Lack of Character

Personality and Moral Behavior

JOHN M. DORIS
University of California, Santa Cruz

CAMBRIDGE
UNIVERSITY PRESS

3

Moral Character, Moral Behavior

> The trouble with Eichmann was precisely that so many were like him, and that the many were neither perverted nor sadistic, that they were, and still are, terribly and terrifyingly normal.
>
> Hannah Arendt

Totalitarianism specializes in the dissolution of fortitude, whether by the extremes of physical torture (Bettelheim 1943) or by the psychological degradation of "thought reform" or "brainwashing" (Lifton 1956; Schein 1956). These practices are repellent, but their effects are not unexpected. Aristotle (1984: 1115b7–9) acknowledged that some things exceed human endurance, and Russell (1945: 267), with another 2,000-odd years of history to consider, remarked that the will withstands the tyrant only so long as the tyrant is unscientific. Situationism teaches something more surprising and, in a sense, more disturbing. The unsettling observation doesn't concern behavior in extremis, but behavior in situations that are rather less than extreme; the problem is not that substantial situational factors have substantial effects on what people do, but that seemingly insubstantial situational factors have substantial effects on what people do. The disproportionate impact of these "insubstantial" situational factors presses charges of empirical inadequacy against characterological moral psychology: If dispositional structures were typically so robust as familiar conceptions of character and personality lead one to believe, insubstantial factors would not so frequently have such impressive effects. In the present chapter, I'll document the evidence for this contention.

Prelude: Character and Compassion

On a March night in 1963, Catherine Genovese was stabbed to death. Her killer attacked her three times over a period of 35 minutes. Despite Genovese's clearly audible screams, 37 of 38 witnesses in her middle-class

Queens neighborhood did not so much as call the police; one, after first calling a friend for advice, notified authorities only when the attacks had ended and Genovese was mortally wounded (Rosenthal 1999). While there is room for controversy over just what compassion consists in, I suspect few would deny that complete inaction when a screaming young woman is slowly butchered nearby problematizes its attribution. As opposed to compassion the emotional syndrome, which may be quite transitory, compassion the character trait is a stable and consistent disposition to perform beneficent actions (Blum 1994: 178–80); failures to behave compassionately when doing so is appropriate and not unduly costly are evidence against attributing the trait.

The experimental and historical records reveal that such omissions, as well as similarly incompassionate actions, commonly occur where the obstacles to compassion and the pressures to incompassion seem remarkably slight: the failures are disproportionate to the pressures. In the first instance, this problematizes thinking about compassion in terms of a robust character trait. If I'm right, however, compassion exemplifies a general problem for characterological moral psychology. I'll treat compassion as a sort of test case.

In part, this strategy is opportunistic: There are quantities of empirical work on compassion-relevant behavior. I'm not merely an opportunist, however; as a core ethical concern on a variety of evaluative perspectives, compassion is a natural locus of discussion. Somewhat awkwardly for me, compassion does not appear in Aristotle's discussion of virtues, but I think it would be a mistake to suppose that he had no interest in the sort of concerns associated with compassion.[1] For example, while Aristotle's magnanimous man is decidedly not a compassionate saint, Aristotle (1984: 1123b30–4) insists such a person will not wrong others; it would be surprising if Aristotle expected him to brutalize innocents or stand by while others do so. Behaviors associated with compassion are of substantial interest for any ethical perspective that emphasizes other-regarding concern, that is, most any recognizably ethical perspective. There may be those who reject this characterization of ethics, but there's little doubt that they are in the minority.[2]

My arguments are not contingent on any particular understanding of compassion; I could as easily couch discussion in terms of what psychologists rather colorlessly call "prosocial behavior" (e.g., Bar-Tal 1976: 3–9; Piliavin et al. 1981: 3–4), inasmuch as ethical reflection is preoccupied with such conduct. Moreover, my arguments do not depend on assuming any especially demanding ethical standard. Unlike "heroic" virtues such as courage, compassion is the subject of quite commonplace ethical demands, demands that are customarily applied to ordinary people in ordinary circumstances. The problem that the empirical work presents is not widespread failure to meet heroic standards – perhaps this would come as no surprise – but

widespread failure to meet quite modest standards. All things considered, my test case should resonate rather broadly.

With this backdrop in mind, it's time for the empirical evidence.[3] I beg the reader's indulgence in a long-winded discussion; this is the only way to responsibly assess a vast experimental literature.

Helping Behavior

Mood Effects

Imagine a person making a call in a suburban shopping plaza. As the caller leaves the phone booth, along comes Alice, who drops a folder full of papers that scatter in the caller's path. Will the caller stop and help before the only copy of Alice's magnum opus is trampled by the bargain-hungry throngs? Perhaps it depends on the person: Jeff, an entrepreneur incessantly stalking his next dollar, probably won't, while Nina, a political activist who takes in stray cats, probably will. Nina is the compassionate type; Jeff isn't. In these circumstances we expect their true colors to show. But this may be a mistake, as an experiment by Isen and Levin (1972) shows. There the paper-dropper was an experimental assistant, or "confederate." For one group of callers, a dime was planted in the phone's coin return slot; for the other, the slot was empty. Here are the results (after Isen and Levin 1972: 387):

	Helped	Did Not Help
Found dime	14	2
Did not find dime	1	24

If greedy Jeff finds the dime, he'll likely help, and if compassionate Nina doesn't, she very likely won't. The situation, more than the person, seems to be making the difference.[4]

On Isen and Levin's (1972: 387) reading, the determinative impact of finding the dime proceeds by influencing affective states; apparently, this small bit of good fortune elevates mood, and "feeling good leads to helping."[5] Numerous studies have shown that mood can have powerful impacts on a wide variety of human functioning: risk taking (Isen and Geva 1987), memory (Isen et al. 1978), cooperative behavior (Carnevale and Isen 1986), and problem solving (Taylor 1991; Isen 1987). Most relevantly, positive affect has repeatedly been shown to be related to prosocial behavior (Aderman 1972: 98–9; Isen 1987: 206–7).[6] The crucial observation is not that mood influences behavior – no surprise there – but just how unobtrusive the stimuli that induce the determinative moods can be. Finding a bit of change is something one would hardly bother to remark on in describing one's day, yet it makes the difference between helping and not.[7]

Related studies suggest that people are more likely to help when exposed to pleasant aromas (Baron and Bronfen 1994; Baron and Thomley 1994; Baron 1997). Baron and Thomley (1994: 780) suspect that the mediating factor is positive affect: Good smells induce good moods, which facilitate prosocial behavior. Once again, a rather trivial situational factor may have a nontrivial impact on prosocial behavior; Baron (1997: 500–1) found subjects near a fragrant bakery or coffee shop more likely to change a dollar bill when asked than those near a neutral-smelling dry goods store. If one must have trouble, best to have it where homey scents abound![8]

Back to our troublesome dime. Are Isen and Levin's nonhelpers behaving incompassionately? Scattered papers are a less-than-dire predicament, so the omission is not serious.[9] On the other hand, the cost of action is low: Help round up the papers and be on your way. And if you've endured the humiliation of scrabbling after scattered papers on a busy street, you may regard such a mishap as one where compassionate behavior is appropriate. In numerous instances Isen and Levin's nonhelping subjects literally trampled the fallen papers; while the footprints they left behind may not be evidence of viciousness, they do seem to tell against the attribution of compassion.[10] Of course, the situation presents bystanders some difficulty in interpretation – would she like help, or would I embarrass her?[11] In fact, evidence suggests that situational ambiguity is likely to impede helping behavior: for example, individuals who hear an emergency may be less likely to help than those who both see and hear it (Shotland and Stebbins 1980: 519).[12] This does not undermine Isen and Levin's result, however. While a sensitive look at the circumstances may tell against judging the passive bystanders too harshly, it does not alter the facts: A mere dime strongly influenced compassion relevant behavior.

Unfortunately, the Isen and Levin subjects did not undergo personality evaluations, so there's no direct evidence regarding dispositional differences, or the lack of dispositional differences, between the helpers and the nonhelpers. But think for a moment of the data: Only 13 percent of dime finders failed to help, whereas 96 percent of nonfinders were similarly passive. Given these numbers, doesn't "He found a dime" look like a plausible, if incomplete, explanation of why Jeff the entrepreneur managed to help? Or are we to suppose that, of a more or less random sample of public phone users in a shopping mall, those possessing robust compassionate dispositions happened to luck into the dime, while their callous brethren didn't (cf. Campbell 1999: 39)?

Now one person did help, despite not finding a dime; perhaps the study shows only that compassionate people are few and far between. Virtue, Aristotle (1984: 1105a7–12) tells us, is difficult; the fact that compassion often fails to be manifested in behavior will not surprise any but the most starry-eyed romantic. But the cases I consider here, like the phone booth study, are ones where prosocial behavior looks to be "minimally decent samaritanism" (see Thomson 1971); the deeds in question do not require heroic commitment or sacrifice. I am not establishing a heroic standard for

good character and arguing from the rarity of this standard being achieved to a general skepticism about characterological moral psychology. Rather, there are problems for standards of character that are well short of heroic, and they are often found in very ordinary places, like the coin return of a public phone.

Group Effects

Another unsettling series of findings, partly instigated by public dismay over the Genovese murder, concern the oft-demonstrated inhibition of helping in groups, or "group effect."[13] In a representative experiment by Latané and Darley (1970: 44–54), puffs of artificial smoke were introduced through a wall vent into a room where undergraduate subjects were filling out forms. After several minutes there was enough smoke to "obscure vision, produce a mildly acrid odor, and interfere with breathing." When the subject was alone in the room, 75 percent (18 of 24) reported the smoke to experimenters within four minutes; when the subject was with two passive confederates, only 10 percent of subjects (1 of 10) reported it. In a trial with three naive subjects per group, in only 38 percent of groups did someone report the smoke, as opposed to the 98 percent one would expect statistically based on the 75 percent response rate in the alone condition. Latané and Darley (1970: 48–52) speculate that in this instance the group effect proceeded by influencing interpretative processes: Seeing confederates acting unconcerned, subjects were more inclined to interpret the "ambiguous" stimulus of artificial smoke as "nondangerous" steam or air conditioning vapors, despite the fact that it moved them to cough, rub their eyes, and open windows.[14]

A related study by Latané and Rodin (1969; cf. Latané and Darley 1970: 57–67) solicited Columbia University undergraduates for participation in a market research study. When they reported to the experimental site, an attractive[15] young woman introduced herself as a "market research representative," provided the subjects with some questionnaires to fill out, and withdrew behind a curtain dividing the room. Subjects were subsequently interrupted by a loud crash, followed by the woman's cries of pain. Apparently, this constituted an arresting and realistic impression of a serious fall taking place behind the curtain: Less than 5 percent of subjects reported suspecting that the victim's cries were recorded, as they in fact were. Seventy percent of bystanders offered help when they waited alone, compared with 7 percent in the company of an unresponsive confederate. When two subjects not previously acquainted waited together, in only 40 percent of groups did one of the subjects intervene, compared with the 91 percent expected based on a 70 percent rate when subjects were alone. Here, too, the group effect appeared to operate through the interpretative process: Nonhelpers said they were unsure of what happened or decided it was not serious. Accordingly, postexperimental interviews revealed that passive subjects did not feel as though they had acted callously: They typically claimed they would readily help in a "real" emergency (Latané and Rodin 1969: 197).

Latané and Darley (1970: 95–100) also discovered a somewhat different effect. They asked students to participate in a group discussion of the problems faced by college students in an urban environment. The ostensible "discussion" proceeded by intercom with the experimenter absent and the subject isolated in a cubicle, ostensibly to preserve anonymity; in fact, the other "participants" were tape recordings, and the situation was designed to address a variant of the group effect. One tape-recorded participant described his difficulty with seizures; he later gave an arresting impression of someone suffering a seizure (1970: 97, 100). Again, the group effect: 100 percent of subjects believing themselves alone with the seizure victim intervened, while only 62 percent of subjects in a "group" consisting of subject, victim, and five more tape-recorded participants did so.

Apparently, in this case the inhibiting mechanism consisted at least partly in a "diffusion of responsibility" (Latané and Darley 1970: 101, 111): The presence of others meant that no individual was forced to bear full responsibility for intervention.[16] When the experimenter terminated each trial after 6 minutes, unresponsive subjects in group conditions appeared aroused and conflicted. Isolated in their cubicle, they lacked the social cues necessary to facilitate an interpretation congenial to inaction, but knowing there were other bystanders, it was not clear that intervention was up to them. In contrast, the passive bystanders in the previous two experiments, where social influence rather than diffusion of responsibility was the inhibiting factor, seemed relaxed; the presence of other passive bystanders assured them that their inaction was appropriate despite the considerable evidence to the contrary (Latané and Darley 1970: 111–12). Then the group effect involves more than one sort of effect. It is not simply that numbers of bystanders influence intervention; different configurations of bystanders may influence intervention in different ways.[17] The operative processes are doubtless complicated, but one general implication of the group effect studies seems fairly clear: Mild social pressures can result in neglect of apparently serious ethical demands.

Good Samaritans

In one of the most widely discussed situationist experiments, Darley and Batson (1973) invited students at the Princeton Theological Seminary to participate in a study of "religious education and vocations." Subjects began experimental procedures by filling out questionnaires in one building and then reported to a nearby building for the second part of the experiment, which consisted in their giving a short verbal presentation.[18] Before leaving the first site, subjects were told either that they were running late ("high hurry" condition), were right on time ("medium hurry" condition), or were a little early ("low hurry" condition); thus the conditions exerted a

different degree of time pressure on the subjects.[19] The behavior of interest occurred on the walk between the two sites, when each seminarian passed an experimental confederate slumped in a doorway, apparently in some sort of distress.

One might expect that most individuals training for a "helping profession" like the ministry would be strongly disposed to assist the unfortunate victim or at the very least inquire as to his condition.[20] Instead, helping varied markedly according to degree of hurry (Darley and Batson 1973: 105).[21]

	Degree of Hurry		
	Low	Medium	High
Percentage helping	63	45	10

It's no surprise that haste can have people paying less regard to others. But the apparent disproportion between the seriousness of the situational pressures and the seriousness of the omission is surprising: The thought of being a few minutes late was enough to make subjects not notice or disregard a person's suffering. The imagery recalls the most cynical caricatures of modern life: Darley and Batson (1973: 107) report that in some cases a hurried seminarian literally stepped over the stricken form of the victim as he hurried on his way!

It is difficult to resist situationist conclusions. Subjects were hurried but certainly not coerced. Nor was there special reason to think, in the green fields of 1970s Princeton, New Jersey, that the victim posed some threat, as might be supposed in more threatening urban climes. Similarly, the placid suburban environment should have worked to reduce situational ambiguity. While urbanites who are daily confronted with the homeless may find themselves wondering whether the unfortunate individual lying on the sidewalk is sick or dying as opposed to inebriated or sleeping, such sights were presumably uncommon enough in the Princeton of 1970 to strongly suggest that something was seriously amiss (cf. Campbell 1999: 28). But hurried seminarians failed to help. What was at stake for them? Did they somehow decide that their obligation to the experimenter trumped a general imperative to help others in distress? In its generality, this looks like a plausible interpretation, but it's hard to believe such an obligation could be viewed as very weighty: Subjects were volunteers being paid a modest $2.50, and the experimenter was someone they had only just met.[22] Once again, there is the appearance of disproportion; in this case the demands of punctuality seem rather slight compared with the ethical demand to at least check on the condition of the confederate.[23]

Helping and Personality

Between 1962 and 1982 more than 1,000 studies on helping behavior and altruism were reported in the psychology literature (Dovidio 1984: 362);

I confess with some embarrassment that the preceding discussion has reported only a fraction of the relevant material. However, my sampling is representative of established trends. As I've said, situationism is motivated by a pattern of results, not by the results of any particular study; I'm discussing some high points of the tradition, but there are many other studies that equally support my interpretation. I'll now say something more about how my interpretation goes.

It would be a serious mistake to understand the situationist experiments as empirical evidence against the existence of altruism. While egoistic theories of motivation are common enough in the social sciences,[24] I doubt questions about the possibility of altruism admit of empirical resolution, since the issue concerns what sort of motivations should be counted as altruistic, and this is substantially a conceptual difficulty. Still, there is a sense in which I might be accused of painting a misleadingly dreary picture of human behavior. The studies I've relied on, like most of those in the prosocial literature, involve helping behavior amongst strangers (see McGuire 1994). But of course much helping, and much human kindness, occurs in the context of social bonds: between friends, family, and coworkers. And here, perhaps, we are right to expect more compassion than we do amongst strangers: Surely I don't suppose that 90 percent of mothers in a hurry would step over the stricken form of their own child? Of course not; nothing I've said contradicts the thought that people help most, and are most helped by, the ones they know and love. Where social ties exist, helping is very likely more reliable than among strangers. At risk of churlishness, however, I cannot resist cautionary observations: Lovers cheat, siblings fight, and parents are unresponsive. More important, the situationist can grant even strong claims for the consistency of prosocial behavior in ongoing relationships, for surely the explanation here is substantially situational: Relationships underwrite affective ties and reciprocal structures that facilitate helping behavior. For all that, we find considerable helping even amongst strangers: Numerous studies of staged emergencies have found impressive rates of intervention, in some conditions approaching 100 percent (Piliavin et al. 1969: 292; Clark and Word 1972: 394–7; Harari et al. 1985: 656–7). The situationist point is not that helping is rare, but that helping is situationally sensitive.

As with all psychology experiments, the studies I've cited encounter questions of *ecological validity*: To what extent does a given experimental finding accurately reflect phenomena found in natural contexts?[25] Experimental situations are in many cases radically different from the natural situations they are meant to address; accordingly, applying experimental work to the interpretation of natural situations is an extrapolative process. As a (roughish) rule, the more closely the experimental situation resembles its natural counterpart, the more straightforward the extrapolation will be. At least initially, the experiments we've just considered seem to fare pretty well in this respect; for instance, the situation faced by subjects in the phone booth study bears

a more than passing similarity to the sort of helping situations people encounter in everyday life.

Field studies like the phone booth demonstration are less subject to worries about ecological validity than are lab studies like the seizure experiment, because subjects in laboratory experiments know they are in an "artificial" situation, an awareness that may influence how they judge and behave.[26] But ecological validity does not require that experimental situations resemble the relevant natural situations exactly or even very closely; more important for the purposes of generalization is whether the processes at issue in each case can plausibly be considered analogous. Nobody is arguing that the group effect studies are exactly like the Genovese tragedy; the point is that there is good reason to think closely related social processes are at work in both instances. More generally, it strains credulity not a bit to claim that people are influenced by mood, time pressures, and the presence of others in both natural and experimental contexts.

But I'm in the business of arguing something that does strain credulity a bit: seemingly insubstantial situational factors have extraordinary effects on behavior. This is undeniably true in experimental contexts, but I contend that it is quite generally true. I'm therefore making an extrapolation, but notice what is required to refute it: One would have to show not that the experimental contexts are different, or even vastly different, from the natural contexts, but that there are differences suggesting that situational factors are less powerful in natural contexts than they are in experimental contexts. Perhaps this can be argued in particular cases, but I suspect this is going to be difficult to establish for a preponderance of relevant experiments and, most especially, for the field studies: Is there some reason to suspect that Isen and Levin's dimes were unnaturally potent?

Indeed, there's an obvious explanation for why the disconcerting potency of small situational variations is more evident in experiments than in life. Given how counterintuitive it is to suppose that such factors powerfully influence behavior, it is no surprise that people typically pay them little attention, and even in the unlikely event that people developed situationist suspicions in the ordinary course of things, it would be difficult for them to engage in the systematic observation required to put such suspicions to the test. Conversely, this is just what experimental observation is designed to do; it's not that the experimentally identified phenomena are not present in natural contexts, but that they are not as readily there adduced.

Then I won't much worry here about ecological validity; for my purposes, the central interpretive issue concerns what experimental work on helping can tell us about the behavioral ramifications of character. Consider first the role of demographic variables like sex and socioeconomic status, a topic that has been the subject of some study. Now these demographic variables are not quite the same thing as character or personality

traits, but if it were shown that such variables impacted helping behavior, it would appear to give the character theorist a foot in the door. Suppose women were reliably more helpful than men. It might then be tempting to conclude that women tend to have more robust compassionate dispositions than men, which is to say that variance along a trait dimension accounts for variance in helping behavior. However, the empirical evidence for a conjecture of this kind is rather weak. Some studies have found no relationship between sex and prosocial behavior, others have found more prosocial behavior on the part of men, and still others have found more prosocial behavior on the part of women.[27] In particular, this pattern or, rather, lack of a pattern, has been found over numerous studies of the group effect (Latané and Darley 1970: 104; Latané and Nida 1981: 315–16). In investigating other demographic correlates of helping, Latané and Darley (1970: 117–19) found that socioeconomic status is not strongly associated with helping behavior, although they do report a modest relationship between bystanders' hometowns and helping behavior, with bystanders hailing from smaller communities being more likely to help than bystanders from larger communities.[28] Perhaps the character theorist can find a glimmer of hope here – it might be argued that rural environments can effectively nurture robust compassionate dispositions – but overall the evidence provides little indication that demographic characteristics are an important determinant of helping behavior.

For the most part, attempts to directly relate personality evaluations to helping behavior have had similarly uncertain results (Krebs 1970: 284–5; Piliavin et al. 1981: 185–92). Darley and Batson (1973: 106) found little relationship between personality measures tapping "types of religiosity" and helping on the part of their seminarians.[29] Yakimovich and Saltz (1971: 428) found that various trait measures – including those for trustworthiness, independence, and altruism – were unrelated to helping in a staged accident paradigm. In the Latané and Darley (1970: 114–15) seizure study, measures of various personality traits – including authoritarianism, Machiavellianism, and social responsibility – failed to predict helping; in a variation conducted by Korte (1971: 155–6), measures of deference, autonomy, and ascendance did not predict helping behavior.

On the other hand, Denner (1968: 461–2) found that subjects exhibiting a low tolerance for ambiguity were less reluctant to report a theft than individuals with high tolerance, while Michelini and associates (1975: 256–7) discovered that individuals manifesting a high concern for esteem were more likely to assist someone who had dropped an armload of books than were individuals with high concern for safety. Based on a suggestive series of studies, Schwartz (e.g., Schwartz and Ben David 1976; Schwartz 1977) argues that individual tendencies to accept rather than deny responsibility are positively related to a range of prosocial behavior, including emergency

intervention and volunteer work. While there is empirical evidence for Schwartz's view, his results do not in every case seem especially strong (e.g., Schwartz and Clausen 1970: 306; Schwartz and Ben David 1976: 410–11), and they have not always been substantiated by other investigators (e.g., Zuckerman and Reis 1978: 505).

I do not contend that there is nothing to recommend personological approaches to prosocial behavior, but it seems more than fair to conclude that the results of this work are equivocal. As is often the case, interpretation of the evidence is to some extent a question of taste: One commentator's equivocal results are another's suggestive results. Obviously, I find evidence for the power of the situation highly suggestive and evidence for the power of personality highly equivocal; others might take the opposite view. I don't really think it's a tie, though: The situationist results we have seen, and those we see below, form a body of research that is undeniably striking, even on the most casual reading, while results having to do with personality and helping often seem rather modest even after application of powerful statistical techniques by sympathetic practitioners.

I must acknowledge an important limitation in the studies I've described: They typically address not patterns of behavior but a particular behavior in a particular situation. While such studies show that insubstantial situational factors may powerfully impact behavior, they can tell us nothing directly about the consistency of the subjects: Direct evidence for or against any particular individual's behavioral consistency requires systematic observation of that individual's behavioral patterns. To gather this sort of evidence, one requires longitudinal studies that observe individuals over a period of many years in numerous and diverse situations.[30] It cannot be denied that there is a dearth of such studies; they are all but prohibited by logistical obstacles, including high cost and professional pressure on academic investigators to "get quick results." Nevertheless, the situationist has a powerful indirect argument against the existence of widespread consistency in helping behavior. The prosocial literature provides unequivocal evidence that situations have powerful determinative impacts on behavior. Add to this the highly plausible speculation that people will typically experience situations with highly variable levels of conduciveness to prosocial behavior, and it seems eminently reasonable to conclude that people will typically exhibit inconsistent prosocial behavior.[31]

If I am right, then, characterological moral psychology is an empirically inadequate approach to the determinants of helping behavior. But the point needs to be put carefully. Flanagan (1991: 295, 302), a generally sympathetic commentator on situationism, cautions that results like Darley and Batson's have "no implications whatsoever for the general issue of whether there are personality traits." True enough. But the question concerns the most perspicuous characterization of personality traits, not their existence. The situationist does not deny that people have personality traits; she instead

denies that people typically have highly general personality traits that effect behavior manifesting a high degree of cross-situational consistency. It is not often going to be the case, as philosophers might be tempted to allege (see Feinberg 1992: 178), that those emerging as Failed Samaritans in some situation suffer a general "character flaw," while those presenting as Good Samaritans are motivated by a general "surplus of benevolence."

Of course, the research we've considered generates skepticism only about personality measures actually subjected to behavioral investigation. As I've said, my skepticism is inductive; accordingly, it leaves open the possibility of highly general personal influences on prosocial behavior that investigators have hitherto failed to discover. An inductive skepticism is a defeasible skepticism. All the same, folks have been at it a while; a situationist bet on future developments doesn't seem a wild gamble.

Destructive Behavior

The Milgram Experiments

So far, we have examined experimental manipulations which appear to generate omissions of compassion, failures to act where one might fairly expect a person of ordinary moral stature to do so. Social psychologists have also performed experimental manipulations of active harming behavior, laboratory inducements to destructive behaviors one would expect a person of ordinary moral stature to quite readily avoid. The classic studies in this vein are the famous, or infamous, "obedience experiments" conducted by Stanley Milgram.[32] While they are among the most widely recognized, and among the most important, of all psychological demonstrations, it is not obvious that we have come fully to grips with the notorious "experiments where they shocked people." Nor is it the case that philosophers have been especially engaged with Milgram's work, despite its apparent ethical significance.[33] Even among those intimately acquainted with the experiments, their interpretation is a matter of controversy, so I shall, at the risk of belaboring some well-known points, go into considerable detail.

For the impatient reader, I offer my main conclusions in advance.[34] Milgram's experiments show how apparently noncoercive situational factors may induce destructive behavior despite the apparent presence of contrary evaluative and dispositional structures. Furthermore, personality research has failed to find a convincing explanation of the Milgram results that references individual differences. Accordingly, Milgram gives us reason to doubt the robustness of dispositions implicated in compassion-relevant moral behavior; his experiments are powerful evidence for situationism. For the patient reader, I'll now substantiate these conclusions in considerable detail.

From 1960 to 1963, Milgram (1974: 1–26) ran various permutations of his experiment with approximately 1,000 subjects drawn from various socioeconomic groups in the New Haven area – postal clerks, high school

teachers, salesmen, engineers, and laborers – who responded to newspaper and mail solicitations seeking paid participants for a study of memory and learning at Yale University. Here's how the story goes.

On arrival at the site, the subject is met by a lab-coated "experimenter" who introduces him to another ostensible subject, actually a confederate, and explains that the study concerns the effects of punishment on learning. There is a drawing to determine experimental roles, rigged so that the subject is designated "teacher" and the confederate "learner." The learner, an affable middle-aged accountant, is strapped into a chair "to prevent excessive movement." An electrode is attached to his wrist with electrode paste "to avoid blisters and burns." The experimenter assures participants that the shocks used as punishment, although they can be extremely painful, will cause no "permanent tissue damage." The teacher is administered an uncomfortable sample shock to convince him of the scenario's authenticity; however, the "shocks" administered the learner are fake, and he experiences no pain.

The teacher is then led to another room and seated in front of an imposing "shock generator" that the experimenter explains is wired to the electrode on the learner, who is now hidden from view in the first room. The teacher next remotely administers a word-association test to the learner; the learner's answers are displayed above the shock generator, and with each wrong answer, the teacher administers a shock, which is increased in intensity one increment for each wrong answer. The learner responds incorrectly on a prearranged schedule, so that shock intensity steadily increases. If the teacher expresses concern about this process, as many subjects did, the experimenter responds with a standardized series of verbal prods: (1) "Please continue," (2) "The experiment requires that you continue," (3) "It is absolutely essential that you continue," (4) "You have no other choice, you *must* go on." The sequence begins again at (1) each time the subject balks and progresses through (4) if he continues to refuse. The experimenter may also repeat the reassurance that the shocks "cause no permanent tissue damage" if the subject expresses concern over the learner's safety. If the subject refuses to continue after prod (4), the experiment terminates, and the subject is counted "disobedient"; subjects who comply with all instructions and proceed to the maximum shock are termed "obedient." According to Milgram (1963: 374), the experimenter's tone of voice is "at all times firm, but not impolite."

Consider the following table, which reports the experiment's fifth permutation. In the left column are the shock levels subjects saw on the generator. At center is the learner's response as played in a standardized recording. On the right is the number of teacher-subjects at each shock level who refused to obey the experimenter's instructions and were counted disobedient.

It is instructive to read the responses with care and imagine hearing them, as the teacher did. With each error, the teacher ups the voltage one

Experiment Number 5 "New Baseline Condition"[35]

Voltage Level	Learner Response		Number Disobedient
Slight shock	15		
	30		
	45		
	60		
Moderate shock	75	Ugh!	
	90	Ugh!	1
	105	Ugh! (*louder*)	
	120	Ugh! Hey, *this* really hurts.	
Strong shock	135	Ugh!	
	150	Ugh!!! Experimenter! That's all. Get me out of here. I told you I had heart trouble. My heart's starting to bother me now. Get me out of here, please. My heart's starting to bother me. I refuse to go on. Let me out.	6
	165	Ugh! Let me out! (*shouting*)	
	180	Ugh! I can't stand the pain. Let me out of here! (*shouting*)	1
Very strong shock	195	Ugh! Let me out of here. Let me out of here. My heart's bothering me. Let me me out of here! You have no right to keep me here! Let me out! Let me out of here! Let me out! Let me out of here! My heart's bothering me. Let me out! Let me out!	
	210	Ugh! Experimenter! *Get* me out of here. I've had enough. I *won't* be in the experiment any more.	
	225	Ugh!	
	240	Ugh!	
Intense shock	255	Ugh! Get me *out* of here.	
	270	(*Agonized scream.*) Let me out of here. Let me out of here. Let me out of here. Let me out? Do you hear? Let me out of here.	2
	285	(*Agonized scream.*)	
	300	(*Agonized scream.*) I absolutely refuse to answer any more. Get me out of here. You can't hold me here. Get me out. Get me out of here.	1
Extreme intensity shock	315	(*Intensely agonized scream.*) *I told you I refuse to answer.* I'm no longer part of this experiment.	1

(continued)

(continued)

Voltage Level	Learner Response	Number Disobedient	
	330	(*Intense and prolonged agonized scream.*) Let me out of here. Let me out of here. My heart's bothering me. Let me out, I tell you. (*Hysterically.*) Let me out of here. Let me out of here. You have no right to hold me here. Let me out! Let me out! Let me out! Let me out of here! Let me out! Let me out!	1
	345		
	360		
Danger:			
Severe shock	375		1
	390		
	405		
	420		
XXX	435		
	450		
Total disobedient			14
Total obedient			26
Percentage obedient			65

increment, gradually moving across the control panel to increasingly ominous designations and increasingly vehement protests. After the last protest, at 330 volts, the learner is unresponsive. Has he had a heart attack? What is the teacher to do? According to the experimenter, the teacher is to treat no answer as a wrong answer and continue the progression. We are left with the indelible image of two-thirds doing so until the bitter end.[36]

The experiment does not suggest that Milgram had stumbled onto an aberrant pocket of sadists in the New Haven area and still less does it suggest that all of us are a bunch of meanies. Trait-contrary behavior does not necessarily signal the possession of a contrary trait; even active failures of compassion do not necessarily imply sadism. What the experiments do highlight, once more, is the power of the situation: the majority of subjects were willing to torture another individual to what seemed the door of death without any more direct pressure than the polite insistence of the experimenter. But it is badly mistaken to think that the obedient subjects generally found their job easy – the experiment does not show, as is sometimes suggested (Goldhagen 1996: 383), that people are blindly obedient to authority. The most striking feature of the demonstration is not blind obedience but *conflicted* obedience. Horribly conflicted obedience: Subjects were often observed to "sweat, tremble, stutter, bite their lips, groan, and

dig their fingernails into their flesh" (Milgram 1963: 375). One onlooker offered this description:

> 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. He constantly pulled on his earlobe, and twisted his hands. At one point he pushed his fist into his forehead and muttered: "Oh God, let's stop it." And yet he continued to respond to every word of the experimenter, and obeyed to the end. (Quoted in Milgram 1963: 377)

On its face, the fact that the experimenter's "firm, but not impolite" prodding generated such grotesque compliance is merely ridiculous. Indeed, this has convinced some observers that the experimental behavior must be a laboratory artifact unrelated to destructive obedience in natural contexts; for them, the best explanation of the compliant behavior is that subjects were not taken in by the hoax and were instead humoring the experimenter with a kind of play-acting (Orne and Holland 1968; Patten 1977a, b). Now the reason for favoring this explanation had better not be that the "preposterousness" of the situation suggests that the subjects could not have thought the shocks were genuine (see Patten 1977b: 432–3). Many social organizations with strong ritual elements – fraternities, sports teams, street gangs, and military outfits – may seem more than faintly preposterous when viewed from the outside, but participants very often view the proceedings with deadly earnest.

A better reason for doubting the success of Milgram's deception would be skepticism explicitly voiced by the subjects. A follow-up questionnaire Milgram distributed about a year after the study provides limited evidence of such skepticism: Of over 600 subjects responding, 80 percent felt it certain or probable that the learner was receiving painful shocks, while the remaining 19 percent were either (1) not sure, (2) doubtful, or (3) certain that the shocks were fake, with only 2.4 percent of these expressing certainty (Milgram 1974: 172–3; cf. Elms 1972: 121). If we consider the attractiveness of an "I wasn't fooled" rationalization for obedient subjects who may very well have been dismayed by their own conduct, the 80 percent figure for credulous subjects seems impressive indeed (Milgram 1974: 173–4). But I'd be among the first to question the diagnostic efficacy of self-reports, so I won't lean too heavily on these results (cf. Patten 1977b: 431–2).

As Milgram (1974: 43, 171) observed, the best evidence for the experimental realism of his paradigm is the extraordinary anxiety of the subjects, amply documented by experiment transcripts (e.g., Milgram 1974: 73–84) and Milgram's (1965) instructive film of the experiment. In fact, the subjects' evident suffering provoked heated ethical criticisms of Milgram's research (Patten 1977a; cf. Miller 1986: 88–138); research ethics are not my concern here, but it's hard to see why there should have been an

ethical outcry if the subjects were happily going on a lark.[37] If the hoax was obvious, why the trembling, stuttering, and groaning? (Indeed, if the hoax was obvious, why wasn't there 100 percent untroubled obedience?) This point is not always sufficiently appreciated by critics.[38] Patten (1977b: 432), in a surprisingly brief discussion amidst a sustained critique of Milgram, argues that subject stress is not particularly suggestive, because even a robust skepticism is compatible with moments of uncertainty: The anxiety of the mostly skeptical subjects is to be explained by occasional "trials of self doubt." For a moment, grant this explanation; the experiment still raises grave concerns about the surprising extent of destructive obedience. Suppose subjects thought it was only probable that the learner was in real distress – say, 4 chances in 5? Or suppose they thought it only somewhat likely, say 1 chance in 3, or even 1 in 10? Would you feel comfortable taking such a chance? The subjects, quite apparently, did not. Obedience where one believes the probability is relatively slight that one is inflicting serious harm on another human being still looks to be ethically problematic. The anxiety manifested by the subjects strongly suggests that they shared this assessment of their conduct, even on the generous assumption that they were often substantially skeptical about the shocks' authenticity.

Another criticism of Milgram emphasizes not the skepticism of the subjects but their credulity. Subjects' faith in the experimenter, who after all was standing by more or less impassively, and the larger institution of science assured subjects that the learner suffered no real harm (Orne and Holland 1968: 287, 291; cf. Darley 1995: 129). There's something to this: Experimental subjects volunteer to participate in an institution they evidently hold in high regard; one should expect them to believe that the experimental environment is a safe one.[39] Apparently, the idea is that the subjects' confidence in the experimenter is what leads them to believe that the shocks cannot be real – their trusted leader could not be ordering them to harm others (Orne and Holland 1968: 287). Notice that if we juxtapose this argument and that of the previous paragraph, as some critics seem to (e.g., Orne and Holland 1968: 287), we attribute to the subjects a rather remarkable attitude toward the proceedings. The same people who trust the experimenter implicitly despite the alarming responses from the learner easily ferret out the elaborate deception and then are polite enough to claim credulity on the subsequent questionnaire! Such Rube Goldberg psychological complexity is not impossible, but this explanation of the phenomenon is not to be preferred on the grounds of simplicity. Appeals to subjects' confidence in the experimenter are also undermined by an extension conducted by Ring and colleagues (1970: 72), who replaced Milgram's impassive "everything's under control" experimenter with an experimenter who exhibited surprise at the alarming proceedings, shaking his head and mopping his brow: Obedience was 91 percent (52 of 57).[40] Even when presented with direct cues undermining the experimenter's competence, the overwhelming majority of the subjects

went along; if the best explanation of their behavior is their faith in the experimenter, their faith must have been a pious one indeed.

Perhaps most telling against the suggestion that subjects doubted the shocks is a variation by Sheridan and King (1972: 165–6), where *actual* shocks were administered to a "cute, fluffy puppy" that exhibited "continuous barking and howling" at higher shock levels! And in this instance, the experimenter did not assure subjects that the learner would suffer no permanent injury, as Milgram's experimenter did. Yet 77 percent of subjects were fully obedient. Did these subjects discover a deception that didn't exist? Perhaps obedience here was partly due the fact that the victim was only a puppy – however cute – and not a person. But just as in Milgram, substantial stress was observed: "pacing from foot to foot, puffing, and even weeping" (Sheridan and King 1972: 166). The subjects apparently (and quite accurately) believed they were hurting the puppy and they didn't much like it. In sum, there are numerous reasons to think that widespread obedience may obtain with credulous subjects.

Although ethical considerations have limited research (Blass 1991: 408; 1996: 1530; Elms 1995: 27), there exist numerous other replications and extensions of Milgram's paradigm. According to Brown (1986: 4), "two-thirds obedient, everywhere the experiment has been tried, is a fair summary."[41] Mantell (1971) found 85 percent of his German subjects to be fully obedient, while Shanab and Yahya's (1978) replication with Jordanian subjects found 63 percent fully obedient. In the Netherlands, Meeus and Raaijmakers' (1986: 313–19)[42] "experimenter" had subjects administer a preemployment test by intercom to a confederate "job applicant." During the test subjects delivered a scripted series of fifteen increasingly disparaging "stress-remarks" to the applicant, for example, "According to the test, it would be better for you to apply for lower functions." The confederate-applicant responded with a scripted series of protests expressing increasing levels of agitation and dismay. In addition, the experimental scenario informed subjects of the following: (1) the applicant agreed to participate in the experiment on the condition that it would not impair test performance, (2) the stress remarks might impair performance, (3) stress performance was unrelated to the job requirements, and (4) the applicant would not get the job if he failed the test. Subjects who administered all fifteen "stress remarks" were termed obedient; 92 percent of subjects (22 of 24) did so.[43] The affinities with Milgram's work are obvious, but the results are in a way more chilling, despite the lower level of gore, because the opportunities for administrative-psychological brutalization of innocents are so pervasive in bureaucratic societies. Neither the level of obedience nor the observation that "there was no real opposition to the experimenter" (Meeus and Raaijmakers 1986: 318) would be very likely to surprise anyone who has been ground in the bureaucratic mill.[44]

After voluminous commentary, there is a substantial consensus that Milgram's methodology is sound (Ross 1988: 102; A. Miller 1986: 139–78;

1995: 38–40). As we've just seen, his results are not an artifact of his laboratory; related demonstrations of laboratory obedience are common enough.[45] Nevertheless, it is possible that the entire body of obedience results is substantially a function of laboratory artifacts.[46] I'll eventually address these concerns by relating the results to natural contexts, but for the moment, I'll consider the implications of the studies themselves for thinking on character.

I claimed at the outset that interpretations of the Milgram results appealing to individual differences are not especially promising; instead, the experiments provide compelling evidence for the power of the situation. As evidence for this, consider the widely varying levels of obedience obtained across variations in the situation Milgram's subjects faced (Miller 1986: 210):

- When subjects were free to choose the shock levels to administer to the victim, only 3 percent delivered the maximum shock (Milgram 1974: 61).
- When the experimenter was physically absent and gave his orders by phone, obedience was 21 percent (Milgram 1974: 60).
- In a "touch-proximity" condition where the subject was instructed to press the victim's hand onto a "shock plate" to administer the punishment, obedience was 30 percent (Milgram 1974: 35).
- When a confederate "peer" administered the shock while the subject performed only subsidiary tasks such as administering the test, obedience was 93 percent (Milgram 1974: 119).

These variations admit of plausible explanations. For most people, it will likely be easier to harm a distant victim than a near one, easier to defy a distant authority than a near one, and easier to perform tasks subsidiary to harming than actual harming. But the central observation remains: The variation in obedience across experimental conditions – from near negligible to near total – is powerful evidence that situational variation can swamp individual differences. Or is it to be supposed that 39 virtuous subjects and one vicious subject were assigned to the 3 percent obedient "subject chooses shock level" condition, while 37 vicious subjects and three virtuous subjects were assigned to the 93 percent obedient "peer administers shocks" condition?

But there were not significant differences between all of Milgram's variations, even in cases where such differences would intuitively be expected.[47] And even where it is evident that a variation has substantial impact, the manipulations do not effect complete uniformity of behavior; therefore, individual dispositional differences must be doing some of the work (see Blass 1991: 402). True enough, but there is a wrinkle worth noting: Subjects in ostensibly identical experimental conditions may experience different situational pressures. For example, Milgram's experimenter did not treat every subject the same, despite the scripted prods. As Darley (1995: 130; cf. Milgram 1974: e.g., 74, 76) points out, in contrast to the recording of learner protests, which never varied, the experimenter exercised

considerable latitude for improvisation in his "prods," apparently in attempts to secure the maximum possible obedience from each subject. There is little reason to think the experimenter's stratagems were equally potent in each instance; thus each trial represents a different "microsituation" that may have a different impact on behavior (see Modigliani and Rochat 1995: 108–9; Rochat and Modigliani 1995: 206). This is an important point, but it risks cutting things too fine: With carefully designed studies like Milgram's, we can have some confidence that, for at least a substantial percentage of trials, subjects in the same experimental conditions experienced a good approximation of *relevantly similar* situations. So differences amongst individuals matter. What differences? How much?

Sex of subject might be expected to make a difference: Miller et al. (1974: 27; cf. Sheridan and King 1972: 166) found that undergraduate subjects expected males to give higher levels of shock than females.[48] Not so: Milgram's (1974: 61–3) trial of experiment 5 with all women subjects yielded 65 percent obedient, just as in the all-male trial, while Ring et al. (1970) obtained 91 percent obedience with female subjects.[49] In replications with both children and adults, Shanab and Yahya (1977, 1978) observed no sex differences in obedience, while in their extensions, Kilham and Mann (1974) found greater obedience for males and Sheridan and King (1972: 166) found greater obedience for females.

In addition, subject age seems to be unimportant. Shanab and Yahya (1978) found 73 percent of Jordanian children fully obedient in a replication of Milgram. In an extension by Martin and colleagues (1976: 349) involving thirteen- and fourteen-year-old boys, 54 percent were willing to administer a full series of fictitious "ultra-high frequency" sounds to *themselves*, despite being informed of potential hearing loss. Perhaps the critic will not be much impressed; one should expect children to be highly compliant with adult experimenters. But take things up from the other end: Shouldn't one expect adults to be less obedient? With increasing age, one expects increasing autonomy or, to wax Aristotelian, fuller character development, but the evidence regarding obedience does not support this. Milgram (1974: 205) considered various other demographic variables such as education and occupation, and while his book does not report the results in detail, he found them "generally weak."

Work on the relation of personality measures to obedience is similarly unimpressive. Elms and Milgram (1966: 287–8; cf. Elms 1972: 132–3) administered standard personality instruments to samples of obedient and defiant subjects; while they did not find a "single personality pattern" expressed in one behavior of the other,[50] they found that obedient subjects tended to score higher on an F-scale for authoritarianism. Insofar as the "authoritarian personality" is expected to be more subordinate with superiors, this is just the sort of difference one might expect in Milgram's subjects. However, Milgram (1974: 205) later remarked that the relationship between the F-scale and obedience, "although suggestive, is not very strong."[51]

Kohlberg (1984: 546–8; cf. Kohlberg and Candee 1984: 67–70) reports that disobedient subjects were further advanced on his moral development scale than were obedients. Unsurprisingly, Kohlberg's findings have been frequently cited by those who favor a personological interpretation of Milgram's results (e.g., Miale and Selzer 1975: 12; see Miller 1986: 240). But Kohlberg describes his results in minimal detail, so caution seems appropriate.[52] Elms (1972: 135) voiced skepticism, as did Milgram (1974: 205). The ambivalence of commentators here is especially striking, since the Elms and Milgram and Kohlberg studies are perhaps the most prominent evidence for the influence of personality variables in the obedience experiments.

Overall, there is a paucity of evidence favoring personological approaches to Milgram's experiments; surveys by both the skeptical Miller (1986: 238–42) and the more sympathetic Blass (1991: 402–3) do not adduce a large body of systematic research with impressive results. Of course, this might simply show that standard personality instruments are not as nuanced as we would like, especially when we are seeking explanations of behavior in unusual conditions (Elms and Milgram 1966: 288). Milgram (1974: 205) himself was "certain" that there is a "complex personality basis to obedience"; his was not a general skepticism about personality, but a skepticism about psychologists' ability to measure personality.[53] As will become evident in Chapter 4, I doubt matters have improved much since Milgram; to put it more precisely, I believe that the substance of Mischel's 1968 critique of personality psychology essentially stands.[54] But I certainly cannot rule out the possibility that different methods or the investigation of different personal variables would have motivated a conclusion other than the one Milgram reached.

There are commentators who favor characterological approaches to Milgram despite the paucity of systematic evidence. Miale and Selzer (1975: 10) suspect that disobedient subjects were "more moral" and more averse to inflicting suffering on others than obedient subjects, while Patten (1977b: 439) concludes that the "Socratic skills of self-mastery, courage and moral stubbornness" are the requisites for avoiding destructive obedience. As vague generalities, such observations have a pleasing ring, but the experiments tell us little about the character of individual subjects; they concern behavior in isolated trails and are therefore silent on the crucial question of consistency. It is true that the Milgram situation looks to be what I've called diagnostic; given the difficulty of behaving compassionately or otherwise admirably in such circumstances, disobedience looks to be evidence for the attribution of some morally admirable trait or traits. Nevertheless, each subject was observed only in a single trial. Damn the obedients and hail the defiants if you will; the experiment does not motivate confidence about how particular subjects would behave in markedly dissimilar situations. There's little reason for confidence that the disobedient subjects, however inspiring

their behavior in the experiment (e.g., Milgram 1974: 48, 85), could be counted on to exhibit Socratic self-mastery in other situations. Conversely, do we think that the obedient subjects were in the habit, say, of applying severe shock to friends and family?[55]

One can fairly assume that in a more or less representative sample of "normal" Americans such as Milgram's subjects, most will have internalized norms prohibiting the behavior of the obedient subjects (Milgram 1963: 376; Ross 1988: 102; Gibbard 1990: 58–60).[56] In fact, none of those Milgram (1974: 27–31) surveyed predicted they would be fully obedient were they subjects, and their typical prediction for others was 1 or 2 percent fully obedient. Apparently, the subjects themselves would have antecedently regarded their behavior as aberrant. Further, remember that only 3 percent of subjects were fully obedient when allowed to choose the shock level themselves. But in variations involving experimenter command, obedient subjects behaved in ways radically at odds with the predilections manifested in the choice condition. This suggests that whatever compassionate dispositions the subjects had were not especially robust. Now I've said that dissonance between behavior and conviction in the face of *extreme* situational pressures should not be taken as evidence against notions of robust traits, since any psychologically plausible theory of character acknowledges limits to fortitude. How "extreme" is the Milgram situation?

Flanagan (1991: 298) remarks that "Milgram's subjects wanted out and were disposed to get out but were *not allowed out*" (my emphasis). But as Milgram (1963: 376) noted, obedience was effected under no threat of punishment or material loss,[57] nor by unambiguously coercive manipulation of the sort found in torture and thought reform. Perhaps it is true that the Milgram paradigm employed coercive mechanisms less obvious than gun and lash, but this doesn't cause me much concern. For my argument requires only that the effects of situational stimuli often seem quite disproportionate to their intuitive magnitude, and such disproportion clearly obtained between the experimenter's instructions and the shocking behavior that they produced, even if one is inclined to insist that the experimental milieu somehow imbued the instructions with subtle coercive powers.

It is true that many of Milgram's experiments took place at the imposing institution of Yale University; perhaps obedience was effected through institutional intimidation. But obedience did not significantly drop when Milgram (1974: 55) relocated his experiment from an impressive laboratory to a rather unprepossessing basement, nor when the experiment was moved from Yale to a dingy "Research Institute" in a run-down section of Bridgeport, Connecticut (Milgram 1974: 66–70).[58] Diminishing the trappings of institutional authority did not significantly decrease obedience; institutional intimidation is at best a very partial explanation of the data. More generally, how much power should the experimenter be thought to have over

volunteer subjects in a short experiment? The experimenter certainly occupied a position of relative power, but his actual coercive tools were sorely limited, and his position was highly transient.

Nevertheless, the experiment's authority structure may have functioned to assure subjects that it was the experimenter, and not they, who bore responsibility for any negative outcome; indeed, for at least some trials, the experimenter explicitly provided the subject with assurances to that effect (Milgram 1974: e.g., 76).[59] Whether explicitly stated or implicit in the experimental dynamic, such perceived absolutions may have helped secure obedience; individuals who believed that they did not bear responsibility for the proceedings might have been more likely to go along with them. In fact, Milgram's (1974: 203–4) analysis indicates that defiant subjects saw themselves as somewhat more responsible, and obedients saw themselves as slightly less responsible, than the experimenter, so there is at least something to the thought that obedience was facilitated by perceptions of diminished responsibility. But this thought takes us only so far: Obedients tended to see themselves as *sharing* responsibility, not as *absolved of* responsibility. Despite the experiment's authority structure, obedients saw themselves as at least partly responsible actors in proceedings that they believed to be – as their manifest anxiety attests – morally objectionable.

Perhaps more potent than subjects' perceptions of authority was their "stepwise" progression through increasing shock levels.[60] The subject is first asked to do something seemingly rather trivial, administering a very slight shock, followed by only a relatively slight increase in voltage each time. If the subject eventually balks, he is faced with a "justification problem" (Flanagan 1991: 297; cf. Sabini and Silver 1982: 70): Why is it wrong to administer this level of shock and not the shocks previously administered? Such justification was available at only one point in the experiment, when the victim first withdrew his implied consent (at 150 volts, level 10); in fact, for most permutations of the experiment, this was the single point at which most defiance occurred (see Ross 1988: 103).[61] On the other hand, the verbal designations on the shock generator would seem to provide some justification for noncompliance: Why wouldn't it seem reasonable to break off between "strong shock" and "very strong shock," for example?

The story is not quite so depressing as it sounds. Ross (1988: 103) imagines a "panic button" placed on Milgram's shock generator together with prominent instructions from a "human subjects committee" stating that the subject should push the button if he wants to stop. Actual human subjects review boards would very likely prohibit putting matters to the test, but Ross conjectures that obedience rates would be much lower than those obtained by Milgram, because the panic button would provide a situational "channel" facilitating subjects acting on their distress (see Flanagan 1991: 297). This seems exactly right. Milgram's lesson is not simply that situational pressures may induce particular *undesirable* behaviors, but that situational

pressures may induce particular behaviors, *period.* Situational sensitivity is not always a bad thing. But in bad situations, it may very well result in bad behavior.

The Stanford Prison Experiment

In the early 1970s, Zimbardo and colleagues devised a "functional representation" of an American prison in the basement of the Stanford University psychology building (Haney et al. 1973: 71–3).[62] Male college students with no history of crime, emotional disability, physical handicap, or intellectual and social disadvantage were selected from a pool of 75 applicants; those chosen were "judged to be most stable (physically and mentally), most mature, and least involved in anti-social behavior" (Haney et al. 1973: 71–3).[63] The 21 participants were randomly assigned the role of "prisoner" or "guard"; prisoners were confined 24 hours a day in a simulated penitentiary complete with barred cells and a small closet for solitary confinement, which became known as the "Hole" (Haney et al. 1973: 72–3). This is what happened.

Five prisoners were released prematurely due to "extreme emotional depression, crying, rage and acute anxiety," symptoms that developed as early as two days into the experiment; one subject developed a psychosomatic rash over portions of his body (Haney et al. 1973: 81). Conversely, most of the guards seemed rather to enjoy their roles (Haney et al. 1973: 81). Prohibited by experimenters from employing physical punishment, they improvised all manner of creative sadisms such as requiring prisoners to clean out toilets with their bare hands (Haney and Zimbardo 1977: 208; cf. Faber 1971: 83). On the second day there was a prisoner insurrection quashed by guards hosing down prisoners with fire extinguishers (Zimbardo et al. 1973). At the end of six days, the alarmed investigators terminated the scheduled two-week experiment (Haney and Zimbardo 1998: 709).

It is difficult for academic commentary to adequately portray the impact of this demonstration. Refer instead to the extraordinary film of the experiment (Zimbardo 1992) or subject diaries like this one (quoted in Haney and Zimbardo 1977: 207–9):

Prior to start of experiment
As I am a pacifist and non-aggressive individual, I cannot see a time when I might maltreat other living things.

On day five
This new prisoner, 416, refuses to eat. That is a violation of Rule Two: "Prisoners must eat at mealtimes," and we are not going to have any of that kind of shit. . . . Obviously we have a troublemaker on our hands. If that's the way he wants it, that's the way he gets it. We throw him into the Hole ordering him to hold greasy sausages in each hand. After an hour, he still refuses. . . . I decide to force feed him, but he won't eat. I let the food slide down his face. I don't believe it is me doing it. I just hate him more for not eating [than I hate myself for doing it].[64]

Once again, it appears that persons are swamped by situations. Administration of personality instruments did not uncover evidence of extreme dispositions consonant with the extreme behaviors; all subjects scored within the "normal-average range" (Haney et al. 1973: 89–90). Subjects were administered a Machiavellianism scale for manipulativeness, an F-scale for conventionality and authoritarianism, and the Comrey Personality Inventory, including subscales for trustworthiness, conformity, and stability; there were no significant differences between prisoners and guards on any of these measures (Haney et al. 1973: 81–4).[65] Whatever factors caused the "guards" to behave as guards and the "prisoners" to behave as prisoners, they are not captured on standard personological approaches to differential functioning.[66]

The Stanford study was of unusual design, involving not the controlled manipulation of a small number of variables, as is typical in social psychology experiments, but countless variables only loosely structured by the policies and physical environment of the prison simulation. Since particular variables could not be effectively isolated, the study presents methodological difficulty (Banuazizi and Movahedi 1975: 154; cf. Haney et al. 1973: 77). But is there reason to doubt general claims about the impact of the experimental environment? Most important, are there grounds for suspicion regarding claims as to the relative unimportance of personal variables? Banuazizi and Movahedi (1975: 154–6) argue that the experimental environment was suffused with reminders that participants were not in an actual prison; hence subjects were merely engaging in a sort of role playing and the simulation was not the functional equivalent of an actual prison. But just as in Milgram's experiments, the extreme reactions of subjects strongly imply that they were taking things very seriously: Psychosomatic rashes are not typical results of laboratory role playing (Dejong 1975; cf. Thayer and Saarni 1975). Furthermore, recordings revealed that 90 percent of all conversations between prisoners were related to prison topics such as visitors, escape plans, and guard harassment, which indicates that at least the prisoners were deeply immersed in the simulation (Haney et al. 1973: 86). There are also telling self-reports, such as this one by a former prisoner: "[I]t was a prison to me, it *still* is a prison to me, I don't regard it as an experiment or simulation" (Haney et al. 1973: 88; cf. Zimbardo 1992).

At bottom, the question of whether a prison "reality" was successfully simulated is of little concern to me; for if the experiment is a failure in this regard, its implications are all the more telling. The situational pressures in a failed simulation are plausibly thought *less* extreme than those in a successful simulation would be. Thus, the disproportion between the extremity of situational factors and the extremity of the resulting behavior is *greater* if the experimental environment was not a "functional representation" of a prison, and the situationist message is therefore *strengthened.* Indeed, the experiment's "unreality" is what makes it so shocking. The participants were volunteers in a short-term experiment; unlike individuals in actual corrections systems, this was not "their life." Still, there was a precipitous descent into barbarism.

The Stanford guards, unlike Milgram's teachers, were not under direct orders to maltreat others; much of the abuse resulted from the guards' initiative and creativity – a sort of entrepreneurial cruelty (see Sabini and Silver 1982: 78–9). But like the Milgram subjects, the Stanford guards did not always endorse their behavior. We have already seen the "pacifist" guard's dismay as he force-fed an inmate, and his sentiments accorded with those expressed by other guards: "I was surprised at myself. I was a real crumb" (Faber 1971: 83). These guards reacted to themselves much as observers may react to them – with alarm and disgust. Importantly, their sentiments invoke the experience of wrongdoers outside the lab. Former corrections officer Roscoe Pondexter, nicknamed "Bonecrusher" by fellow guards at California's Corcoran state prison in honor of his brutality toward inmates, lamented, "I was taught better than that" (Stratton 1999; cf. Haney and Zimbardo 1977: 215–9). Vanardo Simpson, who by his own account murdered some twenty women, children, and elderly men during the Vietnam War's My Lai massacre, insisted, "I wasn't raised up to kill" (Sim and Bilton 1989). Were these men any less "average-normal" than the Stanford guards? The time has come to look past the confines of experimental environments.[67]

Character and Genocide

The Evil

During the twentieth century more than 100 million people died violently at the hands of others (Katz 1993: 10). In 1994, some 800,000 Rwandans were murdered in a period of 100 days; the dead accumulated at nearly three times the rate of Jewish deaths in the Holocaust (Gourevitch 1998). So many corpses were decomposing in the rivers feeding Lake Victoria, the second largest body of fresh water in the world, that authorities feared its fish and water would be unsafe for human consumption (Lorch 1994). The Hutu slaughter of the Tutsi in Rwanda was quite unlike the industrial mass murder of Nazism; it was largely accomplished by the laborious and intimate expedient of hacking victims to death with machetes: "Neighbors hacked neighbors to death in their homes, and colleagues hacked colleagues to death in their workplaces. Priests killed their parishioners, and elementary-school teachers killed their students" (Gourevitch 1995). Who were these neighbors, teachers, priests, and why did they do what they did?

The obvious place to look for insight into the psychology of genocide is the enormous literature on the Holocaust. Such inquiry can be undertaken only with trepidation: The lens of history grows cloudy with time, and human beings have limited capacities by which to fathom unfathomable evil. I've space to engage only a fraction of the relevant material, and my omissions are

legion. In particular, I focus on psychological dynamics with little reference to their political and historical contexts, an expedient that inevitably curtails the sweep of my analysis.[68] But the material that I manage to cover is material that moral psychology can ill afford to ignore; an incomplete treatment is preferable to no treatment at all.

For SS doctors at the Auschwitz death camp, an important "duty" was to meet arriving transports of prisoners and decide who would be condemned to forced labor in the camp and who would be condemned to immediate death. On one occasion, a doctor refused to participate in these "selections"; Eduard Wirths, chief medical officer at Auschwitz, was reputed to have remarked, "Finally, a person with character" (Lifton 1986: 198). Given the institutional pressures at work in Auschwitz, and under the Third Reich more generally, it's tempting to explain such refusals by appeal to moral character, just as Wirths did. Yet "virtually all" Auschwitz doctors performed selections (Lifton 1986: 193); did only men of bad character find their way to the camp?

A persistent theme in accounts of the Holocaust is the perpetrators' "ordinariness."[69] Matters could hardly be otherwise. It takes a lot of people to kill 800,000, 6 million, or 100 million human beings, and there just aren't enough monsters to go around. Unfortunately, it does not take a monster to do monstrous things; if this were the case, our history and prospects would be much brighter. A plausible conjecture, just as with Milgram's obedients or the Stanford guards, is that a very substantial percentage of perpetrators in the Holocaust had previously led lives characterized by ordinary levels of compassion.

While this conjecture can account for the dutiful destructiveness of "cogs in the machine," it is less comfortably applied to the zealous cruelties undertaken on the perpetrators' own initiative (Blass 1993: 37; Darley 1995: 133). Some people perpetrated cruelties with more energy than required by even the most morally depraved Nazi job descriptions; therefore, the evil had to come from within, not without. But to argue that the presence of self-initiated cruelty itself secures the conclusion that the perpetrators are pathological or evil, as some commentators seem to, risks begging the question against the hypothesis that normal individuals may engage in aberrant behavior.[70] Moreover, the Prison Experiment, where "normal" guards acted sadistically on their own initiative, fairly directly contravenes this contention. In any event, I shall argue that there is good evidence that many Nazi war criminals are not straightforwardly understood as possessed of uniformly evil dispositional structures; much like Milgram's obedients, there is evidence that they experienced substantial conflict.

It is true that claims for the ordinariness of the war criminals are typically rather impressionistic and not the results of systematic and detailed observation; as I've said, life is not a laboratory. However, there has been some more systematic study: The defendants at the Nuremburg war crimes trial, among them the Luftwaffe's Goering, the diplomat Ribbentrop, and

the security chief Kaltenbrunner, were subject to detailed psychiatric evaluation, most notably in the form of the Rorschach ink-blot test. This material has been interpreted and reinterpreted, with some finding the defendants to be generally pathological (Gilbert 1950: 274, 286; Miale and Selzer 1975: 287) and others doubting such claims (Kelley 1946: 47; Harrower 1976).[71] Once more, the evidence regarding personality variables seems equivocal. However, even if we were confident in the deliverances of instruments like the Rorschach and accepted attributions of pathology in these contested cases, this would not affect the ordinariness thesis much. For conclusions drawn about the Nazi leadership, perhaps the greatest scourges in all history, would not tell us much about the uncounted others who participated in and condoned atrocities. The situationist does not deny the existence of monsters, but she does deny that the explanation of their behavior will be applicable to the generality of cases.

In his indispensable study, Lifton (1986: 4–5) is struck by the banality of the Nazi doctors he interviewed years after the war; yet these rather pedestrian medical men committed, with great regularity over a period of years, acts of unspeakable evil. And it is not obvious that they found their job particularly onerous. According to one prisoner doctor, "They did their work just as someone who goes to an office goes about his work" (quoted in Lifton 1986: 193). But it would be a mistake to think that their "work" seemed unremarkable to them. These men had previously devoted their lives to a humanitarian profession, in some cases with compassion and distinction; in general practice before his assignment to Auschwitz, Wirths secretly treated Jewish patients after it had become illegal for Aryan doctors to do so (Lifton 1986: 386). As one Auschwitz doctor said, "In the beginning it was almost impossible: Afterward it became almost routine" (quoted in Lifton 1986: 195, cf. 199). What explains this transformation?[72]

The obvious, but incorrect, answer appeals to explicitly coercive indoctrination and control. Lifton (1986: 198) maintains that a determined doctor could avoid performing selections without repercussions, while Goldhagen (1996: 379; cf. Browning 1992: 170–1) more combatively asserts that "it can be said with certitude that never in the history of the Holocaust was a German, SS man or otherwise, killed, sent to a concentration camp, jailed, or punished in any serious way for refusing to kill Jews." Goldhagen's case seems to me overstated, but I think it fair to conclude that explicit coercion was not a necessary condition for atrocities.[73] Still, it is a mistake to count many perpetrators as "willing executioners" if willing means "eager," rather than "not explicitly coerced."

Remember the conflict exhibited by Milgram's subjects and the Stanford guards; a likely explanation is that the subjects had previously internalized ordinary canons of decency, or to put it another way, they possessed an ordinary complement of compassion.[74] If the Nazi war criminals manifested similar tension, it would, analogously, be evidence of their "ordinariness."

However, many war criminals did not appear conflicted. These individuals may have been relatively untroubled murderers; and this is an important disanalogy with conflicted subjects in experimental work on destructive behavior (Sabini and Silver 1982: 60; cf. Katz 1993: 42). But while the experimental subjects had only the supports contained in the experimental milieu for reassurance, Nazi war criminals operated in an all-encompassing institutional context, with the support of peers and superiors, as well as what must have seemed the tacit approval of countless passive bystanders. These pervasive networks of social reinforcement could mute conflict quite effectively, much more so than any experimental pretense; some percentage of those committing atrocities in such an environment could be expected to do so with little in the way of misgivings.

Nevertheless, many war criminals did exhibit conflict. Major Trapp, commander of Reserve Police Battalion 101, a unit that slaughtered Jews in occupied Poland, was reported to have wept after issuing murderous commands (Browning 1992: 58).[75] Among the men who carried out Trapp's orders, heavy drinking was commonplace, for as one (nondrinking) policeman put it, "such a life was quite intolerable sober" (Browning 1992: 82). Nazi doctors likewise reported drinking excessively when performing selections (Lifton 1986: 193); and the same goes for the SS *Einsatzgruppen* death squads the Reich sent east to murder Jews in conquered territories. The *Einsatzgruppen* shot thousands of Jews in the back of the neck, one by one, so there was very close contact with the victims. They were apparently expected to work for only an hour at a time, despite the fact that this task was not physically demanding, and they were liberally provided with alcohol (Sabini and Silver 1982: 73–4).[76] It is worth noting that Nazi propaganda sometimes took the form of exhortations to onerous but necessary work; evidently the masses were not expected to flock eagerly to their genocidal calling (see Katz 1993: 69). None of this is to deny the undeniable: *These people did profoundly evil things.* But it is to raise some doubts about how enthusiastically they did so.

According to Lifton (1986: 193–213), the Auschwitz doctors underwent an intensive socialization process in order to effect their "adaptation" to life in the death-world of the camp. Doctors frequently drank heavily together and often expressed dissatisfaction with camp practices, but these protests eventuated in group rationalizations; the alcoholic therapy sessions were a means for the doctors to establish consensual validation for behaviors that were strongly dissonant with precamp values (Lifton 1986: 195–7). In addition, there may have in some cases been a system of mentoring, where a doctor new to the camps was taken under the wing of a camp veteran to facilitate his assimilation (Lifton 1986: 310–11).

The fact that the camps were more or less closed environments also helped to facilitate compliance (see Katz 1993: 26), much as the isolation of Milgram's subjects facilitated compliance in the experimental milieu. As Lifton (1986: 196) has it, the "Auschwitz reality" became for doctors the "baseline for all else"; immersed in the camp's institutional structure, it grew increasingly difficult for doctors to adopt and maintain a perspective critical of its governing beliefs and values.

Moreover, just as the stepwise progression of experimental demands left the Milgram subjects with weakened rationales for resistance, the development of the Holocaust into full-blown genocide might itself be thought of as a stepwise progression that developed over a period of years from the first economic sanctions against the Jews to the "final solution" of the extermination camps (Sabini and Silver 1982: 70–1; cf. Katz 1993: 37). With the passage of time, what was once unthinkable became unremarkable; persons and nations alike are subject to "moral drift" – a slide into evil as individuals and groups are gradually acclimated to destructive norms (Sabini and Silver 1982: 78).

Unfortunately, the Nazis were not unique in their ability to facilitate this drift: Governments have very often acculturated people to their dirty work with considerable success. Haritos-Fatouros (1988) interviewed torturers employed by the military dictatorship of Greece during 1967–74. These torturers were not made overnight; after months of brutal training, the perpetrators were gradually desensitized to torture, first interacting with prisoners in relatively innocuous ways, then observing torture, and finally themselves becoming full-fledged torturers (Haritos-Fatouros 1988: 1114–17). This is not to say that such training ignores individual differences; only 1.5 percent of recruits were ultimately selected to become torturers (Haritos-Fatouros 1988: 1114). It is therefore possible that the training was effective by dint of identifying the most sadistic subset of trainees, but the sickening probability is that it could have worked even if the personnel selection failed completely in this regard: Haritos-Fatouros (1988: 1119) concludes that in the right circumstances anyone may become a torturer.[77]

If the foregoing is right, many Nazi war criminals exhibited a kind of diachronic fragmentation: Their behavior during the Holocaust was inconsistent with antecedently manifested dispositions. But there is also evidence of synchronic fragmentation, where war criminals exhibited inconsistent dispositions over temporally limited periods within the problematic environment. Once again, the Auschwitz doctors are illustrative. Eduard Wirths was described by prisoners in terms such as "kind," "decent," and "honest," but he was also the man who closely administered the camp's system of selections and mass murders during the years when most murders were committed (Lifton 1986: 384). According to camp survivors, Wirths could exhibit compassion and act to save lives, but he also participated in inhumane medical experiments and zealously executed his bureaucratic role in mass murder (Lifton 1986: 386–92, 401–3). Yet Wirths may have been the only Auschwitz doctor who did not personally enrich himself through graft, and he was devoted to his wife and family (Lifton 1986: 384, 395–9).

The behavior of Josef Mengele struck prisoners as similarly paradoxical:

> He was capable of being so kind to the children, to have them become fond of him, to bring them sugar, to think of small details in their daily lives, and to do things we would genuinely admire... And then, next to that, ... the crematoria smoke, and these children, tomorrow or in a half hour, he is going to send them there. Well, that is where the anomaly lay. (quoted in Lifton 1986: 337)

Mengele surely earned his infamy: Prisoners remembered him as the "most active" of all the Nazi doctors, from whom he perhaps distinguished himself by the frequency of his direct killing and the flamboyance of his cruelty (Lifton 1986: 341–2). But he did not present as a unity: A prisoner doctor referred to Mengele as "*l'homme double*" (Lifton 1986: 375). Such stories abound. The trait common to all camp guards, says Todorov (1996: 141), was gross inconsistency; Arendt (1966: xxix) concluded that almost all SS guards could claim to have saved lives.[78] As Levi (1989: 56) put it, "Compassion and brutality can coexist in the same individual and in the same moment, despite all logic. . . ."[79] Even for the worst of people, dispositional structures are not evaluatively integrated; they defy the logic of characterological psychology. But let us be clear; if evil is as evil does, the Nazis were the most evil of men. But their evil, I contend, is not easily understood as a function of global character structures.

The Good

The Holocaust saw the worst of human history, but it also saw the very best: The rescuers who risked everything to help Jews avoid persecution. The number of rescuers is not easily estimated, but it was doubtless a tiny fraction of the relevant population – perhaps 50,000 to 500,000 out of nearly 700 million people living in Nazi-occupied territories (Oliner and Oliner 1988: 1–2; Gushee 1993: 373; Fogelman 1994: xvi). Such extraordinary behavior prompts explanations in terms of individual dispositional differences. Rescuers and nonrescuers were often in close proximity and in similar circumstances; therefore, the explanation of why one person helped and her neighbor did not must proceed by looking to the differences in persons rather than the differences in situations. As I've said, there must be something to this style of argument. But once again, the evidence is less than clear.

Rescuer studies are typically based on postwar interviews with rescuers and beneficiaries, sometimes with nonrescuers serving as a control group (e.g., Tec 1986; Oliner and Oliner 1988; Fogelman 1994; Monroe 1996).[80] Although reports of the rescuing behavior are corroborated, much of the information in these studies is based on self-reports. Here, the usual worries about self-reports are exacerbated by concerns about the accuracy of memory, given that the events of interest are often being recounted forty or fifty years after the fact. How exactly to interpret this research is an important question for me, because even researchers who are skeptical of personological accounts of the war criminals have found analogous approaches to the rescuers more compelling (Blass 1993: 40). Nevertheless, systematic investigation offers only equivocal support for personological approaches.

With regard to sociocultural factors like education, occupation, and income, the results are mixed, and no orderly pattern emerges.[81] For example, there is no decisive reason to think that rescuers were especially religious when compared with nonrescuers (Oliner and Oliner 1988: 156, 289; Monroe 1996: 121–2; but see Tec 1986: 145). Nor is there conclusive evidence as to whether rescuers were more likely to feel themselves socially marginal or independent, with some researchers identifying such a trend (Tec 1986: 154) and others failing to find such a result (Oliner and Oliner 1988: 176; Fogelman 1994: 329n2). These are striking nonfindings: Darley and Batson (1973) notwithstanding, we might expect that religious commitment might tend people toward helping and also that "rugged individualists" might better resist the seductions of a pernicious mass movement. On the other hand, investigators seem to agree that rescuers are possessed of a distinctive moral outlook, variously characterized as involving a heightened sense of social responsibility and "extensivity" (Oliner and Oliner 1988: 173, 249, 299), a deep concern with "humanistic values" (Fogelman 1994: 253, 274), or a feeling of "shared humanity" with others (Monroe 1996: 213–6; cf. Tec 1986: 176).

If the rescuers' moral outlook did in fact exhibit a characteristic regard for others, we face a familiar question: How is this sort of attitude related to behavior? Some researchers have concluded that the rescuers' attitudes, as inferred from their self-reports, indicate the existence of an "altruistic personality" with reliable behavioral implications (Oliner and Oliner 1988: 186, 221–2; Monroe 1996: 147–9). Many rescuers performed numerous acts of rescue over a period of years, and the interviews frequently seem to suggest a lifelong practice of prosocial behavior. Yet this is not decisive evidence of the behavioral consistency that globalist conceptions of character demand, for at this point the limitations of self-report methodologies loom large. There is no reason to doubt the rescuers' word, but there is also little reason to think that their recollections amount to anything like a systematic sampling of their behavior. Consider also the conversational dynamics: In interviews about rescue, it seems likely that helping behavior would be the focus for both investigator and interviewer. In fact, some rescuers exhibited strong inconsistencies. Oskar Schindler saved over a thousand Jews in Poland from deportation and murder, but he was also a manipulative, hard-drinking, and womanizing war profiteer who did not particularly distinguish himself either before or after the war.[82] There are even cases of lifelong anti-Semites becoming rescuers (Tec 1986: 99–109). One begins to suspect that rescuer behavior was something of a mixed bag, just as it is for the vast majority of