5

THE PSYCHOLOGY OF POWER

To the Person? To the Situation? To the System?

Philip G. Zimbardo

Evil is intentionally behaving—or causing others to act—in ways that demean, dehumanize, harm, destroy, or kill innocent people. This behaviorally focused definition makes a leader responsible for purposeful, motivated actions that have negative consequences to other people. It excludes accidental or unintended harmful outcomes, as well as the broader, generic forms of institutional evil, such as poverty, prejudice, or destruction of the environment by agents of corporate greed. But it does include responsibility for marketing and selling products with known disease-causing, death-dealing properties, such as cigarette manufacturers, or other drug dealers. It also extends beyond the agent directly accountable for aggression to encompass those in positions of authority whose orders or plans are carried out by subordinates. This is true of political and military leaders, such as Hitler, Stalin, Mao, Pol Pot, Idi Amin, Saddam Hussein, and other tyrants.

The same human mind that creates the most beautiful works of art and extraordinary marvels of technology is equally responsible

This chapter is a modified version of my presentation for the Stanford Center on Ethics Symposium on Moral Leadership February 25, 2005. It relies on my recent chapter, "The Social Psychology of Good and Evil: Understanding Our Capacity for Kindness and Cruelty," published in *The Social Psychology of Good and Evil*, ed. Arthur Miller (New York: Guilford Press, 2004), 21–50.

for the perversion of its own perfection. This most dynamic organ in the universe has been a seemingly endless source for creating ever more vile torture chambers and instruments of horror: the concentration camps of the Third Reich, the "bestial machinery" of Japanese soldiers in their rape of Nanking, and the recent demonstration of "creative evil" of 9/11 by turning commercial airlines into weapons of mass destruction.[1] How can the unimaginable become so readily imagined?

My concern centers around how normal individuals can be recruited, induced, and seduced into behaving in ways that could be classified as evil and the role of leaders in that process. In contrast to the traditional approach of trying to identify "evil people" to account for the evil in our midst, I will focus on the central conditions that underpin the transformation of good, or average, people into perpetrators of evil.

Locating Evil Within Particular People: The Rush to Judgment

"Who is responsible for evil in the world, given that there is an all-powerful, omniscient God who is also all-Good?" That conundrum began the intellectual scaffolding of the Inquisition in the sixteenth and seventeenth centuries in Europe. As revealed in *Malleus Maleficarum*, the handbook of the German Inquisitors from the Roman Catholic church, the inquiry concluded that the devil was the source of all evil. However, these theologians argued that the devil works his evil through intermediaries, lesser demons, and, of course, human witches. Therefore, the hunt for evil focused on those marginalized people who looked or acted differently from ordinary people, who might, under rigorous examination or torture, be exposed as witches and then put to death. As historian Ann Barstow notes, the victims were mostly women who could readily be exploited without sources of defense, especially when they had resources that could be confiscated.[2]

Paradoxically, this early effort of the Inquisition to understand the origins of evil and develop responses to evil instead created new

forms of evil. It exemplifies the risk of simplifying complex problems by blaming individual perpetrators.

The same risk emerged following World War II, when a team of psychologists sought to make sense of the Holocaust and the broad appeal of national fascism and Hitler.[3] Their focus was on the authoritarian personality: a set of traits underlying the fascist mentality. However, what they overlooked was the host of processes operating at political, economic, societal, and historical levels of analysis that influenced so many millions of individuals to revere their dictator and hate Jews.

This tendency to explain observed behavior by reference to dispositions, while ignoring or minimizing the impact of situational variables, is what Stanford psychologist Lee Ross has called the fundamental attribution error.[4] We are all subject to this dual bias of overemphasizing dispositional analyses and underemphasizing situational explanations. We succumb to this effect because so much of our education, training, and law enforcement are geared toward a focus on individual orientations. Dispositional analyses are a central operating feature of cultures that are based on individualistic rather than collectivist values.[5] Thus, it is individuals who receive praise, fame, and wealth for achievement and are honored for their uniqueness, but it is also individuals who are blamed for the ills of society. Our legal, medical, educational, and religious systems are all founded on principles of individualism.

Dispositional analyses always include strategies for behavior modification to assist deviant individuals, by education or therapy, or to exclude them from society by imprisonment, exile, or execution. Locating evil within selected individuals or groups has the virtue of rendering society blameless. The focus on people as causes for evil then exonerates social structures and political decision making for contributing to underlying conditions that foster evil: poverty, racism, sexism, and elitism.

Most of us take comfort in the illusion that there is an impermeable boundary separating the evil (them) from the good (us). That view leaves us with less interest in understanding the motivations and circumstances that contributed to evil behavior. But in

fact, as is clear from the Russian novelist Alexander Solzhenitsyn, a victim of persecution by the Soviet KGB, that the line between good and evil lies in the center of every human heart.

It has been my mission as a psychologist to understand better how virtually anyone could be recruited to engage in evil deeds that deprive other human beings of their lives, dignity, and humanity. So I have always begun my analyses of all sorts of antisocial behavior, even the most horrendous instances of evil, with the question, "What could make me do the same thing?" And furthermore, I wonder what set of situational and structural circumstances empowered others—maybe similar to me—to engage in deeds that they too once thought were alien to their nature.

The answers underscore the extraordinary capacity of the human mind to adapt to virtually any environmental circumstances in order to survive, create, and destroy as necessary. We are not born with tendencies toward good or evil but with mental templates to do either, more gloriously than ever before, or more devastatingly than ever experienced before—as the terrorist attacks of September 11, 2001, revealed. It is only through the recognition of our shared human condition that we can acknowledge vulnerability to situational forces. Although the research summarized in this chapter focuses on identifying processes by which ordinary people can be seduced or initiated into engaging in evil deeds, the time has also come to better understand how ordinary people can resist such forces and promote prosocial behavior. If we want to develop mechanisms for combating transformations of good people into evil perpetrators, it is essential to learn first the causal mechanisms underlying those behavior changes.

Blind Obedience to Authority: The Milgram Investigations

Stanley Milgram developed an ingenious research procedure to demonstrate the extent to which situational forces could overwhelm individual will to resist.[6] He shocked the world with his unexpected finding of extremely high rates of compliance to the demands of an

authority figure to deliver apparently dangerous electric shocks to an innocent victim.[7] He found that about 67 percent of research participants went all the way up to the top shock level of 450 volts in attempting to "help" another person learn appropriate behaviors. Milgram's study revealed that ordinary American citizens could so easily be led to engage in "electrocuting a nice stranger" as the Nazis had been led to murder Jews.

After this initial demonstration with Yale College students, Milgram went on to conduct eighteen experimental variations on more than a thousand subjects from a variety of backgrounds, ages, and educational levels. In each of these studies, he varied one social psychological variable and observed its impact on the extent of obedience by a subject to shock the "learner-victim," who pretended to be suffering. The data told the story of the extreme pliability of human nature. Almost everyone could be totally obedient, or almost everyone could resist authority pressures; it depended on situational differences. Milgram was able to demonstrate that compliance rates could soar to 90 percent of people who delivered the maximum 450 volts to the learner-victim, or could be reduced to less than 10 percent of total obedience by introducing one variable into the compliance recipe.

Want maximum obedience? Provide social models of compliance by having participants observe peers behaving obediently. Want people to resist authority pressures? Provide social models of peers who rebelled. Interestingly, almost no one shocked the learner-victim when he actually asked to be shocked. They refused authority pressure when the target seemed to be a masochist. In each of the other variations on this diverse range of ordinary Connecticut citizens, low, medium, or high levels of compliant obedience could be readily elicited as if one were simply turning a human nature dial.

What is the expected rate of such obedience in the Milgram setting according to experts on human nature? When forty psychiatrists were given the basic description of this experiment, their average estimate of the number of United States citizens who would give the full 450 volts was only 1 percent. Only sadists would engage in such behavior, they believed. These experts on human behavior were

totally wrong because they ignored situational determinants. Their training in psychiatry had led them to overly rely on dispositional explanations. This is a strong instance of the operation of the fundamental attribution error in action.

Ten Steps to Creating Evil Traps for Good People

What were the procedures in this research paradigm that seduces many ordinary citizens to engage in such apparently abusive behavior? These procedures parallel compliance strategies used in many real-world settings by "influence professionals" such as salespeople, cult recruiters, and national leaders.[8]

These are the influences that will lead ordinary people to do things they originally believe they would not:

1. Offering an ideology that justifies any means to achieve a seemingly desirable goal. In clinical experiments like Milgram's, the rationale is helping people improve their memories. For nations, it is often a "threat to national security" that justifies going to war or suppressing dissident political opposition. This is the excuse that fascist governments and military juntas have frequently used to destroy socialist or communist opposition. When citizens fear that their national security is being threatened, they are all too willing to surrender their basic freedoms. In the United States, the fear of terrorism has led many citizens to accept torture of prisoners as a necessary tactic for securing information that could prevent further attacks. As research by Susan Fiske and her colleagues indicates, that reasoning contributes to ordinary people's willingness to torture enemy prisoners in contexts like Abu Ghraib prison.[9]
2. Arranging some form of contractual obligation, verbal or written.
3. Giving participants meaningful roles to play (teacher, student) that carry with them previously learned positive values and response scripts.

4. Presenting basic rules to be followed that seem to make sense prior to their actual use, but then can be arbitrarily used to justify mindless compliance. Make the rules vague, and change them as necessary.
5. Altering the semantics of the act, the actor, and the action (from hurting victims to helping learners); replace reality with desirable rhetoric.
6. Creating opportunities for diffusion of responsibility or suggesting that others will be responsible or that the actor will not be held liable.
7. Starting the path toward the ultimate evil act with a small, insignificant first step (only 15 volts in the Milgram experiment).
8. Gradually increasing steps on the pathway to abuse, so that they appear no different from prior actions. Increases of only 30 volts presented no noticeable difference in harm to the Milgram participants.
9. Changing the nature of the influence authority from initially "just" and reasonable to "unjust" and demanding, which elicits confusion but continued obedience.
10. Making the exit costs high by allowing the usual forms of verbal dissent (that make people feel good about themselves), while insisting on behavioral compliance ("I know you are not that kind of person; just keep doing as I tell you").

Such procedures can prepare people psychologically to do the unthinkable.

On Being Anonymous: Deindividuation and Destructiveness

The idea for my doing research that used anonymity as an independent variable in the study of aggressive behavior came not from a psychological theory but rather from William Golding's Nobel

prize–winning novel, *Lord of the Flies*. It chronicles the transformation of good British, Christian choir boys into murderous little beasts by centering on changes in external physical appearance that lead to changes in mental states and behaviors.[10] Painting their faces, changing their outward appearance, made it possible for some boys to kill a pig for food. Once that alien deed of killing another creature was accomplished, they could continue to kill with pleasure, both animals and people alike. In a real-world setting, could similar changes in external appearance affect internal and behavioral processes? That is the question I answered with a set of experiments and field studies on the psychology of deindividuation.[11]

The basic procedure involved having young women deliver a series of painful electric shocks to each of two other young women whom they could see and hear in a one-way mirror before them. Half the subjects were randomly assigned to a condition of anonymity, or deindividuation, and half to one of uniqueness, or individuation. The four college student subjects in each deindividuation cluster were treated as members of a group, not as individuals; their appearances were concealed by hoods, and their names were replaced by numbers. The comparison group consisted of individuals who wore name tags and were made to feel unique. Both sets of subjects were asked to make the same responses of shocking each of two woman "victims" over the course of twenty trials. The cover story was that these "victims" were trying to be creative under stress, so the job of our subjects was to provide stress through painful electric shocks while I, as the experimenter, gave them the creativity test. Unlike the Milgram paradigm, there was no authority insisting on their aggressive behavior because I was in the adjacent room, seen in the two-way observation mirror by the subjects along with each of the two alleged women in the creativity study. The dependent variable was the duration of shock administered, not shock level intensity. Again, no shocks were actually administered, although all participants believed they were doing so and that the two victims were suffering.

The results were clear: women in the deindividuation condition delivered twice as much shock to both victims as did the women in the individuated comparison condition. Moreover, they shocked both victims, the one previously rated as pleasant and the other unpleasant victim, more over the course of the twenty trials, while the individuated subjects shocked the pleasant woman less over time than they did the unpleasant one. One important conclusion flows from this research and its various replications and extensions, some using military personnel from the Belgian army: anything that makes people feel anonymous reduces a sense of accountability and increases their propensity to evil under situations inviting violence.

Cultural Wisdom: How to Make Warriors Kill in Battle But Not at Home

Some societies go to war without having young male warriors change their appearance, while others always include ritual transformations of appearance by painting or masking the warriors (as in *Lord of the Flies*). Does that change in external appearance make a difference in how warring enemies are treated? Harvard anthropologist John Watson posed that question.[12] The Human Area Files were his data source to compare societies with different practices in preparing warriors for war and the extent to which they killed, tortured, or mutilated their victims.

The results are striking confirmation of the prediction that anonymity promotes destructive behavior when permission is also given to behave in aggressive ways that are ordinarily prohibited. Of the twenty-three societies for which these two data sets were present, 80 percent (twelve of fifteen) of societies in which warriors changed their appearance were among those noted as most destructive, while that was true of only one of the eight where the warriors did not change appearance before going to battle. Ninety percent of the time when victims were killed, tortured, or mutilated, it was by warriors who had first changed their appearance.

A key ingredient in transforming ordinarily nonaggressive young men into warriors who can kill on command is to change how they look. War is about old men persuading young men to harm and kill other young men. It becomes easier to do so if they first alter their usual external facade by putting on uniforms or masks or painting their faces. As they acquire anonymity, they lose their usual internal focus of compassion and concern for others. When the war is over, the warriors can return to their peaceful status, encouraged by removing their uniform, taking off the mask, and resuming their former facade.

Moral Disengagement and Dehumanization

Psychologist Al Bandura has developed a model of moral disengagement that specifies the conditions under which anyone can be led to act immorally, even those who usually ascribe to high levels of morality.[13] The model outlines a set of cognitive mechanisms that alter a person's (1) perception of the reprehensible conduct (engaging in moral justifications, making palliative comparisons, using euphemistic labeling); (2) sense of the detrimental effects of that conduct (minimizing, ignoring, or misconstruing the consequences); (3) sense of responsibility for the link between reprehensible conduct and its detrimental effects (displacing or diffusing responsibility); and (4) view of victims (dehumanizing or blaming them).

Bandura and his colleagues designed a powerful experiment that is an elegantly simple demonstration of the power of dehumanizing labels.[14] It reveals how easy it is to induce normal, intelligent individuals to accept a dehumanizing label of other people and then to act aggressively based on that classification. A group of four participants were led to believe they were overhearing a research assistant tell the experimenter that the students from another college were present to start the study in which they were to deliver electric shocks of varying intensity (allegedly as part of a group problem-solving study). In one of the three randomly assigned conditions, the subjects overheard the assistant say to the experimenter that the other students seemed "nice." In a second condition, they heard that the other students seemed like "animals," In a third variation, the assistant did not label the students.

This situational manipulation clearly affected behavior. Experimental subjects gave the most shock to those labeled "animals," and their shock level increased over the ten trials. Those labeled "nice" were given the least shock, and the unlabeled group fell between these extremes. Thus, a single word, *animals*, was sufficient to induce intelligent college students to treat others as if they deserved to be harmed.

What is also of interest is the progressive nature of abuse. On the first trial, there is no difference across the three experimental situations in the level of shock administered, but with each successive opportunity, the shock levels diverge. Those shocking the so-called animals shocked them more and more over time, a result comparable to the escalating shock level of the deindividuated female students in my earlier study. That increase in aggression over time, with practice or with experience, illustrates a self-reinforcing effect of abuse. Perhaps its appeal is not so much in inflicting pain to others as in the sense of power and control in such a situation of dominance.

A more positive finding was that individuals receive more respectful treatment if someone in authority labels them positively. Those perceived as "nice" were least harmed. There is an important message here about the power of words, to be used for good or evil.

Suspension of the Usual Cognitive Controls Guiding Moral Action

Part of what is necessary to get good people to engage in evil is to minimize or reorient normal cognitive control processes. That process suspends conscience, self-awareness, sense of personal

responsibility, obligation, commitment, liability, morality, and analyses in terms of costs and benefits of given actions. The two general strategies for accomplishing this objective are reducing cues of social accountability of the actor (no one knows who I am or cares to) and reducing concerns for self-evaluation. The first strategy minimizes concerns for social evaluation and social approval by making the actor feel anonymous. It works in an environment that masks identity and diffuses personal responsibility across others in the situation. The second strategy stops self-monitoring by relying on tactics that alter one's state of consciousness (such as by drugs, strong emotions, or hyperintense activity) and projecting responsibility outward onto others.

The Hostile Imagination Created by Faces of the Enemy

The importance of situational influences is also apparent in the ways that nations prepare soldiers to engage in wars and prepare citizens to support the risks of going to war, especially a war of aggression. This difficult transformation is accomplished by a special form of cognitive conditioning. Images of the "enemy" are created to prepare the minds of soldiers and citizens to hate those who fit the category. This mental conditioning is the military's most potent weapon.

Archetypes of the enemy are created by propaganda. These visual images often create a societal paranoia that focuses on the threats these enemies pose for women, children, homes, and the religion of the soldier's nation, way of life, and so forth. Psychologist Sam Keen's analysis of this propaganda on a worldwide scale reveals that there are a select number of categories used by "homo hostilis" to invent an evil enemy in the minds of good members of righteous tribes.[15] The enemy is aggressor; faceless; rapist; godless; barbarian; greedy; criminal; torturer; death; a dehumanized animal, or just an abstraction. Alternatively, there is a vision of the enemy as worthy, combatant, a powerful opponent to be crushed in "mortal combat"—as in the video game of the same name.

Can Ordinary Old Men Become Murderers Overnight?

One of the clearest illustrations of the transformation of ordinary people into agents of mass atrocities comes from the chronicle of Nazi genocides by British historian Christopher Browning.[16] In March 1942 about 80 percent of all victims of the Holocaust were still alive; eleven months later, about 80 percent were dead. In this short period of time, the Endlösung (Hitler's Final Solution) was energized by mass murder squads in Poland. This genocide required mobilization of a large-scale killing machine at the same time as able-bodied German soldiers were needed on the collapsing Russian front. Since most Polish Jews lived in small towns and not the large cities, the question that Browning raised about the German High Command was, "Where had they found the manpower during this pivotal year of the war for such an astounding logistical achievement in mass murder?"[17]

His answer came from archives of Nazi war crimes, in the form of the activities of Reserve Battalion 101, a unit of about five hundred men from Hamburg, Germany. They were family men from working-class and lower middle-class backgrounds and were too old to be drafted into the army. They had no military or police experience. They were just raw recruits sent to Poland without warning of, or any training in, their secret mission: the total extermination of all Jews living in the remote villages of Poland. In just four months they had shot to death at point-blank range at least thirty-eight thousand Jews and deported another forty-five thousand to the concentration camp at Treblinka.

Initially their commander acknowledged that this was a difficult mission and any individual could refuse to execute these men, women, and children. Records indicate that at first, about half the German police reservists refused and let others engage in the mass murder. But over time, social modeling processes took their toll, as did guilt-induced persuasion by those who had been doing the shooting. By the end of their journey, up to 90 percent of the men

in Battalion 101 were involved in the shootings, even proudly taking close-up photographs of their killings. Like the guards at Abu Ghraib Prison, these policemen put themselves in their "trophy photos" as proud exterminators of the Jewish menace.

Browning makes clear that there was no special selection of these men. They were as ordinary as can be imagined—until they were put into a situation in which they had official permission and encouragement to act sadistically against those labeled as the "enemy." He also compares the underlying mechanism operating in that far-off land at that distant time to both the psychological processes at work in the Milgram research and the Stanford Prison Experiment discussed below.

Educating Hatred and Destructive Imaginations

The second broad class of operational principles by which otherwise good people can be recruited into evil is through educational and socialization processes that are sanctioned by the government in power, enacted within school programs, and supported by parents and teachers. A prime example is the way in which German children in the 1930s and 1940s were systematically indoctrinated to hate Jews, to view them as the all-purpose enemy of the new German nation.

As the Nazi party rose to power in 1933, no objective took higher priority than the reeducation of Germany's youth. Hitler wrote, "I will have no intellectual training. Knowledge is ruin to my young men. A violently active, dominating, brutal youth—that is what I am after."[18] To teach the youth about geography and race, special primers were created for elementary schools. These "hate primers" were brightly colored comic books that contrasted the beautiful blond Aryans with the despicably ugly caricatured Jew. They sold in the hundreds of thousands. One was titled: *Trust No Fox in the Green Meadows and No Jew on His Oath*. What is most insidious about this kind of hate conditioning is that it included "facts" to be

learned and to be tested on, or from which to practice new penmanship. In the *Trust No Fox*, a series of cartoons illustrates all the ways in which Jews deceive Aryans, get rich and fat from dominating them, and are lascivious, mean, and without compassion for the plight of poor and elderly Aryans. The final scenarios depict the retribution that Aryan children achieve first by expelling Jewish teachers and children from their school, so that "proper discipline and order" can be taught, and then prohibiting them from community areas like public parks, and finally expelling them from Germany. The sign in the cartoon reads ominously, "one-way street." Indeed, it was a unidirectional street that led eventually to the concentration camps and crematoria that were the center pieces of Hitler's Final Solution. Thus, this institutionalized evil was spread, diverting education from its central mission of encouraging critical thinking and opening student minds to new ideas.

The institutionalized evil that George Orwell vividly portrays in *1984*, his fictional account of state dominance, now strikes us as prophetic. For example, there are direct parallels between the mind control strategies and tactics Orwell attributes to "The Party" and those that Reverend Jim Jones used in dominating the members of his religious/political cult, Peoples Temple.[19] Jones orchestrated the suicide/murders of more than nine hundred American citizens in the jungles of Guyana twenty-five years ago, the finale of his grand experiment in institutionalized mind control. Not only did Jones read *1984*, he talked about it often and had a song commissioned by the church's singer entitled "1984 Is Coming," which everyone had to sing at some services.

The Stanford Prison Experiment: Institutional and Systemic Power to Corrupt

My own 1971 prison experiment synthesized many of the processes and variables outlined earlier: anonymity of place and person, dehumanization of victims, and a setting with differentials in control and power. This experiment was designed to extend over a two-week

period to provide our research participants with sufficient time to become fully engaged in their assigned roles of either guards or prisoners. Having participants live in that setting day and night as prisoners, or work there for long eight-hour shifts as guards, would also allow sufficient time for situational norms to develop and patterns of social interaction to emerge, change, and become crystallized. A second feature of this study involved screening all research participants to find individuals as normal as possible: healthy physically and mentally, and without any history in drugs, crime, or violence. These preconditions were essential if we were to untangle the situational versus dispositional knot. A third feature of the study was the absence of any prior training in how to play the randomly assigned roles of prisoner and guard, to leave that up to each subject's prior societal learning. A fourth feature was to make the experimental setting as close to a functional simulation of the psychology of imprisonment as possible.[20]

Central to this mind-set were issues of power and powerlessness, dominance and submission, freedom and servitude, control and rebellion, identity and anonymity, coercive rules and restrictive roles. We gave these social-psychological constructs a practical reality by putting all subjects in appropriate uniforms, using assorted props (handcuffs, police clubs, whistles, signs on doors and halls), replacing corridor hall doors with bars to create prison cells, having no windows or clocks to tell time of day, replacing individual names with numbers (prisoners) or titles for staff (Mr. Correctional Officer, Warden, Superintendent), and giving guards control power over prisoners.

Subjects were recruited from among nearly one hundred who answered our advertisements in the local city newspaper. They were given a background evaluation that consisted of a battery of five psychological tests, personal history, and in-depth interviews. The twenty-four who were evaluated as most normal and healthy in every respect were randomly assigned to the role of prisoner or guard. The student-prisoners underwent a realistic surprise arrest by officers from the Palo Alto Police Department, who cooperated with our plan. The arresting officer took "felons" to the police station for booking, after which they came to the prison in the reconstructed basement of Stanford's Psychology Department. The prisoner's uniform was a smock/dress with a prison ID number. The guards wore military-style uniforms and silver-reflecting sunglasses to enhance anonymity. Data were collected through systematic video recordings, secret audio recordings of conversations of prisoners in their cells, interviews and tests at various times during the study, postexperiment reports, and direct, concealed observations.[21]

In essence, the situational forces overwhelmed individuals' dispositional tendencies. The Evil situation triumphed over the Good people. Our projected two-week experiment had to be terminated after only six days because of the pathology we were witnessing. Normal students were behaving sadistically in their role of guards, inflicting humiliation, pain, and suffering. Some guards even reported they were enjoying doing so. Others had "emotional breakdowns" and stress disorders so extreme that five of them had to be excused within that first week. Those who adapted better to the situation were prisoners who mindlessly followed orders, became blindly obedient to authority, and allowed the guards to dehumanize and degrade them.

I terminated the experiment not only because of the escalating level of violence and degradation by the guards against the prisoners, but also because I became aware of my own personal transformation. I had become a Prison Superintendent, the second role I played in addition to that of principal investigator. I began to talk, walk, and act like a rigid institutional authority figure more concerned about the security of "my prison" than the needs of the young men entrusted to my care as a psychological researcher. In a sense, I consider that the most profound measure of the power of this situation was the extent to which it transformed my own personality. Fortunately, there appeared to be no lasting negative consequences of this powerful experience. At the end of the study, we had extended debriefing sessions of guards and prisoners and periodic checkups over many years to promote a healthy response to the unhealthy environment we had simulated.

The Evil of Inaction

British statesman Edmund Burke aptly observed, "The only thing necessary for evil to triumph is for good men to do nothing." Our usual take on evil focuses on violent, destructive actions, but nonaction can also become a form of evil, when assistance, dissent, or disobedience are called for. Social psychologists heeded the alarm when the infamous Kitty Genovese case made national headlines. As a young woman was being stalked, stabbed, and eventually murdered, thirty-nine people in a housing complex heard her screams and did nothing to help. It seemed obvious that this was a prime example of the callousness of New Yorkers, as many media accounts reported. A counter to this dispositional analysis came in the form of a series of classic studies by Bibb Latane and John Darley on bystander intervention.[22] One key finding was that people are less likely to help when they are in a group, when they perceive others are available who could help, than when those people are alone. The presence of others diffuses the sense of personal responsibility of any individual.

A powerful demonstration of the failure to help strangers in distress was staged by John Darley and Dan Batson.[23] Imagine you are a theology student on your way to deliver the sermon of the Good Samaritan in order to have it videotaped for a psychology study on effective communication. Further imagine that as you are heading from the Psychology Department to the videotaping center, you pass a stranger huddled in an alley in dire distress. Are there any conditions that you could conceive that would not make you stop to be that Good Samaritan? What about the press of time? Would it make a difference if you were late to give that sermon? Most of us would like to believe it would not make a difference, that we would stop and help no matter what the circumstances. That may be particularly the case for theology students, thinking about helping a stranger in distress, which is amply rewarded in the biblical account.

The researchers randomly assigned students of the Princeton Theological Seminary to three conditions that varied how much

time they had to get to the Communication Department to tape their Good Samaritan speeches. The conclusion: don't be a victim in distress when people are late and in a hurry, because 90 percent of them are likely to pass you by. The more time the seminarians believed they had, the more likely they were to stop and help. So the situational variable of time press accounted for the major variance in helping, without any need to resort to dispositional explanations about theology students being callous, cynical, or indifferent, as Kitty Genovese's nonhelpers were assumed to be.

In situations of evil, there are almost always those who know what is going on and do not intervene to help. There were "good" guards in the Stanford Prison Experiment who did no harm to the prisoners, but they never once opposed the demeaning deeds of the bad guards. In the recent Abu Ghraib Prison abuse case, it is clear that many people knew of the abuse, even doctors and nurses, but never intervened to prevent it.[24]

Torturers and Executioners: Pathological Types or Situational Imperatives?

There is little debate but that the systematic torture by men of their fellow men and women represents one of the darkest sides of human nature. Surely, people assume, here is a place where dispositional evil would be manifest. To test that assumption, my colleagues and I focused on Brazil policemen who long tortured "enemies of the state." We began by focusing on torturers, trying to understand both their characters and circumstances, but we had to expand our analytical net to capture their comrades-in-arms, who chose or were assigned to another branch of violence: death squad executioners. They shared a "common enemy"—men, women, and children who, though citizens of their state, even neighbors, were declared by "the authorities" to be threats to the country's national security. Some had to be eliminated efficiently, while others who might hold secret information had to be made to yield it up and confess to their treason.

In carrying out this mission, these torturers could rely in part on the "creative evil" embodied in torture devices and techniques that had been refined over centuries since the Inquisition by religious and later government agents. But the Brazilians added a measure of improvisation to accommodate the particular resistances and resiliencies of the enemies standing before them, claiming innocence, refusing to be intimidated or acknowledge culpability. It took time and emerging insights into human weaknesses for these torturers to become adept at their craft, in contrast to the task of the death squads, who with hoods for anonymity, good guns, and group support, could dispatch their duty swiftly and impersonally. For the torturer, it could never be just business. Torture always involves a personal relationship, essential for understanding what kind of torture to employ, what intensity of torture to use on this person at this time. With the wrong kind or too little, there will be no confession. With too much, the victim dies before confessing. In either case, the torturer fails to deliver the goods. Learning to select the right kind and degree of torture makes rewards abound and praise flow from the superiors.

What kind of men could do such deeds? Did they need to rely on sadistic impulses and a history of sociopathic life experiences to rip and tear the flesh of fellow beings day in and day out for years on end? Were these violence workers a breed apart from the rest of humanity? Or is it conceivable that they could be schooled in sadism by some identifiable and replicable training program? Could a set of external conditions, situational variables that contributed to the making of these torturers and killers, be identified? If their evil deeds were not traceable to inner defects but rather to external forces, such as the political, economic, social, historical, and experiential components of their police training, then we might be able to generalize principles responsible for this remarkable transformation.

Martha Huggins, Mika Haritos-Fatouros, and I interviewed several dozen of these violence workers in depth and have published a summary of our methods and findings.[25] Our results were largely congruent with an earlier study of torturers trained by the Greek military junta.[26] Contrary to conventional wisdom, these torturers

were not motivated by sadistic impulses. Individuals with those tendencies were screened out of the training process because they would not be reliable; they would get off on the pleasure of inflicting pain and not sustain the necessary focus on the goal of confession. From all the evidence we could muster, Brazil's violence workers were not unusual or deviant in any way prior to practicing this new role. Nor did they show any persisting deviant tendencies or pathologies in the years following their work as torturers and executioners. Their transformation was a consequence of the training they were given, group camaraderie, national security ideology, and a belief that socialist-communists were enemies of their state. Torturers were also influenced by their sense of being special, above and better than peers in public service; by the secrecy of their duties; and by the constant pressure to produce desired results regardless of fatigue or personal problems. Such conditions are all too replicable when a nation becomes obsessed with national security and allows fears of terrorism to suspend basic individual rights.

Suicide Bombers: Mindless Fanatics or Mindful Martyrs?

The transformation of Brazilian violence workers is similar to the evolution of Palestinian students to suicide bombers. There have been close to one hundred suicide bombings by Palestinians against Israelis since September 2000. Initially the bombers were all young men, but recently women have joined the ranks. What appears senseless, mindless murder by those under attack is anything but that to those intimately involved. A common assumption is that those bombers are poor, desperate, socially isolated, illiterate young people with no career prospects. That stereotype does not match the actual portraits of these young men and women. Many are students with hopes for a better future, intelligent, attractive, connected with their family and community.

Ariel Merari, an Israeli psychologist, outlines the common steps on the path to these explosive deaths. Leaders of an extremist group first identify particular young people who appear to have an intense

patriotic fervor; declarations at a public rally against Israel or support at some Islamic cause or Palestinian action might set them apart. These individuals are then invited to discuss how serious they are in their love of their country and hatred of Israel; they are asked to translate their commitments into action. Those who are willing join a small group of three to five similar youths who are at varying stages of progress toward becoming agents of death. They learn the tricks of the trade from elders: bomb-making, disguise, and selection of targets. Then they make public their private commitment by preparing a videotape on which they declare themselves to be "living martyrs" for Islam and for the love of Allah. In one hand, they hold the Koran, in the other, a rifle. This video binds them to the final deed, since it is sent home to their family before they execute their final plans. The recruits believe that they will earn a place beside Allah and that their relatives will be similarly blessed because of their martyrdom. A sizable financial payment goes to their family as a gift for their sacrifice.

Their photo is emblazoned on posters that will be put on walls everywhere in the community the moment they succeed in their mission, to become inspirational models. To stifle concerns about the pain from wounds inflicted by exploding nails and other bomb parts, leaders assure them that before the first drop of their blood touches the ground, they will already be seated at the side of Allah, feeling no pain. A further incentive for young males is the promise of heavenly bliss with scores of virgins in the next life. They become heroes and heroines, modeling self-sacrifice to the next cadre of young suicide bombers.

This program relies on various social, psychological, and motivational principles to assist in turning collective hatred into a dedicated, seriously calculated program of indoctrination and martyrdom. It is neither mindless nor senseless, but involves a very different mindset with sensibilities that Americans are not used to witnessing among young adults. A recent television program on female suicide bombers went so far as to describe them as more akin to the girl next door than to alien fanatics. What is so frightening about the

emergence of this new social phenomenon that is spreading, in both the Middle East and recently in Europe, is that so many intelligent young people can be directed toward ending normal, joyous lives in murderous explosive blasts.

To counteract the powerful tactics of these extremist leaders will require providing meaningful life-affirming alternatives to this next generation. It requires new national leadership that explores every negotiating strategy that could lead to peace and not to death. It requires persuading these young people to share their values, their education, and their resources in projects that focus on human commonalities, not differences. Suicide/murder is a gash in the fabric of the human connection that leaders from every nation must unite to prevent. To encourage the sacrifice of youth for the sake of advancing ideologies of the old might itself be considered a form of evil that transcends local politics.

Summing Up Before Moving On

It is a truism in psychology that personality and situation interact to generate behavior, as do cultural and societal influences. Acknowledging the power of situational forces does not excuse the behaviors channeled by their operation. However, a situational perspective provides a knowledge base to shift our attention away from simplistic and ineffective individualistic efforts to change the evildoer and moves our focus toward causal networks that should be modified. Sensitivity to situational determinants of behavior also suggests ways to alter situations of vulnerability. To that end, several related dimensions bear emphasis. First, a range of apparently simple situational factors can function to have an impact on our behavior more compellingly than seems possible. The research outlined here and in other chapters in this book points up the force of influences such as group pressure, social modeling, authoritarian directives, semantic framing, and stereotypical labels.

Second, the situationist approach redefines heroism. When the majority of ordinary people can be overcome by such pressures

toward compliance and conformity, the minority who resist should be considered heroic. Acknowledging the special nature of this resistance points up the need to understand how they are able to withstand compelling pressures.

Third, the situationist approach encourages personal humility. When trying to understand "unthinkable," "unimaginable," "senseless" acts of evil, we should, on this view, avoid embracing the high moral ground that distances us "good folks" from those "bad ones." A situational approach gives others the benefit of attributional charity. It reminds us that any deed, for good or evil, that any human being has ever done, you and I might also do given the same situational forces. If so, it becomes imperative to keep our immediate moral outrage in check and look for the causal factors that could have led individuals in morally reprehensible directions.

An obvious current application of these principles involves the rush to denounce terrorists and suicide bombers instead of also working to understand the psychological, economic, and political conditions that foster such generalized hatred of an enemy nation, including our own. The "war on terrorism" can never be won solely by plans to find and destroy terrorists, since millions of individuals, anywhere, at any time, could become potential terrorist recruits. It is only by understanding the situational determinants of terrorism that programs can be developed to win the hearts and minds of potential terrorists away from destruction and toward creation. This is not a simple task, but it is an essential one that requires implementation of social-psychological perspectives in a comprehensive, long-term plan of attitude, value, and behavior change.

Understanding What Went Wrong in Abu Ghraib Prison

The situational influences of evil came to the fore in recent trials of American prison guards at Abu Ghraib. In October 2004, I testified via closed circuit television to the military trial judge in Baghdad in defense of one of the guards, Sergeant Ivan Frederick. The abuse of

Iraqi prisoners horrified the sensibility of people around the world, in part because it was the first time in history that such conduct was publicized in graphic photographic images. The guards directly responsible were widely condemned as "morally corrupt" and presented by American leaders as exceptions to the norm of American soldiers. The initial focus of the government "to get to the bottom" of this mess clearly reflected a dispositional orientation; blame was attributed to sadistic personalities and related pathologies.

My expert testimony involved an army reserve sergeant in charge of the night shift where all the mayhem occurred. Everything I could learn about the Abu Ghraib Prison, Tier 1-A, the "soft torture" interrogation center of that prison, revealed virtually all of the social psychological processes operating in the Stanford Prison Experiment. In fact, one of the independent investigations, headed by James Shiesinger, specifically details the parallels between the two prisons.

The Abu Ghraib guards were army reservists forced into this role with no mission-specific training. They lacked supervision and personal accountability. Prison norms encouraged the abuse and humiliation of detainees as a way to soften them up for interrogation. These norms were reinforced by civilian contract interrogators, the military police, the Central Intelligence Agency, and the entire chain of military and political command.

I offered evidence in three areas:

Dispositional: Evidence for any personal pathologies or sadistic tendencies, as well as positive traits and values

Situational: Evidence of the working conditions on that prison's night shift

Systemic: Evidence of the broader conditions that spawned and sustained that situation, particularly the leadership and the objectives of that interrogation center

With regard to disposition, this soldier was totally and unequivocally normal on all measures assessed by an army clinical psychologist

(and independently validated by a civilian expert in assessment). There was no evidence of any psychopathology or sadistic tendencies. Rather, Frederick matches our stereotypes of the All-American man. He is a patriotic son of a West Virginia coal miner. He hunts, fishes, plays softball, attends Baptist church services regularly, and has a strong marriage to an African American woman. As a reserve soldier, he had a blameless record as a guard in a low-security, small-town civilian prison with one hundred inmates. He was proud to serve in Iraq and initially worked with children in a small village where he was starting to learn Arabic.

The situational conditions, the behavioral context, involved working conditions that bordered on the inhumane for both guards and prisoners. The processes operating in that prison were directly comparable to those in the Stanford Prison Experiment: deindividuation, dehumanization, moral disengagement, social modeling, pressures for conformity, passive bystanders, power differentials, use of enforced nakedness, and sexually humiliating tactics. The worst abuses in both settings took place on the night shift. The working conditions for Frederick included twelve-hour night shifts (4:00 P.M. to 4:00 A.M.), seven days a week, for forty days with not a day off, then fourteen days after one day off. Exhaustion and stress were exacerbated by chaotic conditions: unsanitary and filthy surroundings that smelled like a putrid sewer, with limited water for showering. Frequent electrical blackouts created dangerous opportunities for prisoner attacks. This young man with no mission-specific training was put in charge of more than three hundred prisoners initially; that number soon swelled to more than a thousand. He was also in charge of twelve army reserve guards and sixty Iraqi police, who often smuggled contraband to the inmates. He rarely left the prison. When off duty, he slept in a cell in a different part of prison. He missed breakfasts and stopped exercising or socializing. Tier 1-A became his total reference setting.

This of itself would qualify anyone for total job burnout.[27] But guards were also under the stress of frequent insurgency attacks. Five U.S. soldiers and twenty prisoners were killed, and many others were wounded by almost daily shelling during Frederick's service. Finally, there were abusive acts by the prisoners themselves. Seven had rioted in another part of the prison and were sent to Tier 1-A for "safe keeping." Four others had raped a fellow boy prisoner. Five separate military investigations concerning Abu Ghraib acknowledge its horrendous conditions. They also acknowledge "failures of leadership, lack of leadership, indifferent leadership, and conflicting leadership demands." What is clear from these independent investigations is a total absence of accountability and oversight and an encouragement of stress interrogation.

The military judge took none of these conditions into account when he issued his sentence. Frederick received the maximum penalty: imprisonment for eight years, a dishonorable discharge, and loss of twenty-two years of army reserve retirement funds. This verdict represents yet another failure of leadership. It ignores the systemic conditions encouraging abuse and absolves the military and political officials responsible.

Promoting Civic Virtue, Moral Engagement, and Human Goodness

There are no simple solutions for the evils addressed here; if there were, they would already have been enacted by those far wiser than I. But the past half-century of psychological research provides some insight about what might be done at the individual and situational levels.

At the individual level, let us first imagine the reverse of the Milgram experiment. Suppose our objective is to create a setting in which people would comply with ever increasing demands to do good, behave in more altruistic ways, undertake ever more positive, prosocial actions. Instead of a paradigm to facilitate the slow descent into evil, we need a paradigm for the slow ascent into goodness.

Consider an altruism scale, which begins by pressuring subjects to provide thirty minutes or an hour of time for a good deed. The demands would then escalate to longer periods and regular

commitments to worthy causes. One could imagine, for example, a progression on behalf of a sustainable environment. Individuals would be encouraged to engage in activities that required increasing commitment of time and money to "green" causes.

Some of the same social-psychological forces that fostered the abuses described could also be harnessed for positive ends. We could, for example, construct an eleven-step plan for promoting civic virtue that parallels the ten steps toward evil outlined earlier:

1. Openly acknowledge errors in judgments. This reduces the need to justify mistakes and continue immoral action. It undercuts the motivation to reduce dissonance by reaffirming a bad decision.
2. Encourage mindfulness. People need reminders not to live their lives on automatic pilot, but to reflect on the situation and consider its ethical implications before acting.[28]
3. Promote a sense of personal responsibility and accountability for all of one's actions. People need a better understanding of how conditions of diffused responsibility disguise their own individual role in the outcomes of their actions.
4. Discourage minor transgressions. Small acts—cheating, gossiping, lying, teasing, and bullying—provide the first steps toward escalating abuses.
5. Distinguish between just and unjust authority. The fact that individuals occupy a position of authority, as in the Milgram experiment, does not entitle them to obedience in unethical actions.
6. Support critical thinking. People need to be encouraged to demand evidence and moral justifications and to evaluate their credibility.
7. Reward moral behavior. More recognition needs to be available for those who do the right thing under difficult situations, such as whistle-blowers in public and private sector positions.

8. Respect human diversity. Appreciating difference is key to reducing in-group biases and prejudices.
9. Change social conditions that promote anonymity. Making people feel special and accountable can promote socially desirable actions and reinforce individuals' sense of self-worth.
10. Challenge pressures for conformity. Individuals need strategies for resisting group influences and maintaining their own moral compass.
11. Refuse to sacrifice crucial freedoms for elusive promises of security. These sacrifices are often the first step toward fascism, and the price is often prohibitive.[29]