, Holland, Mixon and all the contemporary scholars who have provided evidence in favour of the *trust-ambiguity-expertise nexus*, it should be remembered that it was Milgram's weak premise—that the dramaturgical credibility of the obedience studies was of paramount importance—that led these scholars down this distracting path.

If, as we argue, the *trust-ambiguity-expertise nexus* is of little or no importance, what then is? In light of the above argument, we believe the significant issue is whether or not subjects placed the learner's well-being at risk. When viewed from this perspective, the 35% of subjects who stopped 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 the experiment—whether fully deceived or suspicious—were willing to place the learner's well-being at risk (Russell, 2018a: p. 126). If this is a valid claim, it overcomes what has proven to be a major methodological sticking point that for many years has plagued the obedience studies: "how are we to tell an obedient subject who believes in the cover story from one who does not?" (Hoffman et al., 2015: p. 679). Our answer to this question is that it simply does not matter. The question is irrelevant.

Although we are critical of Milgram's flawed logic, we also believe the contemporary scholars—specifically Hollander & Turowetz, Perry et al., and Haslam & Reicher—have been too willing to accept as undeniably truthful the obedient subjects' post-experimental justifications for completing. In the following section we will explain why mostobedient subjects' post-experimental justifications should be treated with skepticism.

5.2. Milgram's Trap

When obedient subjects justified their actions during the post-experimental interviews and surveys, obviously the contemporary scholars believed their statements to be truthful. When subjects said they completed to help advance scien-

tific knowledge, Haslam & Reicher believed them. When subjects said, more commonly, they completed because they thought the experiment was fake, Hollander & Turowetz and Perry et al. believed them. As mentioned, [Hollander & Turowetz \(2017: pp. 659-660\)](#) also criticized the present authors' earlier research for our alleged incredulity. According to them, unlike the Nazi perpetrators after World War Two, the obedient subjects had no or little reason to lie. Thus, Hollander & Turowetz believe us to have been too skeptical and suggest we should have been more willing to accept most subjects' justifications for completing. We disagree and maintain our original position. That is, we suspect most obedient subjects were aware quite early on in the experiment that they were engaging in wrongdoing. However, when the actual purpose of the research was revealed to them during the post-experimental interviews, they then tried to save face by providing inaccurate justifications for completing. Our incredulity surrounding the reliability of most obedient subjects' justifications traces back to a design feature within Milgram's basic experimental procedure—his cunning psychological trap—which we believe lured many of them to lie about their reason/s for completing. Because Hollander & Turowetz (and Perry et al., and Haslam & Reicher for that matter) pay little or no attention to the importance of this trap in adulterating most obedient subjects' post-experimental justifications, the following section—which is rather more conceptual than strictly empirical—will present it.

When Milgram was a Harvard graduate student during the 1956-57 academic year, he was visiting scholar Solomon Asch's research and teaching assistant. Asch subsequently became Milgram's strongest intellectual influence ([Tavris, 1974: p. 77](#)). In his student notes circa 1956-59, Milgram reveals why he believed Asch's Group Pressure Conformity study "is a Great Experiment":

[the subject] beli[e]ves that the conflict with him is a purely private issue which concerns no one but him, and of which all others are totally unaware. He dares not betray his secret, yet by his actions, he is betrayed. The yielding subject makes frantic efforts to conceal his conflict, yet by these efforts is the conflict betrayed. (As cited in [Russell, 2018a: p. 41](#))

In 1960, when Milgram was honing his basic obedience study procedure, he incorporated this feature of Asch's research into his own project. To best ensure that subjects at the beginning of the experiment started inflicting initially light shocks on the learner, the *New Baseline* procedure bombarded subjects with an array of subtle techniques of coercion. [Erdoes \(2013: p. 123\)](#) aptly terms this stage of the experiment the "persuasion phase" (see also [Gibson, 2013; Russell, 2018a: pp. 199-238](#)). One of many such examples of these coercive techniques surrounded the real 45-volt test shock all subjects received in order to bolster the verisimilitude of the shock machine. Arguably, this light test shock may have encouraged most subjects to perceive the first three shocks (15 to 45-volts) they were soon afterwards asked to inflict on the learner as fairly harmless. Thus, very subtly the test shock increased the probability of subjects starting an experiment

that required them to ostensibly inflict intensifying shocks on an innocent person with a mild heart condition.

When the subjects inflicted these initial shocks, the learner grunted in response. However, the next few shocks elicited a slightly more pained reaction from the learner. For example, at the 75 and 90-volt switches, the learner yelps “Ugh!” (Milgram, 1974: p. 56). If subjects then showed any hesitancy, the experimenter released the first two prods: “Please continue,” and “The experiment requires that you continue.” All subjects at this early stage of the *New Baseline*-condition inflicted the first six shocks (15 - 90 volts). Doing so, however, meant all subjects met the most important criterion of Freedman & Fraser’s (1966) foot-in-the-door phenomenon: the fulfillment of one or several small requests which, unbeknown to them, were about to be followed by some far more demanding ones. In relation to Milgram’s experiments, Gilbert (1981: p. 692) explains that this manipulative technique had:

two important consequences: 1) it engages subjects in committing precedent-setting acts of obedience before they realize the “momentum” which the situation is capable of creating, and the “ugly direction” in which that momentum is driving them; and 2) it erects and reinforces the impression that quitting at any particular level of shock is unjustified (since consecutive shock levels differ only slightly and quantitatively).

Subjects could not foresee the “ugly direction” in which they were heading. Milgram (1965a: p. 73) was not exaggerating when he described the experiment as akin to being “thrown into a swift-flowing stream [...] The individual, upon entering the laboratory, becomes integrated into a situation that carries its own momentum.” All subjects up to this point found it easier to swim along with, rather than to resist, Milgram’s metaphorical current.

After completing this section of the experiment, the learner’s reactions to subsequent shocks intensified further:

105 volts Ugh! (louder)
120 volts Ugh! Hey, *this* really hurts.
135 volts Ugh!! (Milgram, 1974: p. 56)

If the first two prods failed to motivate the subject the experimenter issued (not necessarily for the first time) the more forceful third and fourth prods: “It is absolutely essential that you continue,” and “You have no other choice, you *must* go on.”

Because the experimenter was pushing them into inflicting these shocks (Perry, 2012: pp. 134-135), some subjects around this point in the experiment may have felt that they were being bullied by the experimenter and may have convinced themselves that only the experimenter—not themselves—was responsible for harming the learner. And, if a subject genuinely did not *feel* responsible for the learner’s pain, then they were also unlikely to experience the guilt that normally accompanies such responsibility. The first reaction for many of these

subjects was to neutralize the greatest source of *their* pain—the experimenter’s stress-inducing prods—by inflicting *more* shocks. But doing so only drew them a little deeper into Milgram’s trap.

As the intensity of the learner’s reactions to the first nine shocks intensified, the harder it became for subjects to continue viewing themselves as bullied victims. Because the subject flicked the switches, they were clearly most responsible for the learner’s “pain.” Nevertheless, subjects were also aware that they would never have inflicted the shocks of their own accord (as the very low completion rate from the previously mentioned *Subject Chooses Shock Level* Condition suggests). So, in the eyes of subjects, the experimenter must have been solely or mostly responsible. The subjects’ mixed feelings surrounding the issue of responsibility can be traced back to the experimenter’s prods: he said it was “essential” subjects continue inflicting shocks and that they had “no choice.” Unbeknown to the subject, right from the start of the experiment the prods were designed to instill in subjects’ minds the perception that the experimenter did not seem to believe they could be held personally responsible for hurting the learner.

Around this early stage in the experiment, subjects probably no longer perceived themselves as victims, but neither would they have viewed themselves as perpetrators. Such a subtle change in self-perception could have stimulated a shift in their perceptions of responsibility from none to some. But as long as these subjects regarded the experimenter as *mostly* responsible, many could continue a little further up the shock board because, not feeling most responsible, they unlikely believed that it was up to them to stop the experiment.

It is likely, however, that these gradually shifting perceptions of personal responsibility changed significantly after subjects inflicted—as nearly 98% of them did—the 150-volt shock. Suddenly the panicked learner yelled: “Ugh!!! Experimenter! That’s all. Get me out of here. I told you I had heart trouble. My heart’s starting to bother me now. [...] I refuse to go on. Let me out” (Milgram, 1974: p. 56). For many subjects, the learner’s intensified reaction and the sudden mention of heart difficulties made the 150-volt switch a pivotal juncture in the procedure (see Burger, 2009; Packer, 2008). From this point subjects could no longer construe the infliction of shocks as “probably” harmless. The close relationship between cause (flicking switches) and effect (the learner’s heightened “shock” reactions) ensured that the majority of subjects started to *feel* most responsible for what was happening. Subjects could no longer view themselves as innocent victims, but instead as potential perpetrators implicated in the pursuit of wrongdoing. Indeed, many subjects probably suspected that the learner was not being shocked (surely Yale would never have allowed the infliction of real shocks!). But this suspicion implies an awareness of an alternative possibility: the shocks might have been real.

Consequently, around the 150-volt switch and at each subsequent switch thereafter, subjects faced Milgram’s previously mentioned moral dilemma.

Should they help the experimenter collect a full data set, *or* was it more important to stop inflicting potentially dangerous electric shocks on an innocent person complaining of heart trouble? If they chose to end the experiment, the learner's apparent "pain" would cease. However, the latter option required them to do something they greatly feared: they would have to instigate an uncomfortable, impolite and socially awkward confrontation with a predictably frustrated, perhaps irate, Yale scientist³. A simultaneous means of meeting the demands of both the experimenter and the learner was not apparent. To viewers of [Milgram's \(1965b\)](#) video documentary, the solution to the dilemma seems obvious: stand up to the experimenter and stop the experiment. This solution, however, was rare: 33 out of 40 subjects went on and inflicted the 165-volt shock.

So desirous of avoiding a direct confrontation with the experimenter, many obedient subjects sought to devise a non-confrontational, polite and thus inoffensive means to end the experiment, one capable of pleasing both the experimenter and the learner ([Gibson, 2013: pp. 295-298](#); [Russell, 2018a: pp. 215-221](#)). One common strategy involved subjects pursuing covert acts that aimed at sabotaging the entire experimental trial. As [Milgram \(1974: p. 159\)](#) notes,

Some subjects could be observed signaling the correct answer to the victim by stressing it vocally as they read the multiple-choice words aloud. That is, they attempted to prompt the learner and thus prevent his receiving shocks. These subjects are willing to undermine the experiment but not to cause an open break with authority.

Similarly, other subjects tried to deceive the experimenter by failing to fully depress the shock switch so as to not "punish" the learner (see [Perry, 2012: p. 197](#)). These varied attempts at sabotage were so common that Milgram had to instruct the experimenter to ensure subjects performed their tasks correctly (see [Russell, 2009: pp. 152-153](#)). Sabotage was attractive to subjects because it enabled them to (seemingly) meet simultaneously both the experimenter's and the learner's desires.

Other confrontation-fearing subjects were not as creative (deceptive?) as the saboteurs and, according to [Akerlof \(1991: p. 9\)](#), instead procrastinated with a "plan to disobey in the future." That is, these subjects hoped that procrastination might provide them with sufficient time needed to devise an effective non-confrontational means of ending the experiment (see [Rochat](#) cited in [Perry, 2012: p. 380](#))⁴, preferably one capable of pleasing both the experimenter and the learner. However, unbeknown to these subjects, during the pilot studies Milgram had already encountered test subjects' most common non-confrontational exit strategies, thus enabling him to anticipate what most subjects were likely to say during the official experiments ([Russell, 2011; 2018a: p. 217](#)). Consequently,

³As [Miller et al. \(1995: p. 9\)](#) argued, concerns about "being 'impolite' ... would seem absurd. However, in the actual context of the situation, these concerns are influential."

⁴As Milgram argued: "...there is a continual effort on the part of some subjects to 'break out' of the role assigned to them by the experimenter" (SMP, Box 46, Folder 168).

Milgram was able to arm his experimenter with various prods specifically designed to counter the most common non-confrontational attempts at extrication. For example, as the learner's screams intensified, many subjects enquired about the effect of the shocks, presumably sensing it would be reasonable to stop if, as the situation clearly indicated, the learner was being harmed. Subjects were often surprised when the experimenter responded: "Although the shocks may be painful, there is no permanent tissue damage, so please go on". Again, in the hope that further procrastination might help them to devise a more successful non-confrontational exit strategy, many subjects tentatively continued up the shock board. Since subjects were the only ones inflicting "shocks," some requested that the experimenter explicitly establish the direct lines of responsibility for what was clearly *his* experiment. But this "tactic," which politely aimed "to bring the authority to account" (Lunt, 2009: p. 18), also failed, with the experimenter responding: "I'm responsible for anything that happens to him. Continue, please". This statement explicitly confirmed the obedient subjects' emerging suspicion: the experimenter alone believes himself to be responsible for the subjects' actions. Thus, despite most *New Baseline* subjects feeling from about the 150-volt switch onwards that they were engaged in wrongdoing, Milgram's carefully designed prods were incrementally luring them into suspecting that they may not have *appeared* to be most responsible for the learner's pain.

Typically, the subjects' search for a non-confrontational exit strategy backfired, by drawing them into inflicting even more shocks, which *they knew* they should not have done. Even if they had the courage to confront the experimenter, they would also be faced with the burden of explaining their failure to stop upon earlier realising it was wrong to have continued. The easiest, albeit self-interested, option that remained was to relent to the experimenter's pressure and unenthusiastically accept his tempting offer: continue doing what the experimenter wanted, and if later questioned about their immoral decision, then simply blame him for their actions. When subjects contemplated this option, they began to enter into a Faustian bargain with the Mephistophelian experimenter.

Given this opportunity, all a confrontation-fearing subject needed to do was to pretend that they genuinely believed what the experimenter was telling them: it was (apparently) "essential" to "continue," that they had "no choice," that the shocks were "harmless," and that only the experimenter was "responsible." For obedient subjects, doing so came with numerous self-interested advantages: it enabled them to avoid a feared confrontation, (seemingly) absolved them from moral and legal culpability for continuing, and cemented the belief that they could probably inflict more shocks with total impunity. They could continue inflicting further shocks safe in the belief that they could evade blame because they were (apparently) *just following orders*. And if, at any future point, such subjects did not think that blaming the experimenter for their actions would satisfy any interrogators, they could simply blame anyone or anything else as more respon-

sible than themselves⁵. If all else failed, the subject could always argue they were sure the shocks were not real, so they (apparently) felt free to submit to the experimenter's will. And crucially, *only they would know about their making of this decision* (or so Milgram led them to believe). As Milgram said earlier about why he thought Asch's experiment was "[g]reat": the subject believes "that the conflict with him is a purely private issue which concerns no one but him, and of which all others are totally unaware. He dares not betray his secret, yet by his actions, he is betrayed. The yielding subject makes frantic efforts to conceal his conflict, yet by these efforts is the conflict betrayed."

One such action that betrayed the obedient subjects' knowledge that they were engaging in wrongdoing was their common requests for authorization. For example, before administering the 345-volt shock one subject "points to the label, DANGER on the shock generator." After inflicting this shock and "no sound is heard from the learner, she queries, 'Do I go right to the end, sir?'" (Milgram, 1974: p. 77; see also pp. 49, 51, 53, 73-76, 80, 84, 86). As Milgram (1974: p. 187) noted: "repeated requests for authorization are always an early sign that the subordinate senses, at some level, that the transgression of a moral rule is involved." Damico (1982: pp. 424-425) agrees:

... most revealing ... in the Milgram experiment is not the inability of his subjects to understand the difference between right and wrong—anxiety was often the most visible emotion—but their failure to care about the difference in a way that would have made it the controlling factor in their behavior.

However, Damico in the early 1980s was unaware of the then unpublished *Relationship* condition—the only variation where the victim was not a stranger—which tested the limits of the obedient subjects' "failure to care ...". Probably because in this condition subjects were much more likely to be concerned about the implications their actions might have on the learner, 85% flatly refused to do as they were told. Furthermore, and as mentioned, the closer the relationship, the lower the completion rate—all three subjects related to their learner refused to complete. Perhaps unsurprisingly, these particular findings show that the well-being of strangers matters much less than the well-being of emotionally close individuals. This assertion was confirmed by one of the three *Relationship* condition subjects, who explained why he refused to inflict intense shocks on his relative:

Subject: But since he was my brother-in-law I stopped...

Milgram: Why do you think you stopped for a brother-in-law?

Subject: Well ahhh... Why should I keep on going? It's not that necessary to keep on going right? That's the reason why I stopped.

What would he have done had the learner been a stranger?

⁵For example, one subject in the post-experimental interview blamed his military training for his having completed the experiment (Russell, 2018a: p. 224).

Subject: Well ... they told me I should keep on going, I keep on going.

Milgram: Why? What's the difference?

Subject: Well ... [inaudible] ... is not dangerous, nothing will happen to me.

Milgram: But that's what you were told ... with your brother-in-law.

Subject: Yeah, but there's a difference.

Milgram: What's the difference?

Subject: If it is a stranger I don't listen. Right? We are doing an experiment ... They told me to do it, I keep on doing it. He told me keep on going, I keep on going (Russell & Gregory, 2011: pp. 501-502). Several other *Relationship* condition subjects made the same point (Russell & Gregory, 2011: p. 515).

Nonetheless, if a subject in the *New Baseline* decided to accept the experimenter's tempting offer to shoulder full responsibility for their shock-inflicting actions, they thereafter moved from Erdos' (2013: p. 123) "persuasion phase" to his "after capitulation phase", where they stop half-heartedly resisting the experimenter's demands and instead commit to his desire that every shock be inflicted. But unbeknown to these subjects, the experimenter's offer was a trap: the experiments were actually about whether or not they would prioritize their lesser important desire to avoid a confrontation (by following potentially harmful orders) over bravely standing up for and protecting a fellow person's well-being. Milgram was simply testing if subjects would take or reject the immoral bait, and as the incrementally manipulative *New Baseline* procedure illustrates, most—albeit unenthusiastically—complied. In Milgram's own words: "the experimental set up relies ... on *seduction*, the systematic ensarement [*sic*] of the subject into a web of obligation and uncritically from which he is unable to escape" (as cited in Russell & Gregory, 2011: p. 508). We therefore argue that Milgram's obedience experiments are about *how most ordinary people can be tempted into resolving a moral dilemma when they are simultaneously led to suspect they will personally benefit and can probably act with impunity*⁶.

After inflicting every shock, the actual rationale for the experiment was revealed to the subjects: would they follow or reject the experimenter's orders? Those who "obeyed", in an attempt to save face—*pace* Hollander & Turowetz, Perry et al., and Haslam & Reicher—began their post-experimental prevarication, obfuscation and outright lying. Consider the veritable cover-up pursued by the pseudonymous Elinor Rosenblum, a self-styled pillar of the community who did volunteer work with "dropouts" and "leather-jacket guys" (Milgram, 1974: p. 81). Not long into the experiment she encountered the moral dilemma, later admitting, "I was tempted so much to stop and to say: 'Look I'm not going to do it anymore. Sorry. I'm not going to do it'" (Milgram, 1974: p. 83). But as her repeated infliction of shocks indicated, her thoughts were never translated into ac-

⁶With regard to the above account detailing why we think most subjects completed the obedience studies, it must be acknowledged that some likely behaved in similar or identical ways but for different reasons.

tion. After inflicting the 270-volt switch she began to shake uncontrollably, stating “Must I go on? Oh, I’m worried about him ... Can’t we stop? I’m shaking. I’m shaking” (Milgram, 1974: p. 80). In the end, Rosenblum went on to inflict every shock asked of her. During the debrief Rosenblum was awkwardly reintroduced to her unharmed victim and said:

You’re an actor, boy. You’re marvelous! Oh, my God, what he [the experimenter] did to me. I’m exhausted. I didn’t want to go on with it. You don’t know what I went through here. A person like me hurting you, my God. I didn’t want to do it to you. Forgive me, please. I can’t get over this. My face is beet red. I wouldn’t hurt a fly. (Milgram, 1974: pp. 82-83)

Although Rosenblum clearly did not “want to do it” and was repeatedly tempted to challenge the experimenter, the most she was willing to do for the learner was to surreptitiously try and sabotage the experiment. As she said to the learner/actor: “Did you hear me *stressing* the [correct] word[?] I was hoping that you would hear me” (Milgram, 1974: p. 82). This admission of sabotage, however, only invalidated her post-experimental justification for completing: “It is an experiment. [...] So I had to do it. You said so” (1974: p. 83). If she were, as Haslam & Reicher would argue, just following orders for the betterment of science, why did she try to sabotage the very purpose she (apparently) so identified with?⁷ On being told the learner never actually received any shocks, Rosenblum exclaimed: “You’re kidding! He didn’t get what I got. (She squeals) I can’t believe this” (Milgram, 1974: p. 82). Despite her “beet red” face and squeal of surprise, in a subsequent questionnaire, Rosenblum claimed—this time in support of Hollander & Turowetz and Perry et al.—that apparently her “‘mature and well-educated brain’ had not believed the learner was getting shocks” (Milgram, 1974: p. 84).⁸ Lauren Slater (2004: p. 40) perhaps said it best: “The power of Milgram’s experiments lies, perhaps... in the great gap between what we think about ourselves, and who we frankly are.” We thus believe Rosenblum’s inconsistent and contradictory response is a classic example of a subject who—like many others—fell into Milgram’s trap.

Rosenblum’s pre-experimental altruistic self-perception was overwhelmingly positive (“I’m unusual; I’m softhearted”; Milgram, 1974: p. 83), and her self centered actions during the obedience studies made the post-experimental reality—perhaps I’m not as caring as I think I am?—too bitter a pill to swallow. (Self-?) deception offered Elinor Rosenblum a pathway of least emotional resistance.

Unlike her, other obedient subjects were later more honest with themselves. As one subject admitted, “I thought the ‘shocks’ might harm the other ‘subject’ however, I mentally ‘passed the buck’ feeling the one running the experiment would take all responsibility” (as cited in Russell & Gregory, 2011: p. 508). Some were also self-reflexively willing to acknowledge the disconcerting gap between

⁷These common attempts to sabotage the experiment would seem to undermine the validity of Haslam & Reicher’s *Engaged Followership* theory as a significant explanation of the obedience studies.

their positive self-image and that of their behavior during the experiment:

The thought also occurred to me that for a supposedly highly civilized and, in my mind, “soft-hearted” person I had carried the experiment on a lot longer than I should have were I as “soft hearted” as I had led myself to believe. (I try to make myself believe that it was because I had agreed to complete the experiment but without much success). It makes me wonder if I would be a real “resistance fighter” in the event that our country should ever find itself in the position of a France or a Denmark under occupation. (SMP, Box 44, Divider “9”, #0202)

Because Milgram had exposed (and recorded) the selfishness inherent in their moral choices, few subjects were willing to make such humbling admissions. In conflict with Hollander & Turowetz’s (2017: p. 659) interpretation that after the experiment obedient subjects were left “sincerely struggling... to make sense of what had happened”, we instead argue that Milgram’s trap ensured most knew they were engaged in wrongdoing, and also were likely to have been highly motivated to lie about this awareness afterwards. We question the reliability of the contemporary research like Hollander & Turowetz, Perry et al. and Haslam & Reicher, all of whom have uncritically accepted as accurate obedient subjects’ justifications for completing the experiment. Of course, if Milgram’s obedient subjects had reason to lie after the experiment, then—pace Hollander & Turowetz—it can be argued they behaved in the same way that so many Nazi war criminals did after Hitler’s downfall.

6. Conclusion

Milgram believed that deceiving most of his subjects into believing they were inflicting real shocks on the learner was of critical importance to the internal validity of the obedience studies. Recently published work argues that most obedient subjects completed the experiment because they did not believe that the learner was being harmed (Perry et al., 2020; Hollander & Turowetz, 2017). Perry et al. claim that if it were true that Milgram failed to deceive most obedient subjects into believing they were inflicting real shocks, the *trust ambiguity-expertise nexus* undermines the internal validity of the obedience studies.

However, we do not agree that Milgram’s belief that deceiving most subject into believing the learner was being shocked was of critical importance to the internal validity of his experiments. Instead, we argue that what was crucial to Milgram’s research paradigm was that he succeeded in ensuring his subjects could not be certain if the experiments were fake. So, in conflict with Milgram’s logic, we argue that it actually does not matter that some, many, or even all the obedient subjects strongly suspected the experiment was fake. What matters is that most subjects completed the experiment despite there being a real risk of them being wrong. And here the *New Baseline* results are clear: 65% of the *New Baseline* subjects—whether fully deceived or suspicious—placed the learner’s

well-being at risk by choosing to inflict every shock. Thus, in conflict with the *trust ambiguity-expertise nexus* (along with any research that supports it), it still matters a great deal if most subjects chose to inflict every shock. If our above claim is valid, it overcomes and renders redundant what has proven to be a major methodological sticking point that, for over half a century, has plagued Milgram's obedience studies: "how are we to tell an obedient subject who believes in the cover story from one who does not?" (Hoffman et al., 2015: p. 679). In our views, the answer to this question simply does not matter.

Nevertheless, it is important to note that other methodological criticisms remain. Perhaps the best such example is the criticism that Milgram failed to fully standardize his score of baseline variations (see Gibson, 2013: pp. 298-299; Perry, 2012: pp. 134-135). More specifically, Perry and Gibson have shown that after Milgram ran his first few variations, the experimenter stopped following his preconceived actor's script and began improvising his own prods (presumably with the intention of ensuring that subjects continued inflicting shocks). This, we believe, is a valid methodological criticism of Milgram's obedience studies⁸. Yet we do not believe it to be particularly important criticism, because many of the independent replications of the original studies which strictly followed Milgram's technically inaccurate published procedural instructions still obtained high completion rates (see Blass, 2012).

Finally, in conflict with some contemporary scholars, we believe it would be unwise to throw out Milgram's obedience study baby with the bathwater of methodological imperfection. There remains much to learn from Milgram's research, and because his obedience studies remain, for the most part, internally valid, current attempts to externalise aspects of his findings to real-life settings—like the Holocaust and even climate catastrophe—remain eminently plausible (see, for example, Russell, 2018a, 2018b; Russell & Bolton, 2019, the latter two of which are Open Access publications).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Akerlof, G. (1991). Procrastination and Obedience. *The American Economic Review*, 81, 1-19.
- Baumrind, D. (1964). Some Thoughts on Ethics of Research: After Reading Milgram's "Behavioral Study of Obedience". *American Psychologist*, 19, 421-423.
<https://doi.org/10.1037/h0040128>

⁸Milgram's unscientific failure to stop his experimenter from innovating is, in part, why the first author concluded in his book that "Milgram behaved less like a social scientist and more like a goal-orientated project manager trying to socially engineer a preconceived result" (Russell, 2018b: p. 280). Because the same thing can be said of the most destructive Nazi bureaucrats, this conclusion provided Russell (2018b) with the foundation on which to build his Milgram-Holocaust linkage.

- Baumrind, D. (1985). Research Using Intentional Deception: Ethical Issues Revisited. *American Psychologist*, 40, 165-174. <https://doi.org/10.1037/0003-066X.40.2.165>
- Baumrind, D. (2013). Is Milgram's Deceptive Research Ethically Acceptable? *Theoretical & Applied Ethics*, 2, 1-18.
- Baumrind, D. (2015). When Subjects Become Objects: The Lies behind the Milgram Legend. *Theory and Psychology*, 25, 690-696. <https://doi.org/10.1177/0959354315592062>
- Blass, T. (2012). A Cross-Cultural Comparison of Studies of Obedience Using the Milgram Paradigm: A Review. *Social and Personality Psychology Compass*, 6, 196-205. <https://doi.org/10.1111/j.1751-9004.2011.00417.x>
- Brannigan, A. (1997). The Postmodern Experiment: Science and Ontology in Experimental Social Psychology. *The British Journal of Sociology*, 48, 594-610. <https://doi.org/10.2307/591598>
- Brannigan, A. (2013). *Beyond the Banality of Evil: Criminology and Genocide*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199674626.001.0001>
- Brannigan, A. (2020). *The Use and Misuse of the Experimental Method in Social Psychology: A Critical Examination of Classical Research*. London: Routledge. <https://doi.org/10.4324/9781003034803>
- Brannigan, A., Nicholson, I., & Cherry F. (2015). Introduction to the Special Issue: Unplugging the Milgram Machine. *Theory & Psychology*, 25, 551-563. <https://doi.org/10.1177/0959354315604408>
- Burger, J. M. (2009). Replicating Milgram: Would People Still Obey Today? *American Psychologist*, 64, 1-11. <https://doi.org/10.1037/a0010932>
- Damico, A. J. (1982). The Sociology of Justice: Kohlberg and Milgram. *Political Theory*, 10, 409-433. <https://doi.org/10.1177/0090591782010003005>
- Darley, J. M. (1995). Constructive and Destructive Obedience: A Taxonomy of Principal-Agent Relationships. *Journal of Social Issues*, 51, 125-154. <https://doi.org/10.1111/j.1540-4560.1995.tb01338.x>
- De Swaan, A. (2015). *The Killing Compartments: The Mentality of Mass Murder*. New Haven, CT: Yale University Press.
- Eckman, B. K. (1977). Stanley Milgram's "Obedience" Studies. *Et Cetera*, 34, 88-99.
- Erdos, E. (2013). The Milgram Trap. *Theoretical & Applied Ethics*, 2, 123-142.
- Gibson, S. (2013). Milgram's Obedience Experiments: A Rhetorical Analysis. *British Journal of Social Psychology*, 52, 290-309. <https://doi.org/10.1111/j.2044-8309.2011.02070.x>
- Gilbert, S. J. (1981). Another Look at the Milgram Obedience Studies: The Role of the Graded Series of Shocks. *Personality and Social Psychology Bulletin*, 7, 690-695. <https://doi.org/10.1177/014616728174028>
- Griggs, R. A. & Whitehead, G. I. (2015). Coverage of Milgram's Obedience Experiments in Social Psychology Textbooks: Where Have All the Criticisms Gone? *Teaching of Psychology*, 42, 315-322. <https://doi.org/10.1177/0098628315603065>
- Harré, R. (1979). *Social Being: A Theory for Social Psychology*. Oxford, UK: Basil Blackwell.
- Haslam, S. A. & Reicher, S. D. (2012). Tyranny: Revisiting Zimbardo's Stanford Prison Experiment. In J. R. Smith & S. A. Haslam (Eds.), *Social Psychology: Revisiting the Classic Studies* (pp. 126-141). London: Sage.
- Haslam, S. A., Reicher, S. D., & Birney, M. E. (2014). Nothing by Mere Authority: Evidence That in an Experimental Analogue of the Milgram Paradigm Participants Are Motivated Not by Orders but by Appeals to Science. *Journal of Social Issues*, 70, 473-488.

- <https://doi.org/10.1111/josi.12072>
- Haslam, S. A., Reicher, S. D., Millard, K., & McDonald, R. (2015). "Happy to Have Been of Service": The Yale Archive as a Window into the Engaged Followership of Participants in Milgram's "Obedience" Experiments. *British Journal of Social Psychology*, 54, 55-83. <https://doi.org/10.1111/bjso.12074>
- Hoffman, E., Myerberg, N. R., & Morawski, J. G. (2015). Acting Otherwise: Resistance, Agency, and Subjectivities in Milgram's Studies of Obedience. *Theory & Psychology*, 25, 670-689. <https://doi.org/10.1177/0959354315608705>
- Hollander, M. M. & Turowetz, J. (2017). Normalizing Trust: Participants' Immediately Post-Hoc Explanations of Behaviour in Milgram's "Obedience" Experiments. *British Journal of Social Psychology*, 56, 655-674. <https://doi.org/10.1111/bjso.12206>
- Kaposi, D. (2017). The Resistance Experiments: Morality, Authority and Obedience in Stanley Milgram's Account. *Journal for the Theory of Social Behaviour*, 47, 382-401. <https://doi.org/10.1111/jtsb.12137>
- Le Texier, T. (2019). Debunking the Stanford Prison Experiment. *American Psychologist*, 74, 823-839. <https://doi.org/10.1037/amp0000401>
- Lunt, P. (2009). *Stanley Milgram: Understanding Obedience and Its Implications*. Basingstoke: Macmillan International Higher Education.
- Milgram, S. (1963). Behavioral Study of Obedience. *Journal of Abnormal and Social Psychology*, 67, 371-378. <https://doi.org/10.1037/h0040525>
- Milgram, S. (1965a). Some Conditions of Obedience and Disobedience to Authority. *Human Relations*, 18, 57-76. <https://doi.org/10.1177/001872676501800105>
- Milgram, S. (1965b). *Obedience (a Filmed Experiment)*. Distributed by New York University Film Library.
- Milgram, S. (1972). Interpreting Obedience: Error and Evidence. A Reply to Orne and Holland. In A. G. Miller (Eds.), *The Social Psychology of Psychological Research* (pp. 138-154). New York: Free Press.
- Milgram, S. (1974). *Obedience to Authority: An Experimental View*. New York: Harper & Row.
- Miller, A. G. (1986). *The Obedience Experiments: A Case Study of Controversy in Social Science*. New York: Praeger.
- Miller, A. G., Collins, B. E., & Brief, D. E. (1995). Perspectives on Obedience to Authority: The Legacy of the Milgram Experiments. *Journal of Social Issues*, 51, 1-19. <https://doi.org/10.1111/j.1540-4560.1995.tb01331.x>
- Mixon, D. (1972). Instead of Deception. *Journal of the Theory of Social Behavior*, 2, 145-178. <https://doi.org/10.1111/j.1468-5914.1972.tb00309.x>
- Mixon, D. (1976). Studying Feignable Behavior. *Representative Research in Social Psychology*, 7, 89-104.
- Mixon, D. (1989). *Obedience and Civilization: Authorized Crime and the Normality of Evil*. London: Pluto Press.
- Nicholson, I. (2011). "Torture at Yale": Experimental Subjects, Laboratory Torment and the "Rehabilitation" of Milgram's "Obedience to Authority. *Theory & Psychology*, 21, 737-761. <https://doi.org/10.1177/0959354311420199>
- Nussbaum, M. C. (2007). Texts for Torturers: From Stanford to Abu Ghraib—What Turns Ordinary People into Oppressors? *The Times Literary Supplement*. http://entertainment.timesonline.co.uk/tol/arts_and_entertainment/the_tls/article2677344.ece

- O’riordan, T. & Jordan, A. (1995). The Precautionary Principle in Contemporary Environmental Politics. *Environmental Values*, 4, 191-212.
<https://doi.org/10.3197/096327195776679475>
- Orne, M. T. & Holland, C. H. (1968). On the Ecological Validity of Laboratory Deceptions. *International Journal of Psychiatry*, 6, 282-293.
- Orne, M. T. (1962). On the Social Psychology of the Psychology Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, 17, 776-783. <https://doi.org/10.1037/h0043424>
- Packer, D. J. (2008). Identifying Systematic Disobedience in Milgram’s Obedience Experiments: A Meta-Analytic Review. *Perspectives on Psychological Science*, 3, 301-304. <https://doi.org/10.1111/j.1745-6924.2008.00080.x>
- Perry, G. (2012). *Beyond the Shock Machine: The Untold Story of the Milgram Obedience Experiments*. Melbourne: Scribe.
- Perry, G. (2015). Seeing Is Believing: The Role of the Film Obedience in Shaping Perceptions of Milgram’s Obedience to Authority Experiments. *Theory & Psychology*, 25, 622-638. <https://doi.org/10.1177/0959354315604235>
- Perry, G. (2018). *The Lost Boys: Inside Muzafer Sherif’s Robbers Cave Experiment*. London: Scribe Books.
- Perry, G., Brannigan, A., Wanner, R. A., & Henderikus S. (2020). Credibility and Incredulity in Milgram’s Obedience Experiments: A Reanalysis of an Unpublished Test. *Social Psychology Quarterly*, 83, 88-106. <https://doi.org/10.1177/0190272519861952>
- Reicher, S. D., Haslam, S. A., & Smith, J. R. (2012). Working toward the Experimenter: Reconceptualizing Obedience within the Milgram Paradigm as Identification-Based Followership. *Perspectives on Psychological Science*, 7, 315-324. <https://doi.org/10.1177/1745691612448482>
- Rochat, F. & Modigliani, A. (1997). Authority: Obedience, Defiance, and Identification in Experimental and Historical Contexts. In M. Gold (Ed.), *A New Outline of Social Psychology* (pp. 235-246). Washington DC: American Psychological Association. <https://doi.org/10.1037/10225-013>
- Russell, N. & Bolton, A. (2019). Climate Catastrophe and Stanley Milgram’s Electric Shock “Obedience” Experiments: An Uncanny Analogy. *Social Sciences*, 8, 178-204. <https://doi.org/10.3390/socsci8060178>
- Russell, N. & Gregory, R. J. (2011). Spinning an Organizational “Web of Obligation”? Moral Choice in Stanley Milgram’s “Obedience” Experiments. *The American Review of Public Administration*, 41, 495-518. <https://doi.org/10.1177/0275074010384129>
- Russell, N. & Gregory, R. J. (2015). The Milgram-Holocaust Linkage: Challenging the Present Consensus. *State Crime Journal*, 4, 128-153. <https://doi.org/10.13169/statecrime.4.2.0128>
- Russell, N. (2009). *Stanley Milgram’s Obedience to Authority Experiments: Towards an Understanding of Their Relevance in Explaining Aspects of the Nazi Holocaust*. Ph.D. Thesis, Wellington: Victoria University of Wellington.
- Russell, N. (2011). Milgram’s Obedience to Authority Experiments: Origins and Early Evolution. *British Journal of Social Psychology*, 50, 140-162. <https://doi.org/10.1348/014466610X492205>
- Russell, N. (2014a). The Emergence of Milgram’s Bureaucratic Machine. *Journal of Social Issues*, 70, 409-423. <https://doi.org/10.1111/josi.12068>
- Russell, N. (2014b). Stanley Milgram’s Obedience to Authority “Relationship” Condition: Some Methodological and Theoretical Implications. *Social Sciences*, 3, 194-214.

<https://doi.org/10.3390/socsci3020194>

Russell, N. (2018a). *Understanding Willing Participants: Milgram's Obedience Experiments and the Holocaust, Volume 1*. Cham: Palgrave Macmillan.

<https://doi.org/10.1007/978-3-319-95816-3>

Russell, N. (2018b). *Understanding Willing Participants: Milgram's Obedience Experiments and the Holocaust, Volume 2*. Cham: Palgrave Macmillan.

<https://doi.org/10.1007/978-3-319-97999-1>

Slater, L. (2004). *Opening Skinner's Box: Great Psychological Experiments of the Twentieth Century*. New York: W. W. Norton.

Tavris, C. (1974). A Sketch of Stanley Milgram: A Man of 1,000 Ideas. *Psychology Today*, 8, 74-75.