Beyond the Banality of Evil

Criminology and Genocide

AUGUSTINE BRANNIGAN

OXFORD
UNIVERSITY PRESS

Contents

List of Abbreviations

ANC	African National Congress
AR	African Rights
ASP	Assembly of States Parties
CDR	Coalition for the Defence of the Republic (Rwanda)
DPKO	Department of Peacekeeping Operations (UN)
DRC	Democratic Republic of Congo
ECCC	Extraordinary Chambers in the Courts of Cambodia
ECOWAS	Economic Community of West Africa
FAR	armed forces of Rwanda
FMLN	Farabundi Marti National Liberation Front
IAC	International Auschwitz Committee
ICC	International Criminal Court
ICISS	International Commission on Intervention and State Sovereignty
ICJ	International Court of Justice
ICTR	International Criminal Tribunal for Rwanda
ICTY	International Criminal Tribunal for the Former Yugoslavia
IFP	Inkata Freedom Party
IMTFE	International Military Tribunal for the Far East
R2P	Responsibility to Protect
RPF	Rwandan Patriotic Front
RUF	Revolutionary United Front (Sierra Leone)
SATRC	South African Truth and Reconciliation Commission
SCSL	Special Court for Sierra Leone
SPSC	Special Panels for Serious Crimes (East Timor)
TFV	Trust Fund for Victims
TRC	Truth and Reconciliation Commission
UNHCR	UN High Commission for Refugees
UNPROFOR	United Nations Protection Force
UNSC	United Nations Security Council

URNG   Revolutionary National Unity of Guatemala
VPRS   Victims' Participation and Reparation Section
VRS   the army of the Republika Srpska
VWU   Victims and Witnesses Unit

1

Genocide and the Obedience Paradigm

Introduction: From the Holocaust to Genocide

My inquiry into genocide began in the 1990s with the appearance of new scholarship on the Holocaust. Two books in particular caught my attention. Christopher Browning's *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland* (1992) and Daniel Goldhagen's *Hitler's Willing Executioners: Ordinary Germans and the Holocaust* (1996) were based on research into the role of police battalions in genocide in Eastern Europe beginning in 1941. Browning begins his account with a story about the first mass shooting of civilians assigned to Police Battalion 101 in Józefów, Poland. The men were told that they had received orders from the highest authorities to enter the village and to remove the Jews completely. All the old men, the infirm, all the women and children were to be shot dead. The able-bodied men were to be arrested for slave labour and shipped to a concentration camp in Lublin. But Major Trapp tempered his orders with this consideration: if the men did not feel they were up for the assignment, they could step out without recrimination. Some 10–20 per cent did so. What that implied to Browning was that the men who *did* participate were not coerced, or forced, to do so. Goldhagen, who recounts the same incident, goes so far as to say that they participated because they thought it was the right thing to do—that the Jews deserved to die. These two studies reopened the issue of agency in the ordinary soldier's cooperation in mass murder.

The ensuing 'Goldhagen debates' generated an enormous outpouring of historical reconsideration (Shandley 1998). While this academic conversation was underway, two new genocides occurred. In 1994 in Rwanda, over 550 000 civilians were massacred by the national army, political militias, and peasants recruited to genocide by elements of the Rwandan government. It occurred with lightning

speed, and received indifferent media coverage in Europe and America. A year later, in Srebrenica, Bosnian Serbs reportedly murdered an estimated 7 000 to 8 000 Bosniac men and boys. News stories circulated of young Serb soldiers making digital films on their cell phones of these executions, and posting them on websites as mementos,[1] just as an earlier generation of German policemen took photographs of ghetto clearings in Jósefów, Łomazy, and Warsaw to circulate at home some fifty-odd years earlier (Klee, Dressen and Riess 1996; Browning 2000: 154). At the time, few criminologists grasped the enormity of these events.

In criminology, mass murder, atrocities, and genocide appear to attract little attention. Instead, there is a preoccupation with individual-level predatory behaviour. In North American criminology, there is a consensus around the utility of what have come to be known as 'control' theories of criminal behaviours (Ellis and Walsh, 1999). These are basically Benthamite theories of human nature that assume that the pursuit of pleasure is driven by natural appetites checked only by internal and external sources of pain. Hence, control theories suggest that crimes arise when the individual's attachment to the community becomes undermined, and antisocial impulses go unchecked. In the late 19th century, Durkheim (1893) diagnosed the condition of 'anomie' brought about by rapid industrialization, and the demise of institutions that effectively bonded individuals and families to the economic structure through craft guilds and traditional modes of production. For Durkheim, criminal impulses were suppressed because villainy evoked the deepest condemnations of society and integrated the population around common moral feelings.

The success of this line of thinking suggests that the direction of criminology over the past few decades has led to a mindset in which we have become almost incapable of grasping the phenomenon of state-initiated crimes, such as those associated with inter-ethnic conflict. This is ironic inasmuch as individuals were far more likely to be killed in the past century as a result of collective crimes such as aggressive war, genocide, and state-initiated mass slaughter than to die at the hands of an individual perpetrator. We keep official statistics on the latter, but not the former. We develop theories on

[1] Associated Press, 3 June 2005 on MSNBC <www.msnbc.com/id/8085091/> retrieved 19 January 2009. The video was broadcast on Serbian television in June 2005.

crime causation, and develop policies to ameliorate garden-variety crimes, while oblivious to crimes associated with political violence. Because criminology has focused on the crimes of individual perpetrators, we have few intellectual leads in explaining the most prevalent forms of killing in modern times, crimes which appear to be more, not less prevalent. Genocide and analogous behaviours are largely uncharted water for criminology. They have become orphaned in the field devoted to the study of crime. Criminology has already tackled what makes individual murder, rape, and robbery possible. Can we capture the structural and agentic processes that operate in genocide?

The social science literature on genocide has its roots in the study of the Holocaust. Polish jurist, Raphael Lemkin, coined the term 'genocide' in 1944 in his analysis of the Nazi domination of Europe. As a result of his lobbying, the United Nations adopted the 'Convention for the Prevention and Punishment of Genocide' in December of 1948 (Chalk and Jonassohn 1990). It defined genocide as:

Any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial, or religious group as such:
  A. Killing members of the group;
  B. Causing serious bodily or mental harm to members of the group;
  C. Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in part;
  D. Imposing measures intended to prevent births within the group;
  E. Forcibly transferring children of the group to another group

Several specific crimes were named: genocide, conspiracy to commit genocide, incitement to commit genocide, attempt to commit genocide, and complicity in genocide. Such crimes could be committed whether they occurred in times of war or peace. All perpetrators became liable whether they were constitutionally responsible rulers, public officials, or private individuals. Trials could be held in the jurisdiction in which the crimes occurred, or in a specifically designed international tribunal. And the contracting parties could call on the UN to initiate the prevention and suppression of genocide when they thought it was occurring.

In the years after the Second World War, the motives and the methods of the Nazis and their allies were documented in many sources, most notably in the work of Raul Hilberg (1985) and William Shirer (1960). There also arose a comparative literature seeking to evaluate the distinctiveness of the Holocaust as well as

some parallels to other 20th-century massacres, most notably the murder of 1.7 million Armenians in Turkey and theft of their wealth, a process that started in the 1890s and continued during the First World War (Hovannisian 1986; Staub 1989). This 'barbarity and vandalism,' as he called it, had preoccupied Lemkin during the 1930s. How could the Turkish state annihilate the constituent Christian community that pre-dated Islam in Anatolia with impunity? After the UN adopted Lemkin's concept of genocide, the definition was broadened substantially by academics to cover the mass killing of civilians by governments—'democide' (Rummel 1991)—and the mass slaughter of political enemies—'politicide' (Harff and Gurr 1988). In addition, there has developed a body of theoretical sociology describing the role of genocide as a function of political struggle (Kuper 1981; Fein 1984, 1993). Attempts have been made to develop typologies of the leading types of genocide, and how they differ in motives and outcomes (e.g. Chalk and Jonassohn 1990). However, in acknowledging the burgeoning field of genocide studies, it would be premature to claim that significant consensus has emerged about the 'hard facts' of genocide in this literature, and how they might figure in a coherent theory of it.

In terms of explaining genocide and democide, the most promising line of thinking is the suggestion that the liability of engaging in genocide is a function of state political development, and particularly that autocratic or authoritarian states (whether fascist or communist) have a far greater proclivity to engage in the systematic murder of unarmed civilians than more democratic states (Rummel 1994; Horowitz 2002). This is not to say that democratic states are blameless, but they are both more restrained in their ability to mobilize mass murder, and may be better able to ensure that their aggression escapes such criminal labelling. Nonetheless, this political tendency is a hard fact of the kind largely absent in the individualistic turn of modern criminology, and reinforces the need to integrate the explanations of crime with the social science literature on genocide. Ironically, the most significant contribution to social science thinking about genocide is associated with the experimental studies of Stanley Milgram in social psychology. As a social psychologist, Milgram did not focus on the state as such, but on the role of bureaucracy and its restraints on individual freedoms. He developed the 'obedience paradigm'—the notion that the Holocaust occurred as a result of the subservience of individuals to hierarchical and bureaucratic social structures. Milgram's work has grown to

mythic proportions in today's social science (Blass 2004). I review this work since it has generated more comment and speculation than any comparable contribution in the social sciences. I also outline why I think this work has become outmoded in advancing our understanding of genocide.

The Holocaust, Obedience, and the Banality of Evil

Stanley Milgram was the son of immigrant Jews, born in New York, and raised in a social environment starkly cognizant of the Holocaust. Like many of his generation, he was deeply troubled by German anti-Semitism. In his research as a doctoral student in psychology, he researched national differences in conformity since this promised to shed light on German mistreatment of the Jewish minority. When Adolph Eichmann was seized by Israeli agents in Argentina, Milgram was already exploring ways to investigate obedience in a psychological setting. The evolution of Milgram's design of the obedience study has been researched through archival materials by N.J.C. Russell (2009, 2011) and Gina Perry (2012).

Milgram closely followed the five *New Yorker* reports filed by Hannah Arendt who covered the trial of Adolph Eichmann in Jerusalem in 1961, and which were the basis of her 1963 book. Adolph Eichmann was captured in Argentina on 11 May 1960, and returned secretly to Jerusalem by the Israeli secret police for trial as a war criminal several days later. He was given the option of being assassinated in Argentina, or standing trial in Jerusalem, and chose the latter. Before capture, he had already recorded dozens of hours of audio tapes with Willem Sassen, a low-ranking SS officer on the run, and produced thousands of unpublished pages of transcripts describing the 'German side' of the story. Eichmann was the Nazi 'bureaucrat' who helped orchestrate the mass murder of European Jewry by concentrating the victims in Poland after the Nazi conquest of Poland, France, and most of Western Europe. He advanced to a senior position in the Department of Jewish Affairs, and played a pivotal role in deporting entire Jewish communities to the factories designed for their large-scale extermination at Auschwitz, Belzec, Chelmo, Majdanek, Sobibor, and Treblinka. Over five million innocent people, men, women, and children, were murdered at these death camps with the assistance of ordinary German administrators, policemen, soldiers, and camp guards. In respect of Auschwitz alone, Rudolph Höss signed a confession that

acknowledged that 'at least 2.5 million victims were executed and exterminated there by gassing and burning, and at least another half million succumbed to starvation and disease making a total dead of about 3 million' (quoted in Ramler 2008: 68–71). In Arendt's view, Eichmann was not a psychopathic killer, nor had he exhibited deep animosity towards the Jews. He followed orders with zeal and without any evidence of inhibitions of conscience. Milgram appears to have accepted Arendt's diagnosis of 'the banality of evil'—the idea that Eichmann participated in mass murder simply as an obedient cog in a state bureaucracy.

Milgram developed a laboratory protocol that attempted to capture the essence of behaviour at the core of genocide: individual obedience to malevolent authority. This is the subject of the current chapter. Subjects were recruited for a study of learning. They were informed that the experimenter was testing the effects of punishment on learning. The subjects saw 'The Learner' strapped into a device in which he received electrical shocks as a punishment for failure to learn a series of specific word pairs. All this was simulated, and no one was actually shocked. Milgram's question: at what point would the subject refuse to comply with demands to administer the shocks? The issue of social conformity had been suggested by Solomon Asch's study of 'the line judgment task' in which subjects were pressured by a group to accept a conclusion they knew to be untrue. Subjects were asked to match a stimulus line to one of three choices where two out of three were clearly erroneous, but were chosen nonetheless by the majority. Asch's subjects were more likely to capitulate based on the size of the group, and less likely to capitulate when one other subject resisted group pressure. As a graduate student, Milgram had been assigned to Asch as a research assistant, and became intimate with his famous work. Indeed, Milgram used Asch's protocol to examine national differences in social pressure. Milgram's obedience protocol differed inasmuch as the pressure to comply with such injurious demands came from the orders of an authority figure, a lab-coated 'Scientist'. The experiments attracted enormous attention in the academy and in society at large (Miller 1986; Miller et al. 1999). They were extremely controversial for both empirical and ethical reasons (Orne and Holland 1968; Mixon 1971, 1989; Baumrind 1964; Patten 1977a, 1977b).

In Milgram's experiment, ordinary subjects were made to play the role of aggressive teachers, of people who acted violently against

innocent victims. In what Stam, Radtke, and Lubek (1998) refer to as the 'received view' of this work, Milgram took people from all walks of life and revealed their incipient capacity for evil. Following the Eichmann trial, Milgram suggested that the capacity for evil was fostered in individuals by bureaucratic authorities. The study was advertised as an experiment designed to test the effectiveness of punishment on human learning. The 'Teachers' were paid to teach the 'Learner' to memorize a long series of word pairs. The experimenter explained the rationale for the study: to determine the effectiveness of *negative* reinforcements on learning. The Learner's errors were to result in an electric shock. Each successive error resulted in a small increase in the level of shock. The shocks began at 15 volts and escalated up to 450 volts. The experiment advertised for both teachers and learners, but all the subjects were assigned the role of 'Teachers'. The Teachers were given a sample shock to demonstrate the discomfort that resulted from the device used to discourage errors. The machine was an impressive electrical appliance with switches, lights, and verbal designations describing the severity of the shock (i.e. mild, moderate, high, extremely high, XXX). Over 700 subjects were drawn from a wide range of occupations and professions to participate.

Milgram hired John Williams to play the role of the Scientist, and James McDonough to play the role of Learner. Williams's task was to encourage the Teachers to comply with demands to administer increasingly severe levels of shock. The assignment was designed so that the Learner's performance always failed, and attracted increasingly severe (but simulated) levels of punishment. Many subjects experienced tremendous anxiety. Unlike the classical experimental approach, Milgram did not specify specific hypotheses a priori. He did not begin by testing the validity of any particular theory of behaviour or hypothesis. There are no references to earlier studies of group influence (i.e. Asch and Sherif) in his articles. Milgram proceeds as though his work was generated without influence from the earlier studies. Milgram consulted many groups to determine how they thought normal individuals would react to the situation, and how many would refuse to take part. Everyone predicted that all of the subjects would defy the authority figure and refuse to administer severe shocks.

The Blackwell Reader in Social Psychology summarized the study as follows: 'there is no experimental design as such; no factors are manipulated. No statistics are reported on the data nor are

they needed since no experimental variations were compared'
(Hewstone, Manstead, and Stroebe 1997: 54). This does not give
Milgram his due. Milgram studied various conditions of aggres-
sion, the best known of which was proximity. He argued that the
closer the victim to the context of aggression, the lower the levels of
compliance. Some argue that this was his most salient discovery
(Russell and Gregory 2005). He also tested other effects. In fact, in
his 1974 book, he reports eighteen different conditions of obedi-
ence, although he had completed twenty-three (Perry 2012). In his
baseline study, Milgram found that the majority of subjects *did*
administer the maximum level of shock (65 per cent), and that this
did not decline even when the Learner reportedly suffered from a
cardiac irregularity. He concluded that compliance of individuals
in these conditions resulted from the force of authorities over their
subordinates. His experiment extracted this general human ten-
dency from the reports of the Holocaust killers who reported, ini-
tially at Nuremberg and later in Jerusalem, that their role in mass
murder was a result of 'following orders'. That has been the para-
digmatic view of the obedience studies over the last five decades.

Criticisms were raised both in terms of internal and external
validity. Internal validity depends on whether the protocol employed
by the experimenter actually succeeds in defining the situation for
the subject as intended; external validity depends on whether the
protocol corresponds credibly to features of everyday life to which
the experiment might be generalized. As for internal validity, con-
trary to the paradigmatic view, Orne and Holland (1968), Mixon
(1971), Darley (1995) and other critics argued that, in psychology
experiments, subjects presume that no one will actually get injured.
In this study, Milgram assumed that subjects would define the
administration of shocks as tantamount to assault or cruelty. How-
ever, in the pretests of the study, Milgram reported 'in the absence
of protests from the learner, every subject in the pilot study went
blithely to the end of the board' (1974: 22). Every subject in the
pretest administered the maximum shock level without pressure
from anyone. Presumably, subjects did not assume the worst about
administering electrical shocks. It was only at this point that
Milgram introduced the various feedback conditions, initially a
knock on the wall, to indicate that the Learner receiving the shocks
was actually experiencing discomfort. In the *Obedience* film, it
is evident that when the Learner-actor exhibits pain by actually
calling out loud, the real subjects *initially* laugh, and appear to be

startled that anyone is actually being hurt. In the later designs, when the subjects hear similar complaints from the Learner testifying to the painfulness of the shocks, they also have in their presence the 'authority figure', the Scientist, who contradicts their perceptions that something is going wrong, and who reacts passively as people are audibly suffering. The experimental design is ambiguous. The subject is drawn between what is heard—a suffering victim—and what is seen—a non-plussed authority figure subject to the same information, but not alarmed by it. This causes enormous conflict for the subjects. Subjects frequently sweated, stuttered, and trembled. They may have started with an assumption that nothing-can-go-wrong only to have this contradicted by what they could hear from the Learner, but not by what they could see from the scientific authority. The design of internal validity is questionable since the subjects are exposed to *conflicting* information. As Orne and Holland (1968: 287) noted, the most incongruous aspect of the experiment was the behaviour of the experimenter who sat by indifferently when the Learner called out in agony, and demanded to be released. Orne and Holland concluded that subjects must have inferred that the harm being experienced through the shock administration was not what it appeared to be, just as the audience at a magic performance knows that the magician's assistant is not being cut in half with a saw, and cannot be suspended in thin air without support.

The credibility of the experiment in terms of external validity may have been further eroded by the fact that the role of the Teacher was actually superfluous in the experiment since the teaching could obviously be carried out without volunteers. In the same vein, it could not have escaped notice by all the subjects that the learning task was simply impossible. Mantel's analysis of the external validity of the experiment was highly critical:

Every experiment was basically preposterous . . . the entire experimental procedure from beginning to end could make no sense at all, even to the laymen. A person is strapped to a chair and immobilized and is explicitly told he is going to be exposed to extremely painful electric shocks . . . The task the student is to learn is evidently impossible. He can't learn it in such a short space of time . . . No one could learn it . . . This experiment becomes more incredulous and senseless the further it is carried (Mantel 1971:110–11).

In a similar vein, Baumrind (1985: 171) noted that 'far from illuminating real life, as he claimed, Milgram in fact appeared to have

constructed a set of conditions so internally inconsistent that they could not occur in real life. His application of his results to destructive obedience in military settings or Nazi Germany . . . is metaphoric rather than scientific.'

Don Mixon suggests that every experimental manipulation that Milgram developed which introduced less ambiguous evidence that a subject was being hurt reduced the aggression of the Teacher. When the Learner's pain was signalled through knocking on the wall, compliance dropped from 100 per cent to 65 per cent. The slightest evidence that harm was occurring produced the largest measure of defiance that Milgram measured. All the elaborate verbal feedback of the Learner's suffering, which formed the 'baseline' measurement, reduced the compliance by only a further 2.5 per cent over the knock on the wall; only one less person in forty resisted going to the highest shock level.

Russell (2011: 153) writes that 'where I have probably added most to the literature was, first, in revealing Milgram's yet, in his publications, unmentioned goal to maximize the first official experiment's completion rate, and second how he set about achieving this goal'. The completion rate refers to the percentage complying with the highest level of aggression described in the first publication. It was achieved by the ad hoc introduction of tweaks designed to bind the subjects to the assignment, and to reduce their stress in complying during the trial runs. The 'binding factors' included the presentation of the task as a legitimate university pursuit, represented by a mature Scientist, in a task that created foot-in-the-door 'momentum' through numerous incremental steps in shock administration. Several 'stress reduction mechanisms' were introduced to offset the subjects' instinctual resistance: the subjects were advised that the procedure produced no permanent damage, it was presented as a legitimate learning experiment, the 'lethal' label on the ultimate switch was replaced with 'extreme', and the experimenter assumed responsibility for any adverse effects. Under these protocols, Milgram produced his best-known result (Russell 2011: 160). The received view focuses only on the authority figure, and ignores the framing that generated the provocative 65 per cent outcome.

Reicher and Haslam (2011: 166) reject the received view and suggest that the levels of compliance 'depend upon participant's exposure to the voices of different constituents'. When the victim's suffering was brought into the room and portrayed dramatically by an actor in a real subject's presence, although the authority figure's

comportment suggested no harm, the aggression declined. When the authority figure was totally removed from the lab and apparently out of earshot, the pain feedback information from the Learner reduced the shocks from 65 per cent to about 20 per cent. The more evident the painfulness of the procedure to the Learner, the lower the levels of obedience (Brannigan 2004).

Milgram investigated this issue in a post hoc questionnaire completed by 658 former subjects a year after the study. He asked subjects whether they thought anyone was actually harmed—a good check on internal validity (1974: 172). Fifty-six per cent suggested that they fully believed the Learner was receiving painful shocks; 24 per cent thought the Learner was *probably* getting the shocks. So far, so good. However, this belief was *not* spread evenly across the obedient and defiant groups. Those who were convinced that the shocks were real were more likely to be *defiant* of authority. I have re-analysed Milgram's table to highlight the link between the perceptions of harm and the resulting defiance of authority. For the sake of simplicity, in Table 1.1, I omit the category in which subjects reported 'I just wasn't sure whether the learner was getting the shocks or not'. This permits us to dichotomize the subject responses. The omitted category comprised 6.1 per cent (n = 40) of the total pool of respondents. The following table reports the numbers of persons falling into each of the other categories. For the sake of simplicity, I analyse the view of the *believers* versus the *sceptics* by summing the first and second rows and contrasting them to the third and fourth rows.

Table 1.1 Defiance gauged by perception of harm

Subjects' reported perception during the experiment measured afterwards	Subject was defiant	Subject was obedient
1. 'I fully believed the learner was getting painful shocks.'	230	139
2. 'Although I had some doubts, I believed the learner was *probably* getting the shocks.'	83	75
3. 'Although I had some doubts, I thought the learner was probably not getting the shocks.'	28	47
4. 'I was certain the learner was not getting the shocks.'	5	11

The benefit of this procedure is that it permits us to calculate the *odds* of being defiant of authority based on the perception that the shocks were believed to be real (n=230) or were *probably* real (83), versus were believed to be unreal (n=5) or were *probably* unreal (n=28). The odds are calculated as a*c/b*d.[2] If a subject thought the shocks were real or were probably real, this increased their likelihood of defiance 2.57 times. This conclusion is consistent with an internal report written by Taketo Murata for Milgram in 1962. Murata hypothesized that those who reported 'fully believing that the Learner was being shocked would not reach as high shock levels as those not fully believing. This is found to be so' (Murata 1962; Perry 2012). In eighteen out of twenty-three conditions, the believers had lower shock means than sceptics. On this basis, it is quite clear that the experimental protocol was far from being internally valid. Not only did a significant portion of the subjects fail to accept the experimenter's definition of the situation, but also when the subjects *did* accept the harm definition, they tended to be *defiant* of authority.

Milgram and the agentic shift

How did Milgram explain the behaviours he observed in the lab? As noted earlier, he did not begin with a theory and design an experiment to test it. He tested various levels of Learner feedback (distal, proximal), the role of group mediation of response, the role of gender, location, and Teacher-choice of shock levels, and discovered enormous variation in compliance. However, when he summarizes his work, Milgram largely ignores all these conditions in the variability of obedience. He focuses exclusively on the power ascribed to the authority figure and his ability to extract obedience from the subjects. In his explanatory chapters, Milgram proposes that obedience appears to have a biological basis, that it probably confers fitness during evolutionary competition by making coordinated action more effective than the sum of individual actions. Echoing

---

[2] a=230 + 83 (313), b=139 + 75 (214), c=28 + 5 (33), d= 47 + 11 (58). a x c/b x d = 313*58/214*33 = 18,154/7,062 or 2.57. Fisher test p = .000. If we calculate the OR for the 2 extreme groups that accepted the harm or denied it completely, we arrive at the following estimate: a x c/b x d = 230*11/139*5 = 3.64. Hence, among those with the most certain views, the perception of harm increased the likelihood of defiance by 3.64 times.

Hobbes, Milgram writes: 'a curb must be placed on the unregulated appetites, for unless this is done, mutual destruction . . . will result' (1974: 127). In the first instance 'the presence of conscience in men' inhibits destructive competition among those who 'occupy a common territory'. Conscience makes individuals self-regulating, and inhibits mutual exploitation. However, in hierarchical social organizations, the individual conscience is not sufficient and may be anarchical. Here another process comes into play at a higher level: 'the psychology of the ultimate leader demands a different set of explanatory principles' (1974: 130). For this different set of principles, Milgram turns to Freud. Freud (1922: 78) had analysed the psychodynamics of authority in the army, the Church, and the family, and the common patterns of submission to leadership. Freud explained it this way: 'the individual gives up his ego ideal and substitutes for it the group ideal embodied in the leader'. For Freud, subordinates in social hierarchies comply with the demands of leadership because of anxiety associated with the Oedipal desire to challenge the leader. Anxiety is based on *Thanatos*, the destructive instinct, which is directed against the self. It fosters a deep sense of moral obligation to comply, and a sense of dread in defiance, and substitutes the group's ideals for the individual's. Milgram's mechanism is quite different. How is it that otherwise decent and conscientious individuals act so horribly against the Learner in the lab? They do so, writes Milgram, 'because conscience, which regulates impulsive aggressive action, is per force diminished at the point of entering a hierarchical structure' (1974: 132). The individual's moral compass changes when he or she enters a group, and conscience appears to take a holiday when it joins a hierarchy.

By what mechanism does this happen? Conscience appears to undergo 'an agentic shift'. Milgram notes that 'the state of agency is the keystone of our analysis' (p. 133). Something magical or transformative happens when ego enters into a pattern of hierarchical social action—ego moves from an autonomous mode of self-direction to an agentic mode unencumbered by individual conscience. Milgram indicates that this is probably associated with changes in patterns of neural functioning. And while it was then difficult to identify neurochemical changes with certainty, Milgram nonetheless asserted that the 'chemical inhibitors and disinhibitors alter the probability of certain neural pathways and sequences being used'. In addition, he said there is compelling phenomenological evidence of such a shift reflected in 'an alteration of attitude'. When one joins a hierarchy, one sees

one's self as 'an agent for executing the wishes of another person' as opposed to acting on one's own agenda. Given the presence of 'certain critical releasers…the shift is not freely reversible' (p. 134). While this capacity for the agentic shift may have biological origins, it is reinforced throughout the life cycle as individuals move from families, to schools, to employment. Family socialization is premised on child obedience to parents. Education is premised on discipline and compliance to teachers. Jobs come with expectations defined by employers. While social theory applauds the effective bonds between children, their families, schools, and communities, Milgram sees this as a liability. 'The very genesis of our moral ideals is inseparable from the inculcation of an obedient attitude' (1974: 136). If the family, school, and work inculcate an obedient mindset throughout the life cycle, one might ask at what point does the autonomous conscience appear? Despite entering a relationship with an authority figure, many of Milgram's subjects did not experience an agentic shift. Neither Milgram's nor Freud's analysis reflects the findings in the lab: in the received view, subjects appear to be mortified because they fear that someone innocent may have suffered at their hands.

Milgram's theory of the agentic shift emerged years after the conclusion of his experimental work, and it was never itself tested experimentally. Nonetheless, his work still appears to retain relevance in contemporary studies of genocide. In his chapter on 'Ordinary Men', Christopher Browning refers to Milgram at some length (1998: 171–5). While generally sympathetic to Milgram's approach, Browning suggested that, in contrast to the experiment, the 'authority figure' in his analysis, Major Trapp, was actually a rather weak leader, though much loved by his men, and that their participation in mass executions appears to have arisen more from duty and interpersonal loyalty than blind obedience.

There is also significant 'gerrymandering' in the moral assessments associated with Milgram's account. I raise this because it points to a major problem of external validity. If Milgram knew during the course of his experiment that subjects were being hurt, that is to say, were being emotionally traumatized, why did he not terminate the experiments immediately? Milgram (1963: 375) noted:

many subjects showed signs of nervousness in the experimental situation, and especially upon administering the more powerful shocks. In a large number of cases the degree of tension reached extremes that are rarely seen in sociopsychological laboratory studies. Subjects were observed to sweat,

tremble, stutter, bite their lips, groan and dig their fingernails into their flesh. These were characteristic rather than exceptional responses to the experiment...One observer related: 'I observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, who was rapidly approaching a point of nervous collapse'.

If accurate, why did Milgram not terminate the study once he had observed such trauma? It appears he thought science and society might benefit from it in the long run. However, in characterizing the conduct of his teachers as acting in what he described as a 'shockingly immoral way', Milgram overlooks the fact that the subjects might be entitled to the same excuse since they were encouraged to administer electric shocks to advance human knowledge about the effectiveness of punishment. They were also assured that the shocks would not result in any lasting harm. If acting to advance science, would the subjects characterize their conduct as 'immoral aggression' (bad) or 'reinforcement' (good)? Milgram seems to be keeping two sets of books. In one set, he describes the task to subjects as a legitimate exercise, then, in a second, characterizes it as immoral. Abse suggests that if one wants to view the subjects as so many Eichmanns, then 'the experimenter had to act the part, to some extent, of a Himmler' (1973: 29).

There is a further moral ambiguity in respect of Milgram's depictions of authority. We see this in his concluding paragraph to the 1974 monograph where his language mystifies the moral standing of authorities associated with collective violence. He refers to 'the character' of the kind created in modern societies (mentioning America in particular) and its inability to 'insulate its citizens from brutality and inhumane treatment at the direction of *malevolent* authority' (1974: 199, emphasis added). Then he says that a substantial part of the population will act badly 'without limitations of conscience, so long as they perceive that the command comes from a *legitimate* authority' (emphasis added). Within the same paragraph, Milgram conflates malevolence and legitimacy at the highest level of the state. But surely it makes a difference whether the leadership is legitimate or criminal, since political defiance may turn on this perception. Did Eichmann view the rule of the Third Reich as illegal, or did he act with a 'clear conscience', as Erber (2002) suggests?

Duress, duty, and the obedience paradigm

Hannah Arendt focuses on a point that Milgram, as a psychologist, appears to miss. The orders for aggressive war and the special treatment of subject populations were undertaken within the rule of law in Nazi Germany and its conquered territories. Eichmann, as well as the defendants at Nuremberg, invoked obedience to orders. Their behaviour was lawful within the state structure inasmuch as the orders for deportation and extermination came from the head of state or the sovereign, and the sovereign historically has been immune from prosecution by other states for politically sanctioned activities within the sovereign's jurisdiction, save for crimes covered by international conventions. That impunity would have extended to persons acting under delegated authority, such as Eichmann. The 1948 UN Convention on the Prevention and Punishment of Genocide altered this doctrine fundamentally, but it was *ex post facto* law, and did not apply to Eichmann.

This raises a question about the entire way in which Milgram approached the Eichmann case. When Eichmann and other Nazis offered the defence of 'following orders', the obedience paradigm confused *duress* and *duty*. Milgram appears to have associated the concept of 'orders' with the idea that officers and enlisted men who followed them were acting under duress. Where Milgram creates a situation of enormous emotional conflict, and where the authority figure attempts to coerce the subjects, Eichmann's case attracted interest for the opposite reason—because he followed orders with zeal, not because of fear of a superordinate power, or because of coercion. In addition, there was never any evidence of self-doubt, contrition, shame, embarrassment, remorse, or mortification on Eichmann's part. He does not denounce his past, prostrate himself, apologize profusely, and seek forgiveness. He was not a 'desk murderer' distantly removed from the mass killings and disinterested in the fate of the Jewish victims. He was the project's most enthusiastic supporter. Information released in 2011 from tapes recorded in Argentina prior to his arrest corroborates this: 'I was no ordinary recipient of orders. If I had been, I would have been a fool. Instead, I was part of the thought process. I was an idealist' (quoted in both Aderet 2011 and Spiegel 2011).

I noted briefly Browning's references to Milgram. In his analysis of the Order Police, Daniel Goldhagen also dealt at some length with this work. He reports a series of misconceptions about the

Holocaust that have influenced our understanding of how it occurred—again associated with the idea of duress and coercion. First, there was 'a widespread conviction' that any German soldier who refused to participate in the killings would have been killed himself or severely punished (1997:10); and second, that the perpetrators were merely 'blind followers of orders'. According to Goldhagen, the evidence suggested otherwise. Goldhagen also challenged the idea that the Germans were subject to 'tremendous social psychological pressure' arising from 'the institutional roles that individuals occupy' (1997: 12); and that 'the callous disregard for the victims' was a result of the large, impersonal bureaucracy that undermined any personal responsibilities for the killings. These factors appear to suggest that the perpetrators were 'beings moved solely by external forces or by trans-historical and invariant propensities' (1997: 13), a characterization at the root of Arendt's *banality of evil*. For Goldhagen, it is important that the students of the Holocaust appreciate the motives and self-understandings of the perpetrators, the fact that they were Germans, and that the object of their fury was the Jewish community. Milgram takes an historical observation and reduces it to a species-trait. Arendt similarly glosses Eichmann's development as an enthusiastic Nazi and reduces it to totalitarianism (Cesarani 2006). Lipstadt (2011: 163ff) shares Cesarani's suspicions about how Arendt's political ideology led her to overlook the agency of those who endorsed the Nazi philosophy with zeal.

If we follow Goldhagen, Milgram's perspective turns things on their head. Milgram's depiction of the Holocaust transfers our focus away from the real victims by dwelling on the murderers, as though they were the victims of *their* bureaucracies, and reifying their alibi of 'following orders' as though this entailed coercion. Milgram's conceptualization seems to depict the Germans as unwilling executioners, as victims of totalitarianism. In transporting these issues to the laboratory, Milgram's design is based on the supposition that the Teacher's aggression is not only illegitimate, but is *seen* to be illegitimate by the subjects, by implication, suggesting that ordinary Germans participated in genocide involuntarily. This makes our adherence to the Milgram paradigm impossible. When accused Nazis invoked the defence of state orders, they were raising the positive *duty* that empowered them through the command structure to do what they did, that the orders did not originate from them, nor were they in a position to openly subvert them (Osiel 1999). They

had an opportunity to be reassigned, as in Police Battalion 101. Likewise, Himmler told his senior generals on the Eastern Front that if they were unable to follow orders, they could resign and collect their pensions (Browning 1992: 74–5). Goering's statement at Nuremberg was similarly telling. 'We had orders to obey the head of state. We weren't a bunch of criminals meeting in the woods in the dead of night to plan mass murders' (quoted in Ramler 2008: 60). The Final Solution to the Jewish Problem was proposed by Hitler's most senior advisors, Heydrich and Himmler, probably following his private suggestions, and instituted at the Wannsee Conference in Berlin in the early winter of 1942 by the most senior bureaucrats of the German state. After their defeat, the defence of 'following orders' may have been heard as duress, but this conceals the positive agency that came from duty, and the zeal with which the officers and soldiers furthered the political goals of the Reich. To the extent that Milgram and Arendt filtered our understanding of the Holocaust, we require a new approach that is more faithful to the original events.

Replicating Milgram

J.M. Burger (2009) replicated Milgram's work, at least partially. His work was based on a revised protocol in which the Learner reports medical problems with his heart, and the Teacher receives remote voice feedback from shocks appearing to originate in a separate room. Given the grave worries over the potential traumatization of subjects caused in part by Milgram's original work, Burger limited the maximum shock level to 150 volts. In Milgram's original study, 79 per cent of persons who went *beyond* this level of shock showed total obedience. This was also the point at which the Learner first expressed serious complaints, and loudly demanded to be released from the study. Burger measured whether the subjects *tried* to continue after the 150-shock level; all subjects who had not desisted at this point were prevented from continuing. Hence, the experimenter avoided the prolonged pressure on the subjects to comply at higher shock levels, while getting a measure of aggression that correlated with the original conditions and findings. There were 70 subjects (29 men and 41 women) aged from 20 to 81 years (median = 41). Burger eliminated subjects who had prior knowledge of the Milgram paradigm and/or who seemed prone to unpleasant reactions. Subjects were paid $50 for attending two

forty-five-minute sessions at the university, and were debriefed immediately following the completion of their individual tests.

Burger also administered a series of psychometric tests before the experiment. These picked up such tendencies as empathetic concern for the plight of unfortunates; prior levels of anxiety; desires of respondents to master control of their lives; and subclinical symptoms of depression. Burger also explored a protocol in which a second Teacher (a confederate) refused to continue after hearing Learner complaints following the 90-volt switch level. Thirty of the subjects were tested in this variation, each with a 'rebel' of the same gender. The real subjects were asked to replace the first Teacher. Unlike the original experiment in which the real subjects had already administered multiple shocks before two rebels broke off, in this design there was no practice effect before the real subject began to administer shocks. Burger hypothesized that the modelled refusal condition would reduce obedience in the real subjects.

What were the results? Burger reports that 28 out of 40 (70 per cent) attempted to continue after 150 volts. Milgram had found that 32 out of 40 (82.5 per cent) did likewise but, due to differences in sample sizes, these results are not statistically different. Secondly, he failed to find defiance in the model refusal group, contrary to Milgram. Third, there were no gender effects and no effects associated with individual differences based on the psychometric tests. As for whether the Teachers were convinced that the Learners were getting the shocks, Burger did not conduct the same sort of post hoc survey used by Milgram. He reported (personal communication, March 2011) that after personally debriefing every subject, he was of the opinion that not a single person was in doubt about the veracity of the shocks. This was a much higher level of conviction than Milgram reported. His subjects were told explicitly that at any time they could withdraw from the study, as could the Learner. 'Several of the participants who stopped the procedure after hearing the learner's protests pointed out that the confederate had been promised he could stop when he wanted to' (2009: 9). Since the Learner did not, this could be taken as an indication that the shocks were not as bad as they seemed.

The most interesting finding from Burger's replication was reported in a second paper in which he and colleagues analysed subject responses to the prods from the Scientist. Burger argues that Milgram was not really studying obedience to orders at all. In the original study, if a Teacher hesitated after resistance from the Learner, the

Scientist used four escalating prods to get him or her to continue: 'please continue', 'the experiment requires that you continue', 'it is absolutely essential that you continue', and 'you have no other choice, you must go on'. Only the last prod looks anything like an order. 'When participants heard the only prod that we might reasonably consider an order, not one individual "obeyed"' (Burger, Girgis, and Manning 2011). Indeed, the evidence shows that compliance declined with each level of escalation. Burger et al. concluded by noting that alternative interpretations to Milgram's work should be explored and 'the way the research is portrayed to students, scholars, and the public may need to be reassessed' (2011: 6). Burger et al.'s results about the ineffectiveness of direct orders are corroborated by Stephen Gibson's analysis of the audiotapes from two of Milgram's original experimental sets. He examined the exchange between subjects and John Williams in each case when prod 4 was employed. 'The first, and perhaps most striking, observation to make is just how ineffective prod 4 appears to have been' (Gibson 2011: 301). It was used on twenty-three occasions but in only one case was it followed by full obedience. Typically, it resulted in the experimenter acknowledging that the subject indeed had a choice!

Why Milgram fails on the question of genocide: beyond the banality of evil

I believe Milgram will always enjoy an important place in social science simply for bringing the issue of genocide and mass murder so vividly to the attention of scientists and society. His use of the electrical shock device, his casting of the innocent, middle-aged Learner, Mr Wallace, and the grim-faced lab-coated Scientist, Mr Williams, have etched themselves into the memories of successive generations of professors and their students. His work has become something of a touchstone for anyone researching genocide. However, major problems have been identified in his work, and ironically there has been no development of the ideas that he advanced in the 1960s, particularly at the theoretical level in the area of the agentic state. Agency remains a core issue in social sciences, and the constraints on agency leading to compliance in mass murder still haunt the study of genocide as noted in the recent histories of the police battalions in Poland.

   In this chapter I have raised methodological issues pertaining to the continuing utility of Milgram's ideas based on issues of internal

and external validity. I have argued that there exists strong evidence that in the original study, a substantial number of subjects did not actually believe that the Learner was being hurt, making it difficult to attach any significance to their behaviour, and making it difficult to generalize to conditions in real life. This view is shared by many, including Daniel Goldhagen (1997: 592) who wrote that Milgram 'discovered that the more the people who administered the shocks confronted the apparent pain of the person being shocked, the more frequently they were willing to defy the authority of the Yale University experimenter'. Even in his replication of the basic Milgram observations, Burger makes a different but equally worrisome observation: however you characterize their behaviour, the subjects in the Milgram protocol were not following orders at all. Prods that most closely approximated orders were singularly unsuccessful in achieving compliance.

This raises another issue in terms of the application of the research to everyday life. The original case that provoked Milgram (Eichmann) did not have the character of an individual bullied into submission by a bureaucracy or by an authority figure. Nor did the later attempts to fit such atrocities as the My Lai massacre in Vietnam follow the Milgram paradigm of 'just following orders'. Lieutenant Calley did not get orders from 'the highest authorities' to carry out the shooting of civilians; nor did Charlie Company's massacre resemble the routine executions carried out on the Eastern Front by the *Einsatzgruppen*. In his biography of Eichmann, Cesarani (2006: 15) says the following:

Ironically, her book, *Eichmann in Jerusalem*, more than the trial itself shaped Eichmann's legacy. Anyone writing on the subject today works in the shadow of Hannah Arendt. Her notion of 'the banality of evil', combined with Milgram's thesis on the predilection for obedience to authority, straight-jacketed research into Nazi Germany and the persecution of the Jews for two decades.

In the following chapters I sketch an alternative account of genocide, starting with some paradoxes about the nature of genocide from the perspective of criminology, followed by a genealogical account of genocide and an alternative explanation of genocidal behaviour. When we think about genocide, we have been conditioned to think of 'the banality' of evil. Were we to take the perspective of its advocates, there is nothing banal about it. Genocide is a political crime whose architects seek a complete transcendence of their historical circumstances, and pursue their version of The Good through mass atrocities.