1 The Obedience Experiments

Milgram conducted his obedience experiments from 1961–1962 at Yale University, and published his first academic paper reporting his findings in 1963. The reaction was almost instantaneous, with Baumrind's (1964) critique and Milgram's (1964b) response setting the tone for decades of debate and research. As Kaposi (2017) has suggested, the reaction to the obedience experiments can be loosely divided into two 'waves'. A first wave of reaction involved important ethical, methodological and conceptual debates, and can (again, loosely) be said to have lasted until the 1980s. Subsequently, there was something of a hiatus, with a relative paucity of work – especially empirical work – in the 1990s. During the first decade of the twenty-first century, however, there was a reawakening of interest in the experiments, stimulated in part by the increasing availability of material from the experiments in Milgram's archive. In Chapters 1 and 2 I will provide an overview of this work. In Chapter 2 I will focus on the 'new wave' of critique, commentary and analysis. First however, the present chapter will review the initial wave of post-Milgram scholarship, as well as providing a summary of Milgram's experiments themselves.

In reviewing Milgram's experiments and the first wave of extensions, replications and critiques, the aim is not to be comprehensive but rather to survey the main themes and arguments that are apparent in this rich literature. Arthur G. Miller (1986) provided the definitive account of the first 20 or so years of scholarship provoked by Milgram's studies in his comprehensive and scholarly book, *The Obedience Experiments: A Case Study of Controversy in Social Science*. Miller, of course, has his own position on the experiments, and it would not be unfair to describe him as essentially – though not uncritically – of the view that Milgram's studies were, and remain, valuable and important contributions to psychology and the wider social sciences (for a restatement and updating of his position, see Miller, 2016). Regardless, however, of one's own take on the obedience experiments, Miller's (1986) book remains a valuable resource.

In contrast, my approach in the present chapter is necessarily somewhat more selective. In one respect, this is because the job of summary and review of this work has been made much easier by the simple fact that others have done it so well elsewhere. In another, equally important, respect, this is because the burgeoning of a renewed primary and secondary literature on the obedience experiments has moved things along considerably. This is not to say that the arguments made in the first 20 of 30 years of what we might term 'Milgram scholarship' are no longer relevant, and we should always be mindful of ignorance lest we slip it merely pouring old wine into new bottles. However, as will be suggested in Chapter 2, the renewed attempts to revisit Milgram empirically, coupled with the increasing focus on the lessons to be drawn from close scrutiny of Milgram's archives, add layers of complexity to the story of the obedience experiments that were simply not possible until recent years.

A similar argument is necessary in relation to the obedience experiments themselves. Whatever else he was or was not, Milgram was a fine writer who combined accessibility with gravity in order to produce a highly readable account of the obedience experiments (Milgram, 1974). This, together with a handful of earlier empirical papers (Milgram, 1963, 1965a, b), replies and commentaries (1964b, 1967, 1972, 1977, 1983), and – in a slightly different vein – his documentary film of the experiments (Milgram, 1965c), constitutes the 'official' version of the obedience experiments. The purpose of providing an introductory summary of Milgram's experiments here is not so much to orient readers to what the experiments were, or what 'happened' in them, for – as will be argued in subsequent chapters – archival researchers have highlighted several problems with relying on Milgram's account of the studies. Rather, they are summarised precisely to provide an overview of Milgram's account as an *account* of his studies. As Griggs and Whitehead (Griggs, 2017; Griggs & Whitehead, 2015a, b) have recently shown in their analyses of textbook coverage of the obedience experiments, this account is remarkably resistant to change and continues to frame the way in which the obedience experiments are understood. With this is mind, I will now turn to outlining what we might call the 'standard story', or the 'received account' of Milgram's studies. I will begin by outlining the most well known of Milgram's experiments.

The Experiments

Milgram's participants took part in what they were led to believe was a study of the effects of punishment on learning. Shortly after a participant

had arrived at the laboratory, a second person arrived, and was introduced to them as Mr Wallace. Although it appeared that 'Mr Wallace' was another naïve participant, he was in fact a confederate – Jim McDonough – employed by Milgram. The experimenter – played by John Williams – explained to the naïve participant and 'Mr Wallace' that one of them would take the role of teacher and the other the role of learner. A rigged selection process took place in which the naïve participant was always allocated the role of teacher, and the confederate the role of learner. The learner was taken to an adjoining room where, as the teacher looked on, the experimenter strapped electrodes to him in order that he would be able to receive punishment in the form of electric shocks.

Returning to the main laboratory, the experimenter asked the teacher to sit in front of an imposing machine for generating electric shocks. This machine featured a series of levers for administering the shocks, beginning at 15 volts and rising in 15-volt increments to 450 volts. The experimenter administered a sample shock of 45 volts to the teacher (this was the only genuine shock used in the whole experiment), and explained how the experimental procedure was to work.

The teacher was to read a series of word pairs into a microphone. These would be heard by the learner in the next room, who would try to remember the word pairs. The teacher would then need to test the learner on the word pairs, and would do this by reading the first word of each pair in turn, followed by four choices. The learner had to indicate which of the four choices was correct by pressing one of four buttons which would light up the corresponding response on a box in the main laboratory. If the response was correct, the teacher was to move on to the next item in the test. If, however, it was incorrect, he was to administer an electric shock as punishment for the error. To do this he had to say 'wrong', then state the voltage to be delivered, press the appropriate shock lever, and then read the correct answer to the learner.

The learner provided his responses according to a preset order that ensured that he would get many of the word pairs incorrect. As the experiment unfolded, it therefore quickly became apparent to the naïve participants that the learner was going to require increasingly strong shocks. At 75 volts the learner began to yelp following the administration of the shock, and the intensity of these exclamations escalated until he demanded to be released following the 150-volt shock. If participants continued, the protests continued, becoming more aggravated until the learner refused to answer following the 300-volt shock. From 345 volts onwards, each successive shock was met only with silence, leading participants to assume that the learner was unconscious, or worse.

If at any point during the experimental session the participant hesitated or refused to continue, the experimenter had a series of 'prods' at his disposal that he could use in an attempt to get the participant to continue with the experiment. Four of these were sequential prods, and were to be used in order and started afresh for each new attempt at resistance:

> 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.
> (Milgram, 1974, p. 21, italics in original)

Only when a participant had successfully defied the fourth prod was an experimental session terminated. In addition, the experimenter could use two 'special' prods to answer specific queries from participants as appropriate. These were: 'Although the shocks may be painful, there is no permanent tissue damage, so please go on' (Milgram, 1974) 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, p. 22). If participants continued to 450 volts, the experimenter continued with the test until they had administered this shock three times and then discontinued the experimental session. Such participants were recorded as obedient, with those who managed to resist sufficiently to draw the experiment to a close recorded as disobedient.

Other Experimental Conditions

The procedure outlined above is well known, in no small part due to its prominence in Milgram's (1965c) film of his experiments. It was used as the basis of four conditions of the experiments: 'voice-feedback', 'a new baseline', 'change of personnel' and 'women as subjects'. The 'new baseline' and 'change of personnel' conditions added a heart condition for the learner, which he raised with the experimenter while having the electrodes strapped to his arm, and then again as he was protesting at various points during the experiment. 'Change of personnel', as its name implies, used the same procedure but featured different confederates in the roles of teacher and learner, and 'women as subjects', again as implied by its (now rather dated) name, was the only condition in which women took part.

However, this well-known procedure is only one of many variations used by Milgram. In an attempt to identify and test several factors that may influence obedience, Milgram ran numerous variations of his experiment. His first publication on the obedience experiments (Milgram, 1963) outlined what became known as the 'remote' condition.

Table 1.1 *Summary of obedience rates in Milgram's (1974) experimental conditions*

Condition number and name	Obedience % (N)	
1. Remote-victim	65	(26/40)
2. Voice-feedback	62.5	(25/40)
3. Proximity	40	(16/40)
4. Touch-proximity	30	(12/40)
5. A new baseline	65	(26/40)
6. Change of personnel	50	(20/40)
7. Closeness of authority	20.5	(9/40)
8. Women as subjects	65	(26/40)
9. The victim's limited contract	40	(16/40)
10. Institutional context	47.5	(19/40)
11. Subject free to choose shock level	2.5	(1/40)
12. Learner demands to be shocked	0	(0/20)
13. An ordinary man gives orders	20	(4/20)
13a. The subject as bystander	68.75	(11/16)
14. Authority as victim: An ordinary man commanding	0	(0/20)
15. Two authorities: Contradictory commands	0	(0/20)
16. Two authorities: One as victim	65	(13/20)
17. Two peers rebel	10	(4/40)
18. A peer administers shocks	92.50	(37/40)

In this version, the experiment proceeds along similar lines to those described above, but instead of the repeated verbal protests from the learner, there is instead only a pounding on the walls following the 300-volt shock.

Table 1.1 presents a summary of obedience rates in Milgram's (1974) experimental conditions. The signature findings from Milgram's studies are typically identified as the initial finding of 65 per cent obedience in the 'remote' condition (Milgram, 1963), and the finding that the addition of repeated verbal protests did not reduce obedience, yielding 62.5 per cent obedience in the 'voice-feedback' condition, and 65 per cent in the 'new baseline' condition, which was replicated in the only condition in which women took part ('women as subjects'), also yielding a 65 per cent obedience rate.

The number of conditions in which defiance was more common than obedience is notable. In 11 out of the 19 conditions (or 10 out of 17 if we discount conditions 11 and 13a, in which the dependent measure was not really comparable with that used in the other conditions), defiant participants outnumbered obedient ones. The mean rate of obedience was thus 39.17 per cent (or 39.59% excluding 11 and 13a), and the total number of obedient participants was 265 out of 636 (or 253 out

of 580 excluding 11 and 13a; see also Haslam, Loughnan & Perry, 2014 for a meta-analytic overview of Milgram's experimental conditions). We should be wary of making too much of this given that Milgram explicitly aimed to vary the factors that would make obedience more or less likely, and as such the observation that most participants across all conditions were actually defiant does not stand as a challenge to the basic finding that in the 'standard' conditions obedience rates were rather high, but nevertheless it serves as a useful reminder that there is much more to the experiments than the classic 65 per cent finding.

Milgram (1974) outlined the results of his experimental conditions in four thematic stages: conditions 1–4 deal with the proximity of the victim, conditions 5–11 deal with 'further variations and controls' (p. 55), conditions 12–16 with variations in the experimental roles and conditions 17–18 with the influence of group processes. In order to contextualise the information provided in Table 1.1 it is worth briefly outlining each of these four sets of conditions.

Proximity

The proximity series consisted of the remote condition first outlined by Milgram (1963), together with the voice-feedback, proximity and touch-proximity conditions. The results of this series of experiments was first reported by Milgram (1965a), and is typically held to show how bringing the learner physically (and psychologically) closer to the teacher increases the pressure on the perpetrator and thus results in reduced obedience. The remote and voice-feedback conditions were outlined above, and it was also noted that the introduction of verbal protests in the voice-feedback condition did not notably reduce obedience from the remote condition. However, arguably the key conditions in this series are the proximity and touch-proximity conditions. In both these conditions, the learner was seated in the same room as the teacher, and thus the teacher was more immediately confronted with the learner's apparent pain and anguish. In the touch-proximity condition, the teacher also had to physically hold the learner's hand down onto a shock plate in order that the punishment could be administered. These conditions resulted in reduced levels of obedience (see Table 1.1).

Further Variations and Controls

As Milgram (1974) notes, he had to move laboratories during his experiments. Condition 5 ('a new baseline') represents his attempt to replicate what he by now appeared to consider the baseline finding against which

all others were to be compared in his new, somewhat less impressive, facilities. The procedure for this condition was based on the voice-feedback condition (condition 2), but featured the addition of a heart complaint for the learner, which was introduced as the electrodes were being attached to his arm. The new laboratory did not appear to have an impact on obedience rates, but in condition 6 ('change of personnel'), there was a modest reduction. As noted previously, this condition featured different confederates in the roles of teacher and learner, but was in all other respects the same as 'a new baseline'.

Condition 7 extended the logic of the proximity series to the distance between authority and teacher. In this condition, the experimenter left the room and gave his instructions over the telephone, leading to a further reduction in obedience levels. Condition 8 was the only condition in which women took part, and yielded an identical obedience rate as the equivalent condition with male participants ('a new baseline'). Condition 9 featured the learner placing a clear condition on his participation before the experiment began, in which he consented to take part only if he could be released when he said so. Condition 10 (also known as the 'Bridgeport' condition) removed the experiment from Yale University to an office building in the nearby town of Bridgeport, with Milgram's aim being to examine the extent to which the institutional authority associated with Yale might have impacted on obedience. Condition 11 was probably the closest thing that Milgram got to what might be regarded as a genuine experimental control group. In this condition the participants were not instructed to administer shocks of increasing severity, but were instead able to punish the learner using whichever levers they chose. In this experiment, the 'obedience' rate of 2.5 per cent indicates the percentage of participants who administered the 450-volt shock at any point in the procedure, not simply the number who reached the end of the shock scale without defying the experimenter. Moreover, in this condition the majority of participants (95%) exclusively used shocks below 150 volts.

Role Permutations

Conditions 12–16 all feature some form of variation of the roles employed in the experimental scenario. In condition 12, the experimenter sought to draw the session to a close as a result of the learner's cries of pain, but the learner himself insisted they keep going and that the teacher continue administering him shocks. In condition 13, the experimenter left the room and an additional confederate, who appeared to the naïve participant to be simply another member of the public, seemingly took it upon himself to instruct the participant to continue. In both these

situations, Milgram argues that obedience was reduced as the source of the commands to keep administering the shocks was not an authority figure but someone who appeared to be merely another participant. However, Milgram allowed condition 13 to continue after the participant had effected their withdrawal, with the additional 'ordinary' confederate administering the shocks himself (Milgram labelled this condition 13a). Milgram was interested to see how many participants would intervene in this scenario. In these circumstances, the 'obedience' rate of 68.75 per cent refers to the proportion of experimental sessions in which the 'ordinary' confederate managed to reach 450 volts. Although this appears to indicate that few participants were able to resist in this situation, Milgram's (1974) summary of it emphasises resistance: even the majority of participants who did not manage to prevent the confederate from administering all the shocks put up quite a show of resistance, and those who did persist and were able to curtail the experimental session appear to have done so by taking physical action (e.g. unplugging the shock generator; restraining the confederate).

In condition 14, an apparently 'ordinary man' gave orders with the experimenter in the role of learner, and in condition 15 there were two experimenters who gave contradictory commands. In both these conditions obedience was reduced to zero. In condition 16, there were again two experimenters, but this time one of them took on the role of learner with the other in the usual experimenter role of issuing orders. As Milgram (1974, p. 109) notes, in this situation the experimenter 'fares no better than a victim who is not an authority at all.'

Group Processes

The final two conditions outlined in Milgram's (1974) book focussed on the effects of placing the participant in a group of teachers, rather than being alone. In condition 17 ('two peers rebel'), the two additional confederate teachers withdrew as the learner's protests intensified. This led to a notable reduction in obedience when compared to the baseline conditions. By contrast, in condition 18, the naïve participant fulfils a subsidiary role, with the act of administering the electric shocks performed by a confederate. In this condition, removed from the act of delivering the punishment themselves, a greater number of participants remained in the experiment than in any other condition.

It is worth noting that the presentation of Milgram's experimental conditions in these four thematic groupings does not necessarily reflect the order in which the conditions were conducted. Milgram's archive

features a different numbering system in which several conditions are numbered differently, with these original numberings appearing to reflect the order in which the conditions were conducted. For example, 'two peers rebel' is condition 7 in the archival record, rather than condition 17 as in Milgram's (1974) book. Similarly, 'women as subjects' is condition 20, rather than condition 8, in the archives, and 'institutional context' is condition 23 rather than 10.

In addition, there are further conditions that were not reported by Milgram (1974). Some were reported elsewhere, such as a condition on group effects (Milgram, 1964a), which might more properly be understood as dealing with conformity rather than obedience. Others, such as the 'relationship' or 'bring a friend' condition (Perry, 2012; Rochat & Modigliani, 1997; Russell, 2014a), were not published by Milgram at all. For reasons of space, I will not summarise these conditions here (though see Chapter 2 for a consideration of the relationship condition).

These summaries provide an overview of what are typically seen as the key aspects of the Milgram experiments: the experimental procedures and their respective obedience rates. However, Milgram (1974) also includes in his fullest account of his studies a series of vignettes concerning the performance of particular individuals in the experiments, and at various points includes excerpts from transcripts of several of the experimental sessions. A key aspect of my argument in the present volume is that this sort of data from the experiments has been largely neglected, and whilst many have considered the statistical details of the obedience rates in Milgram's different experimental conditions and have interrogated his theoretical account, there has been relatively little attention paid to what, for me, seems to be a much richer form of data from the experiments – the experimental interactions themselves. By considering the excerpts presented by Milgram, we can begin to point the way to a rather different way of thinking about the experiments.

For example, the following extract features a participant from condition 3 who Milgram identifies as a 'Professor of Old Testament'. The exchange occurs after the participant had reached 150 volts on the shock generator:

EXPERIMENTER: It's absolutely essential to the experiment that we continue.
SUBJECT: I understand that statement, but I don't understand why the experiment is placed above this person's life.
EXPERIMENTER: There is no permanent tissue damage.
SUBJECT: Well, that's your opinion. If he doesn't want to continue, I'm taking orders from him.
EXPERIMENTER: You have no other choice, sir, you must go on.
SUBJECT: If this were Russia maybe, but not in America.

(Milgram, 1974, p. 48)

After this exchange, Milgram reports that the experimental session was discontinued, and goes on to point to the way in which the participant 'initially justified his breaking off the experiment not by asserting disobedience but by asserting that he would then take orders from the victim' (Milgram, 1974, p. 49). Milgram is interested in what this exchange tells us about the psychological dynamics of the situation, and speculates that the participant here undergoes a shift in the source of authority, thus he 'does not disobey so much as shifts the person from whom he will take orders' (Milgram, 1974). This may be the case, but of course we have no way of knowing what processes were occurring 'under the skull' of this participant. What we can say, however, is how the participant's words function pragmatically as a means of effecting his withdrawal from the experiment. Whether they reflect some genuine underlying psychological shift is thus secondary to the question of what they *do* for him in the context of their utterance. Importantly, the function is the same whether or not we impute some hypothetical change of mental state.

We might therefore conceptualise this utterance as part of a rhetorical strategy for extricating the participant from the experimental situation. Indeed, when we place the utterance in the context of the participant's whole speaking turn, we can begin to see how it is part of a more general rhetorical strategy of undermining the experimenter's authority. Significantly, the participant responds to the experimenter's 'There is no permanent tissue damage' prompt with a classic rhetorical move which functions to undermine the factual status of the experimenter's statement: He glosses what has been offered as a statement of fact as merely an *opinion*, and thereby challenges the experimenter's authority to pronounce on the physical effects of electric shocks. In providing such a reading of the transcript, we have no need to seek to identify whether the participant *really* thinks that the experimenter's statement about tissue damage is simply a matter of opinion, nor need we be concerned with whether the participant undergoes some sort of psychological shift from obeying the orders of the experimenter to obeying the participant. What is crucial is that these utterances perform a particular rhetorical function in the local context of their use. They enable the participant to subvert the experimenter's authority, and in so doing provide a potential 'way out' of the experimental situation.

Even if we want to retain a concern with the correspondence between what participants say and some putative underlying belief, identity or knowledge, this is arguably still not enough to explain the way in which they attempt to challenge the experimenter. Consider the following example of a participant from condition 2, to whom Milgram gave the pseudonym 'Jan Rensaleer':

EXPERIMENTER: There is no permanent tissue damage.
MR. RENSALEER: Yes, but I know what shocks do to you. I'm an electrical engineer, and I have had shocks ... and you get real shook up by them – especially if you know the next one is coming. I'm sorry.

(Milgram, 1974, p. 51; ellipsis in original)

We might use this as an example of how a participant was able to draw on his occupational identity, which conferred specialist expertise concerning the effects of electric shocks. We may have no particular reason to doubt that Rensaleer is telling the truth – he *is* an electrical engineer, and he *does* know what shocks can do. But many other things will also be true about this man – Milgram tells us that he emigrated from the Netherlands and belongs to the Dutch Reformed Church. Perhaps he also enjoys baseball, or likes to play cards. Maybe he is a husband and a father. Any and all of these may be true, but they are not mentioned here as he attempts to argue his way out of the experiment. Mere truth is inadequate as a criterion for studying the to-and-fro of argumentation; we need to consider what any particular utterance does in the specific context in which it is uttered. And here, again, we can see how Rensaleer, in a quite different way to the 'Professor of Old Testament', uses the self-category of 'electrical engineer', and the associated knowledge claim concerning the effects of electric shocks, in order to undermine the experimenter's authority and to challenge his entitlement to pronounce on the safety of the shocks.

I will return to these issues in subsequent chapters, where they will form a centrepiece of my argument for a respecification of the obedience experiments. For now, however, I will turn to consider Milgram's theoretical account of his findings.

Milgram's Theoretical Account: The Agentic State

The obedience experiments were not designed to test specific hypotheses derived from one or more theoretical perspectives. Instead, they are perhaps best understood as the outcome of an exploratory process of inductive research in which Milgram gradually moved through a series of situational variables that might affect the extent to which people obey orders from an authority figure. Milgram's (1963, 1965a) early obedience publications contained some theoretical ideas, but these were not formed into an integrated theoretical story that helped to organise and make sense of his experimental findings. For example, when considering the findings of the proximity series and of the experimenter-absent condition, Milgram (1965a, p. 66) speculated that 'it would appear that

something akin to fields of force, diminishing in effectiveness with increasing psychological distance from their source, have a controlling effect on the subject's performance.' However, by the time his fullest account of the experiments was published, Milgram (1974) had developed an overarching theoretical account of his findings.

Milgram's (1974) attempt to draw together his empirical findings into a coherent theoretical account relies on the concept of the *agentic state*. Drawing on cybernetics, Milgram suggests that an individual organism functioning in a system of interrelationships with other organisms – such as human society – needs some sort of process through which to regulate autonomous behaviour. If everyone simply went about their business behaving freely all the time, social organisation would be impossible. In particular, Milgram notes the preponderance of hierarchal systems in human society, and considers what is needed for such social structures to operate. Milgram proposed that:

> The critical shift in functioning is reflected in an alteration of attitude. Specifically, the person entering an authority system no longer views himself [*sic*] as acting out of his own purposes but rather comes to see himself as an agent for executing the wishes of another person. Once an individual conceives his action in this light, profound alterations occur in his behavior and his internal functioning. These are so pronounced that one may say that this altered attitude places the individual in a different *state* from the one he was in prior to integration into the hierarchy. I shall term this *the agentic state*, by which I mean the condition a person is in when he sees himself as an agent for carrying out another person's wishes. (Milgram, 1974, p. 133, italics in original)

Importantly, when in the agentic state, a person is rendered 'open to regulation by a person of higher status. In this condition the individual no longer views himself as responsible for his own actions but defines himself as in instrument for carrying out the wishes of others' (Milgram, 1974, p. 134).

Milgram goes on to articulate the factors that lead to obedience as arising from the relationship between what he terms *binding factors* and *sources of strain*. Binding factors are those aspects of the experimental situation that keep participants obeying the experimenter. In contrast, sources of strain are those aspects of the situation that create strain for the participants and lead them towards disobedience. When the binding factors outweigh the strain, obedience is the result; by contrast, when the strain is greater than the binding factors then disobedience will result.

Binding factors include features of the experimental design such as the gradated nature of the shock sequence (see also Gilbert, 1981), the pre-established obligation that the participants feel towards the experimenter and the anxiety experienced by participants as they contemplate defiance.

Sources of strain include the learner's cries and demands to be released, the fear of retaliation (either in the form of direct retribution from the learner, or possible legal ramifications) and the experience of contravening values concerning how others should be treated. Each of these produces in the participants a palpable sense of strain, and in general terms increases the likelihood that participants will defy the experimenter. Milgram thus explains the variations in his findings as a function of the way in which binding factors and sources of strain were given greater or lesser prominence in the structure of the different conditions.

Milgram's theoretical account has important implications for how we might understand notions of choice and responsibility. Milgram does build choice into his account – he argues that participants do have a certain amount of choice as to whether to place themselves in a situation in which they will be subject to hierarchical relations, and even the extent to which they come to take on the agentic state. However, he notes that situational pressures towards entering the agentic state are extremely powerful and as such place strict limitations on individual choice. As a result, Milgram suggests that the extent to which people can be held responsible for their actions when in the agentic state is limited:

> The most frequent defense of the individual who has performed a heinous act under command of an authority is that he has simply done his duty. In asserting this defense, the individual is not introducing an alibi concocted for the moment but is reporting honestly on the psychological attitude induced by submission to authority. (Milgram, 1974, p. 146)

In an appendix, Milgram (1974) provides data that he interprets as supporting this proposition. In the four experimental conditions that make up the proximity series, participants were asked to use an instrument known as the responsibility clock to indicate the proportion of responsibility for shocking the learner against his will that they would allocate to themselves, to the experimenter and to the learner himself. Milgram (1974, p. 204) notes that 'the defiant subjects, more often than the obedient subjects, attribute primary responsibility to themselves. And they attribute less responsibility to the learner.' While cautioning that these are post hoc accounts and should necessarily be treated with caution, Milgram nevertheless suggested that disobedient participants see themselves as primarily responsible for shocking the learner, whereas obedient participants do not.

Milgram's theory thus constitutes an attempt to account for the variation in obedience levels across conditions, and this has important implications for the philosophical and practical way in which we understand agency and responsibility. Notably, and perhaps troublingly, it suggests

that 'just following orders' may have some basis in the reality of underlying psychological processes.

Critical Reaction

Milgram's experiments provoked an almost immediate critical reaction, and in many respects this continues to the present day (see Chapter 2). Key early criticisms of Milgram's experiment highlighted a number of important ethical, methodological and theoretical weaknesses, and it is testament to the controversy generated by the obedience experiments that many of these critiques have gone on to be an integral part of the story of the Milgram experiments.

Ethics

Diana Baumrind's (1964) seminal article set the tone for much of the criticism of the obedience experiments on ethical grounds. As Miller (2013) has recently noted, Baumrind's critique has been almost as influential as Milgram's experiments themselves, both in terms of bringing the experiments to the attention of a wider disciplinary readership, and in framing the terms of the ethical debate at a time when such matters were not the subject of routine discussion in the academic literature. Indeed, it is arguable that the ethical controversy resulting from the experiments played a key role in the debates that led to the more robust codification of research ethics in US psychology a decade later (Stark, 2010).

Baumrind's criticism of Milgram centred on the unacceptability of the use of deception, in particular given the nature of the experimental task to which participants were subjected. She argues that the stress caused to participants was not justified, and indeed challenges Milgram's claims that the importance of the findings mitigates the momentary discomfort caused to participants. Indeed, Baumrind draws attention to methodological problems with the experiments that limit the extent to which they can be seen as providing any useful insight into broader processes anyway (see discussion later in the chapter). Furthermore, Baumrind is not convinced that the participants would have suffered no long-term consequences, and more generally is concerned that such research undermines trust in the discipline of psychology. She concludes that,

> I would not like to see experiments such as Milgram's proceed unless the subjects were fully informed of the dangers of serious aftereffects and his correctives were clearly shown to be effective in restoring their state of well being. (Baumrind, 1964, p. 423)

This critique has been expanded, updated and restated at various points over the years (e.g. Baumrind, 1985, 2013, 2015), and has set the tone for the ethical controversy that has accompanied any discussion of the obedience experiments (although see Nicholson, 2011, for an argument that the obedience experiments underwent something of a 'rehabilitation' that led to a de-emphasis on ethical issues).

Other important critiques came from Kelman (1967) and Patten (1977a). As part of a wider critique of the use of deception in social psychology, Kelman (1967) challenged Milgram's (1964b; see also Kaufmann, 1967) claim that participants in his experiments had been given the opportunity to learn something about themselves:

> If this were a lesson from life, it would indeed constitute an instructive confrontation and provide a valuable insight. But do we, for the purpose of experimentation, have the right to provide such potentially disturbing insights to subjects who do not know that this is what they are coming for? (Kelman, 1967, p. 4)

Patten (1977a) extended the ethical critique of Milgram by arguing for an equivalence between Milgram's actions in conducting his research and those of his participants in administering the shocks. Patten argued that the grounds on which Milgram seeks to exonerate his experiments can also be used to exonerate the actions of his participants. If Milgram can assert that placing his participants in a stressful situation is acceptable owing to the potential utility of the knowledge to be gained, so his participants can in turn be exonerated on the grounds that they too administered shocks in the service of what they thought were the higher ideals of science. In both cases, the ends justify the means. However, if Milgram wishes to cast his participants' actions as immoral then his own actions must be characterised in similar terms. Milgram cannot, therefore, 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.

Milgram's (1964b, 1974, 1977) response to the criticism of the ethics of his research was based around the issue of participants' reactions to the experiment. Milgram (1964b, p. 848) argued that 'The extreme tension induced in some subjects was unexpected.' However, recognising that this only applies to the very early stages of his research programme, after which any defence about being taken by surprise would cease to apply, he goes on to draw a distinction between 'momentary excitement' and 'harm' (p. 849). Thus he argues that while participants may have experienced stress during the experiment, there were no lasting effects. Milgram emphasises the importance of his 'dehoaxing' procedure, and

cites two sources of follow-up data in support of his claims: first, responses to a post-experiment questionnaire in which only 1.3 per cent of participants indicated negative feelings about having participated in the experiment (and 83.7% indicated positive feelings); second, a summary of a report by a psychiatrist (Paul Errera) who had focussed on 40 participants who 'he felt would be most likely to have suffered consequences from participation' (Milgram, 1964b, p. 850). Milgram quotes Errera's conclusions as indicating that no long-term consequences of participation in the experiments had been observed.

However, given the central role of participant reactions, it is worth noting a contribution to these debates that is unusual in that it takes an empirical approach to the investigation of ethical issues. Ring, Wallstone and Corey (1970) explored the effects of different modes of debriefing on participant reactions to a Milgram-esque paradigm. They found that participants who were given a full debriefing that validated the behaviour they had displayed in the experiment were generally positive about their participation, whereas participants who received no debrief were more negative. This appears to bear out Milgram's arguments about the absence of negative reactions, although whether those reactions should be (a) believed and (b) treated as answering all ethical objections if they are believed (Kelman, 1967; Patten, 1977a) is another matter. Moreover, as we will see in Chapter 2, given the use of Milgram's archival data to highlight the inadequacy of his debriefing processes (Nicholson, 2011; Perry, 2012), things are now a little more complex than they previously appeared.

Methodology

The most notable methodological critique of Milgram in the 1960s came in the form of Orne and Holland's (1968) influential argument that the experimental scenario created by Milgram was likely to have produced demand characteristics. Although now a staple of undergraduate methodological training in psychology, at the time the idea of demand characteristics was a relatively novel concept (Orne, 1962), which can be understood as part of a broader movement to consider the psychology of the psychology experiment (Miller, 1972a; Rosenthal, 1966). Orne and Holland (1968) argued that participants, acting on cues embedded in the situation, would have sensed that the situation was not all that it seemed. In particular, they argued that the impassivity of the experimenter in response to the apparently anguished cries of the learner would have acted as an indication to participants that they were the real focus of the investigation. Unsure of precisely what to do in this unusual situation,

participants behaved in a way that they sensed Milgram wanted them to behave so as not to be seen to be disrupting his research. As such, the validity of the experiments is called into question.

In his response, Milgram (1972) noted that Orne and Holland's criticisms are tenuous in that they are not grounded in empirical data, but rather in speculation as to what participants might have thought about the experiments. Milgram draws on his own empirical data to challenge the claim that participants would have seen through the cover story. Specifically, he points to a post-experiment measure in which participants indicated that they believed they had caused extreme pain to the learner, and follow-up questionnaires which show that only a minority of participants had harboured significant doubts as to whether the learner was actually getting the shocks. Moreover, Milgram suggests that even where participants did indicate doubt, this may be as much a product of defensiveness as actual belief. It may be more palatable to convince oneself that you always knew the experiment to be a sham than to confront the uncomfortable fact of having administered what you believed to be painful electric shocks. Milgram also highlights what he sees as a fundamental misconception in Orne and Holland's arguments. Whereas for Orne and Holland the fact that participants trust the experimenter and assume that nothing untoward can really happen is reason to doubt the validity of the experiments, for Milgram it is precisely this feature of the experimenter–participant relationship that makes his obedience situation so powerful. Participants put their trust in the experimenter despite increasing indications that this trust might be misplaced (principally, the reactions of the learner), and this stands as an indication of the power of authority figures to elicit obedience.

Miller (1986) notes that Milgram's point about participants continuing because they trust the experimenter may be slightly misplaced in that Orne and Holland's argument is not that the actions of participants are not directed by the experimenter, but that participants went along with the experimenter because they knew that no one was really getting harmed – *in spite of* the protests from the learner. Miller also points out that, although not explicitly stated by Orne and Holland, the implication of their critique is that whereas participants who were sure the experiment was a set-up were obedient, those who believed the shocks were genuine were defiant. As such, the results of Milgram's experiments would take on a much more optimistic gloss than is usually the case – when they believed they were causing pain to someone, participants *did not* obey the experimenter.

Concerns around the validity of the experiments are also apparent in discussions concerning the extent to which the experiments provide

insights that are of more general applicability beyond the laboratory. These concerns with generalisation raise in a slightly different way the concern with ecological validity highlighted by Orne and Holland's (1968) demand characteristics critique.

As discussed in the Introduction, Milgram (1963, 1974) framed his obedience research as being of direct relevance to the Holocaust. Blass (2004) has documented the ways in which the Holocaust informed Milgram's thinking in relation to the obedience experiments (and, indeed, more broadly in terms of its impact on his intellectual trajectory). The experiments have been drawn on extensively in attempts to make sense of how the Nazis were able to put the final solution into practice, and yet they have again been the source of much controversy on this matter (Miller, 2004). Moreover, the extent to which they can be generalised to other abuses and atrocities has been the subject of much debate (e.g. Kelman & Hamilton, 1989).

It is important to note that Milgram did not conceive of his experiments as providing *the* explanation for the Holocaust – to do so would have been overly simplistic and easily dismissed. Rather he claimed to have developed a technique for studying the essential psychological mechanism underlying the operation of hierarchical command structures, which helps to illuminate how such systems could function to elicit behaviour that would result in atrocities. Fundamentally, Milgram's account leads to the conclusion that seemingly ordinary people can be made to perform heinous acts. This moves us away from the idea that people who carry out such acts do so as a result of individual pathology. In this respect, and has often been remarked, Milgram's perspective complements that of Hannah Arendt (1963/1977), who in observing the trial of Adolf Eichmann in 1961 coined the term *the banality of evil* as a means of capturing the extent to which Eichmann presented himself as a mere bureaucrat, more concerned with doing his job than with any commitment to the extermination of the Jews. It has been argued that this is based on a misreading of Arendt's thesis (Lang, 2014; Reicher, 2014; and see Chapter 2), but nevertheless Milgram himself saw his findings as confirming Arendt's philosophical analysis:

> Arendt's conception of the *banality of evil* comes closer to the truth than one might dare imagine. The ordinary person who shocked his victim did so out of a sense of obligation – a conception of his duties as a subject – and not from any particularly aggressive tendencies. (Milgram, 1974, p. 6, italics in original)

Analyses that are essentially sympathetic with Milgram's have been advanced by a number of scholars, both in relation to the Holocaust itself and to atrocities such as the My Lai massacre in the Vietnam War

(e.g. Blass, 1993; Kelman & Hamilton, 1989; Sabini & Silver, 1980; see Miller, 1986, chapter 7, for a review). However, while Miller (1986) suggests that initial reactions – especially within social psychology – were broadly in favour of seeing the experiments as being able to shed light on some of the psychological dynamics underlying the Holocaust, there was no shortage of dissenting voices.

Many of these dissenting voices have built on other methodological criticisms in order to argue that the obedience studies are of little or no value in generalising beyond the laboratory. In her seminal ethical critique of Milgram's (1963) initial obedience publication, Baumrind (1964) also gave an early airing to the sort of arguments that would come to be used against any attempt to see the experiments as providing an insight into the Holocaust:

> the parallel between authority-subordinate relationships in Hitler's Germany and in Milgram's laboratory is unclear. In the former situation the SS man or member of the German Officer Corps, when obeying orders to slaughter, had no reason to think of his superior officer as benignly disposed towards himself or their victims. The victims were perceived as subhuman and not worthy of consideration. The subordinate officer was an agent in a great cause. He did not need to feel guilt or conflict because within his frame of reference he was acting rightly. (Baumrind, 1964, p. 423)

This argument, in varying forms, has been developed and extended by many scholars who were not convinced of the comparability between the obedience experiments and the Nazi Holocaust – or, indeed, of other real-life atrocities. For Fromm (1973), the role of science as the source of authority in the experiments is crucial, and this limits the extent to which it can be compared with situations where the commands come from an altogether different source of authority. Similarly, Patten (1977b) draws a distinction between authority based on expertise (i.e. that associated with scientists) and authority not based on expertise (e.g. the authority of military commanders). Again, Patten argues that this key difference limits the generalisability of Milgram's findings – which are derived from a context in which the authority figure derived his authority from expertise – to events such as the Holocaust and the My Lai massacre where authority was based on a nonexpert hierarchical system.

This points to a more general sense in which some scholars have argued that the obedience experiments didn't really tell us anything that we don't already know (Mixon, 1989). Even a cursory inspection of history will point to numerous occasions on which atrocities have been committed by people acting under orders that they believe to come from a legitimate authority.

Theory

As outlined above, Milgram (1974) developed a theoretical account of his findings that revolved around the concept of the agentic state. However, Milgram's theory is not generally regarded as a very good explanation for his findings, even by those scholars who are otherwise positively disposed towards the obedience experiments (Blass, 2004; Miller, 1986). Empirical tests of the theory are, perhaps surprisingly, somewhat sparse, and in many respects this explains why the theory is not particularly well-regarded – there have simply been too few attempts to test it. Moreover, what little evidence there is tends to point to the limitations of the theory.

The key arguments pertaining to Milgram's theory were developed by Mantell and Panzarella (1976). They outlined two problems: First, Milgram's (1974) own data do not appear to support it particularly well; second, in their own empirical test of the theory it was found wanting. In relation to Milgram's own 'responsibility clock' data, they note that while obedient participants did appear to allocate a lower proportion of responsibility to themselves than disobedient participants, they point out that they still allocate to themselves almost as great a proportion of responsibility as they do to the experimenter. The obedient participants evidently did not simply relinquish personal responsibility in any straightforward fashion.

Mantell and Panzarella then tested Milgram's theory by exploring responsibility data in relation to their own replications of the obedience experiments. They found that post-experimental attributions of responsibility did not predict performance in the experiments. Specifically, obedient participants did not attribute greater responsibility to the experimenter, and disobedient participants did not attribute greater responsibility to themselves.

A final notable argument against the agentic state theory was provided by Helm and Morelli (1979). As Milgram (1963, 1974) noted, many obedient participants displayed visible signs of tension and stress. Citing the example of 'Fred Prozi', who is the subject of one of Milgram's (1974) case study vignettes, and whose experimental session is featured at length in Milgram's (1965c) film, Helm and Morelli (1979) suggest that such behaviours are inconsistent with the passive, automaton-like image implied by the agentic state theory.

Early Extensions and Replications

In addition to debates concerning ethical, methodological and conceptual issues, the obedience experiments stimulated a number of follow-up

investigations that sought to replicate and/or extend Milgram's paradigm in various ways. However, given the influence of Milgram's original research programme, it is in some respects surprising that the volume of subsequent work by other investigators was rather modest. In other respects, of course, this is perhaps less surprising given the controversy surrounding the experiments. Even for researchers who found the experiments to be of intellectual importance, the ethical issues they raised may have been too serious to countenance an attempt to conduct similar experiments themselves (e.g. Smith, 1976).

The first wave of replications and extensions of the obedience experiments, taking place roughly between the late 1960s and the mid-1980s, is usefully summarised and reviewed by Blass (1999, 2012) and Miller (1986, chapter 4). The details of these various studies will not concern us here, but it is nevertheless instructive to note some of the variations involved. For example, Shanab and Yahya (1977, 1978) conducted cross-cultural replications in Jordan and Mantell (1971) in Germany. Kilham and Mann (1974), in addition to replicating Milgram's paradigm in Australia, explored the extent to which obedience varied depending on whether participants were administering shocks themselves or merely passing on the experimenter's instructions to someone else who would then administer the shocks. Powers and Geen (1972) explored the effects of having observed either an obedient or disobedient model prior to taking part in the experiment. It is also notable that some of these studies raised additional ethical issues over and above those associated with the standard Milgram paradigm. For example, Sheridan and King (1972) conducted a version of Milgram's paradigm with a genuine victim – a puppy – who was actually receiving electric shocks. Shanab and Yahya's (1977) study involved children as participants, some of whom were as young as 6 years of age.

These are perhaps the most notable extensions and replications in part because they were published in peer-reviewed journals. It is an indication of the paucity of such studies that scholars such as Blass (1999, 2012) are compelled to draw on a range of other experiments from doctoral theses and other unpublished sources that never made it into the peer-reviewed scientific literature. Nevertheless, if we leave aside these concerns for a moment, we can draw on the reviews of Miller and Blass to summarise the implications of this body of work. The overriding message that these authors derive from their reviews of this literature is the remarkable stability of obedience rates in the Milgram paradigm across culture and time, as well as across other variables such as age and gender.

Blass (1999) reviewed studies that had explored gender and obedience and found that, in line with Milgram's (1974) own findings, there was

little evidence for gender differences. Of 10 studies that compared obedience rates between men and women, nine showed no evidence of gender differences. Blass (1999) also explored the possibility that obedience rates may have changed over time, and found no relationship between obedience rates and year of publication across studies published between 1963 (Milgram's original publication) and 1985 (the most recent replication at the time of Blass's analysis). In a subsequent review, Blass (2012) focussed specifically on cross-cultural issues, comparing studies conducted in the USA with those conducted in other countries. Using data from studies that involved what he terms Milgram's 'standard' conditions, which include those based on the remote-victim, voice-feedback, new baseline and proximity conditions, Blass shows that although there is variation from study to study, the mean obedience rates are remarkably similar: 61 per cent for US studies and 66 per cent for studies conducted elsewhere (a difference that is statistically non-significant).

Beyond attempts to replicate Milgram's findings using experimental designs that seek to remain reasonably faithful to Milgram's procedure, two other extensions of Milgram's paradigm are worth summarising. First, a tradition of obedience research based on role-playing; second, Meeus and Raaijmakers's studies of administrative obedience.

Role-Playing

In response to the ethical criticism directed at Milgram's research, a number of researchers developed role-playing experiments based on Milgram's procedures (Geller, 1978; Mixon, 1972, 1976; O'Leary, Willis & Tomich, 1970). These were part of a wider debate concerning the merits of role-playing as a more ethically palatable alternative to deception (e.g. Freedman, 1969; Kelman, 1967; Miller, 1972b; Mixon, 1971, 1977). Mixon's (1972, 1976) studies are particularly notable in that as well as exploring the utility of role-playing studies of obedience on purely ethical grounds, he also explored methodological issues, such as those identified by Orne and Holland (1968) around demand characteristics. In some variations, Mixon framed the experiment in such a way as to attenuate the implication that the learner would not be harmed, and in these conditions 'obedience' was reduced substantially. In contrast, when participants were given reason to believe that the learner would not be harmed, 'obedience' increased. Essentially, therefore, Mixon inverts the usual approach to role-playing studies – rather than seeing the role-playing studies as showing that people behave *as if* in a real situation, Mixon shows that in the 'real' situation (i.e. Milgram's original experiments), people may well be behaving as if in a role-playing

situation. If they believe no harm will come to the learner they obey, whereas they disobey when they believe the learner is at risk.

In Miller's (1986) summary of the objections to the role-playing approach, three key concerns can be identified: First, and perhaps most obviously, any role-playing simulation, however involved the participants may appear to be, is necessarily based on hypothetical behaviour; second, the case for role-playing depends upon a comparison with Milgram's original deceptive experiments; third, the similarities in observed outcomes may nevertheless obscure differences in underlying process. Perhaps understandably, role-playing has not typically been seen as a genuinely viable alternative procedure, yet it is notable that the spirit of the role-playing studies lives on in more recent attempts to explore obedience using virtual reality and similar techniques (Dambrun & Vatiné, 2010; Haslam, Reicher & Millard, 2015a; Slater et al., 2006; see Chapter 2).

Administrative Obedience

A final set of extensions worth noting are the studies conducted by Meeus and Raaijmakers (1986, 1987, 1995). These are notable in that they constitute an early example of an attempt to develop a programme of research on obedience that modifies Milgram's procedure to make it less stressful for participants, and therefore attempts to provide an ethically acceptable paradigm for studying obedience in the laboratory. Meeus and Raaijmakers's studies were by no means the first such attempt – Ring et al.'s (1970) study, for example, had used noise blasts instead of electric shocks – but it represents the most well-developed and, to the extent that it is still being used by some researchers today (see Chapter 2), the most influential.

Meeus and Raaijmakers (1986, 1987, 1995) conducted what is almost certainly the most extensive programme of experimental research on obedience other than Milgram's itself. Noting that obedience pressures in modern societies are more likely to occur in administrative contexts than in settings requiring the meting out of physical violence, Meeus and Raaijmakers's paradigm was modelled on Milgram's procedure but required naïve participants to provide negative feedback to someone who they believed was completing an assessment as part of a selection process for a job. The feedback consisted of verbal remarks that participants were instructed to say to the job applicant. The applicant became increasingly agitated as a result of receiving this feedback, and if followed completely, the procedure culminated in the applicant failing the assessment, which would mean failing to get the job. The experimenter was

armed with a similar set of prods as was Milgram's experimenter, which were used whenever participants hesitated or refused to go on.

Meeus and Raaijmakers ran two baseline conditions in which they found 83 per cent and 91 per cent obedience. These findings confirmed their hypothesis that obedience rates would be higher using their procedure than Milgram had found owing to the absence of physical violence in their experiments. Meeus and Raaijmakers conducted several experimental variations on their baseline condition. Two of these mirrored conditions conducted by Milgram – the two peers rebel and experimenter-absent conditions – and found similar reductions in obedience levels as had been found by Milgram. Other variations were based on novel hypotheses and found that providing advance warning to participants about the nature of the experiment did not reduce obedience rates, but that obedience was reduced when participants perceived a risk to themselves – specifically, when they were informed that they were legally responsible for their actions. Meeus and Raaijamkers also conducted a series of role-playing conditions, finding that the more active the role-playing scenario (e.g. actually playing the role of the 'naïve' participant rather than responding to a description) the more similar the results were to the equivalent non–role-playing experimental conditions.

Meeus and Raaijmakers (1987, 1995) interpret their results as supporting Milgram's agentic state theory, citing responsibility data obtained using a similar 'responsibility clock' as that used by Milgram. They identify significant differences between the amount of responsibility attributed to the experimenter and to participants themselves. However, there is an important limitation in their analysis in that they did not distinguish between obedient and defiant participants. In the baseline conditions, in which obedience was over 80 per cent, this may not be too much of a problem, but in other conditions it is. Agentic state theory would predict that defiant participants should attribute greater responsibility to themselves as their resistance is the result of not having entered the agentic state. However, if we look at Meeus and Raaijmakers's experimental conditions in which defiance was the predominant response, we see a similar pattern of results to the baseline conditions: participants attribute greater responsibility to the experimenter than to themselves. Indeed, the condition which was most effective at reducing obedience – the two peers rebel condition, in which only 16 per cent of participants were obedient – features the greatest proportion of responsibility allocated to the experimenter. This appears to imply that participants were in an agentic state *even when they disobeyed*, a finding that clearly provides a further challenge to the explanatory value of Milgram's theory.

Conclusions

Having reviewed Milgram's research programme, the ethical, methodological and theoretical debates that followed, as well as some early attempts to replicate and extend his findings, what can we conclude at this point? How might we characterise the consensus on Milgram's studies that existed before the emergence of a second-wave of critical engagement with Milgram's work in the early part of the current century? As Miller (1986) notes in his earlier summary of this work, one's conclusions on such matters are invariably tied not simply to one's assessment of the empirical evidence and the logic of the arguments involved (if, indeed, anything really is), but – owing to the wider issues raised by Milgram's studies – are intimately tied to one's own views on matters of politics, ethics and the appropriate way to do research. As will become clear in subsequent chapters, of particular relevance here is my own position on social psychological methodology. I am not, and never have been, an experimentalist, and while not denying the value of experiments as part of a broader methodological toolkit, I subscribe to the position that sees social psychology's over-reliance on experimentation as having caused important conceptual problems for the discipline. It follows necessarily from this that I simply do not see the sorts of experiments conducted by Milgram as being worth the ethical problems; in short, subjecting people to stressful experiences in the name of experimental social psychology is not, for me, justifiable given the question marks over the status of the insights to be gained from such an endeavour. Nevertheless, let me try to formulate what I take to be the broad status of Milgram's experiments (in social psychology at least) at around (roughly) the turn of the twenty-first century.

First of all, it is important to note that the basic empirical phenomenon captured by Milgram – the extent to which people go along with instructions to administer what appear to be potentially harmful electric shocks to someone they believe to be another naïve participant – is remarkably robust. To be sure, there are legitimate questions concerning the relatively small number of replications, and in several cases these replications come from sources that have not passed through standard peer-review processes. But, as Blass (1999, 2012) and Miller (1986) observe, the reliability of the phenomenon despite variations in methodological factors and culture is nevertheless notable.

The meaning of the phenomenon is, however, a different matter altogether. The methodological criticisms focussing on both internal validity (i.e. demand characteristics) and external (ecological) validity are in many respects compelling. In no small part because few

investigators – certainly not Milgram himself – have actually been able to provide robust data on how participants interpreted the experimental situation, the criticisms levelled by the likes of Orne and Holland (1968) can, at the very least, not simply be dismissed. Moreover, the matter of ecological validity – the extent to which the studies can be used to generalise to 'real world' contexts – is again an open question.

Notably, Milgram's theoretical account of his studies – based on the concept of the agentic state – is generally seen as the weakest element of his work. It is instructive here that even those scholars, such as Miller (1986) and Blass (2004), who are generally inclined to defend both the value and the ethics of Milgram's work, have raised serious questions concerning the theoretical component of his analysis.

The ethics of the studies have been defended over the years by Milgram's supporters, but the more general tendency has been for social psychology to move away from the sorts of studies that Milgram conducted. This is not so much a matter of deception per se, which is still routinely used in experimental social psychology, but more of the intense stress caused to participants in the Milgram experiments. This general movement away from ambitious but potentially stressful experimental designs has been lamented by some in the discipline (e.g. Zimbardo, 1999), whilst others have attempted to come up with ethically more acceptable ways of conducting ambitious research designs (e.g. Reicher & Haslam, 2006). As we will see in Chapter 2, many researchers have more recently sought to follow the example of Meeus and Raaijmakers (1986, 1987, 1995) in developing paradigms with which to study obedience that involve less stress for participants. However, there are still examples of experimental studies of obedience that rely on both deception and extremely stressful situations (Beauvois, Courbet & Oberlé, 2012).

So, we can perhaps say that the first generation of scholarship generated by the obedience experiments led to a consensus that (a) Milgram's empirical findings were powerful and robust; (b) there were important question marks over the meaning of these findings, and in particular general agreement that Milgram's theoretical account was inadequate and (c) that, regardless of one's own position on the ethics of Milgram's experiments, social psychologists had moved away from conducting research that relied on experimental designs that combined deception with stressful situations.

As Blass (2012) has noted, attempts to directly engage with the empirical phenomena of Milgram's experiments petered out in the mid-1980s and didn't really get going again until the mid-2000s. The combination of increased ethical restrictions on researchers, coupled with the rise of the social cognition perspective which led research attention in different

directions, undoubtedly played a key role in this hiatus. But also I suspect that, despite the heated debate around Milgram's studies, the relative lack of attention given to them for a period of (roughly) 20 years is due in part to a sense in which we thought we knew what they were about, and what the relevant debates were. A nagging sense that there probably wasn't much new to say about the experiments pervaded as a conventionalised image of them crystallised and was promulgated through textbooks and introductory psychology classes. This is not to say that no work on the obedience experiments was done at all during this period – Blass (1991, 1992, 1995, 1996a, b) published a number of articles and an edited collection (Blass, 2000) covering various aspects of the experiments, and a major special issue (Miller, Collins & Brief, 1995) was devoted to the experiments. However, from around the mid-2000s new strands of work were developed that not only reopened the experimental study of obedience, but through an engagement with Milgram's archives began to derive new insights from, and to raise new questions about, the obedience experiments.