

Selective Activation and Disengagement of Moral Control

Albert Bandura

Stanford University

Moral conduct is motivated and regulated mainly by the ongoing exercise of self-reactive influence. But self-regulatory mechanisms do not operate unless they are activated, and there are different psychological mechanisms by which moral control can be selectively activated or disengaged from inhumane conduct. Self-sanctions can be disengaged by reconstruing detrimental conduct through moral justification, euphemistic labeling, and advantageous contrast with other inhumanities; by obscuring personal agency in detrimental activities through diffusion and displacement of responsibility; by disregarding or misrepresenting the harmful consequences of inhumane conduct; and by blaming and dehumanizing the victims. These mechanisms of moral disengagement operate not only in the perpetration of inhumanities under extraordinary circumstances, but in everyday situations where people routinely perform activities that bring personal benefits at injurious costs to others. Given the many psychological devices for disengagement of moral control, societies cannot rely solely on individuals, however honorable their standards, to provide safeguards against inhumanities. To function humanely, societies must establish effective social safeguards against moral disengagement practices that foster exploitive and destructive conduct.

The exercise of self-sanction plays a central role in the regulation of inhumane conduct. In the course of socialization, moral standards are adopted that serve as guides and deterrents for conduct. Once internalized control has developed, people regulate their actions by the sanctions they apply to themselves. They do things that give them self-satisfaction and a sense of self-worth. They

This article includes material from the book, *Social Foundations of Thought and Action: A Social Cognitive Theory* (Bandura, 1986) and from a chapter in Kurtines and Gewirtz (1990), *Handbook of Moral Behavior and Development: Theory, Research and Applications* (Vol. 1).

Correspondence regarding this article should be addressed to Albert Bandura, Department of Psychology, Jordan Hall, Building 420, Stanford University, Stanford, CA 94305.

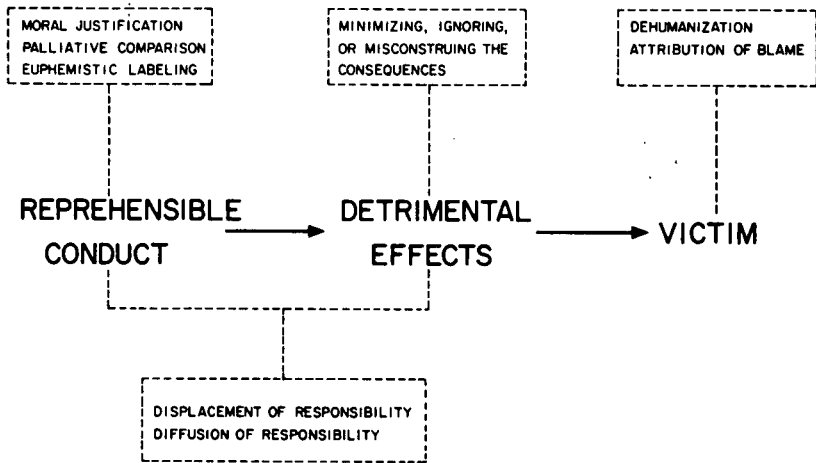


Fig. 1. Mechanisms through which moral self-sanctions are selectively activated or disengaged from reprehensible conduct at different points in the regulatory process. Source: Albert Bandura, *Social Foundations of Thought and Action: A Social Cognitive Theory*, © 1986, p. 376. Reprinted by permission of Prentice-Hall, Englewood Cliffs, NJ.

refrain from behaving in ways that violate their moral standards because such behavior will bring self-condemnation. Self-sanctions thus keep conduct in line with internal standards. But moral standards do not function as fixed internal regulators of conduct. Self-regulatory mechanisms do not operate unless they are activated, and there are many processes by which self-sanctions can be disengaged from inhumane conduct (Bandura, 1986, 1990a). Selective activation or disengagement of self-reactive control permits different types of conduct, given the same moral standards. Figure 1 shows the four major points in the self-regulatory process at which internal moral control can be disengaged from detrimental conduct. Self-sanctions can be disengaged by reconstruing conduct, obscuring causal agency, disregarding or misrepresenting injurious consequences, and blaming and devaluing the victims.

These mechanisms of moral disengagement have been examined most extensively in aggressive conduct. But selective disengagement of moral control is by no means confined to extraordinary inducements to aggression. People often experience conflicts where behavior they personally devalue can serve as the means for securing valued benefits. As long as self-sanctions override the force of external inducements, behavior is kept in line with personal standards. However, in the face of strong external inducements, such conflicts are often resolved by selective disengagement of self-sanctions. This enables otherwise considerate people to perform self-serving activities that have detrimental social effects.

Reconstructing Detrimental Conduct

Moral Justification

One set of disengagement practices operates on the construal of the behavior itself. People do not ordinarily engage in reprehensible conduct until they have justified to themselves the morality of their actions. What is culpable can be made righteous through cognitive reconstrual. In this process, detrimental conduct is made personally and socially acceptable by portraying it as in the service of moral purposes. People then can act on a moral imperative.

Although moral and cognitive restructuring can be easily used to support self-serving and destructive purposes, it can also serve militant action aimed at changing inhumane social conditions. By appealing to morality, social reformers are able to use coercive, and even violent, tactics to force social change. Vigorous disputes arise over the morality of aggressive action directed against institutional practices. Power holders often resist, by forcible means if necessary, pressures to make needed social changes that jeopardize their own self-interests. Resistance to warranted changes invites social activism. Challengers define their militant actions as morally justifiable means to eradicate harmful social practices. Power holders, in turn, condemn such activism as representing irresponsible resort to violent solutions or efforts to coerce changes that lack popular support.

Radical shifts in destructive behavior through moral justification are most strikingly revealed in military conduct. People who have been socialized to deplore killing as morally condemnable can be rapidly transformed into skilled combatants, who may feel little compunction and even a sense of pride in taking human life in combat. The conversion of socialized people into dedicated fighters is achieved not by altering their personality structures, aggressive drives, or moral standards. Rather, it is accomplished by cognitively restructuring the moral value of killing, so that it can be done free from self-censuring restraints (Sanford & Comstock, 1971; Kelman, 1973). Through moral sanction of violent means, people see themselves as fighting ruthless oppressors who have an unquenchable appetite for conquest, or protecting their cherished values and way of life, preserving world peace, saving humanity from subjugation to an evil ideology, or honoring their country's international commitments. The task of making violence morally defensible is facilitated when nonviolent options are judged to have been ineffective, and utilitarian justifications portray the suffering caused by violent counterattacks as greatly outweighed by the human suffering inflicted by the foe.

Over the centuries, much reprehensible and destructive conduct has been perpetrated by ordinary, decent people in the name of religious principles, right-

eous ideologies, and nationalistic imperatives. Individuals espousing high moral principles are inclined to resist arbitrary social demands to behave punitively, but they will aggress against people who violate their personal principles (Keniston, 1970). Throughout history countless people have suffered at the hands of self-righteous crusaders bent on stamping out what they considered evil. Rapoport and Alexander (1982) document the lengthy blood-stained history of holy terror wrought by religious justifications. Acting on moral or ideological imperatives reflects a conscious offense mechanism, not an unconscious defense mechanism.

When viewed from divergent perspectives, violent acts are different things to different people. It is often proclaimed in conflicts of power that one group's terroristic activity is another group's liberation movement fought by heroic freedom fighters. This is why moral appeals against violence usually fall on deaf ears. Adversaries sanctify their own militant actions but condemn those of their antagonists as barbarity masquerading under a mask of outrageous moral reasoning.

Terrorism and counterterrorism. Terrorists invoke moral principles to justify human atrocities (Bandura, 1990b). Moral justification is also brought into play in selecting counterterrorist measures. This poses more troublesome problems for democratic societies than for totalitarian ones. Totalitarian regimes have fewer constraints against using institutional power to control media coverage of terrorist events, to restrict individual rights, to sacrifice individuals for the benefit of the state rather than make concessions to terrorists, and to combat threats with lethal means. Democratic societies face the moral dilemma of how to justify countermeasures that take some toll on innocent lives in the process of curbing terrorists' atrocities, without violating the societies' own fundamental principles and standards of civilized conduct (Carmichael, 1982). Therefore, the use of violent countermeasures is typically justified on utilitarian grounds in terms of the benefits to humanity and the social order that curbing terrorist attacks will bring. The mass media, especially television, provide the best access to the public through their strong drawing power. For this reason, television is increasingly used as the principal vehicle of justification. Struggles to legitimize and gain support for one's causes and to discredit those of one's opponents are now waged more and more through the electronic media (Ball-Rokeach, 1972; Bassiouni, 1981).

Use of nuclear weapons and nuclear threats. The nuclear age has ushered in new magnitudes of risk that create major moral imperatives and paradoxes. Major disputes revolve around the morality of the development of nuclear weaponry and nuclear retaliatory policies (Churchill, 1983; Johnson, 1984; Lackey, 1984). Proponents of the deterrence doctrine justify threat of nuclear retaliation on the grounds that self-defense against grave dangers is morally obligatory.

National security is presumably ensured by maintaining a balance of potential nuclear destructiveness that will be mutually deterring. However, threats of nuclear retaliation have no deterrent effect unless the feuding nations believe their adversary has every intention to use such weapons in the event of a nuclear attack. Since virtually everyone concedes that it would be suicidal to use them, a nuclear deterrence doctrine paradoxically seeks to achieve a deterrent effect with threats of actions that no one in their right mind could conceive of ever taking. Hence, in efforts to add credibility to deterrence policies, nuclear weapons are menacingly deployed and nuclear systems are said to be preprogrammed so that a launch of offensive missiles will trigger a massive nuclear counterstrike semiautomatically.

Opponents of nuclear deterrence policies consider the development of nuclear weaponry and threats to use it, even in retaliation, as morally wrong. They regard a retaliatory strike that would inevitably produce vast human and ecological devastation as a ghastly act of vengeance that is irrational as well as immoral. A counterstrike after a failed deterrence would most likely achieve only massive mutual destruction through a series of nuclear exchanges with surviving missiles. In the aftermath, survivors would find themselves in a largely uninhabitable environment. In short, the moral logic of counterstrike threat is undermined by its self-destructive consequences. Because of the vast scope and magnitude of indiscriminate nuclear devastation, the traditional just-war tenets that sanction self-defense in order to avert grave harm afford little guidance in the use of nuclear weapons. What is immoral to do is immoral to threaten (Kavka, 1988). For opponents of nuclear systems, their indiscriminate destructiveness challenges the moral permissibility of nuclear powers inflicting the catastrophic risks of nuclear deterrence on the people of innocent nations who are granted no say in the matter (Lackey, 1985). What is immoral to do is also immoral to risk.

Euphemistic Labeling

Language shapes people's thought patterns, on which they base many of their actions. Activities can take on a very different appearance depending on what they are called. Euphemistic language thus provides a convenient device for masking reprehensible activities or even conferring a respectable status upon them. Through convoluted verbiage, destructive conduct is made benign and those who engage in it are relieved of a sense of personal agency. Laboratory studies reveal the disinhibitory power of euphemistic language (Diener et al., 1975). Adults behave much more aggressively when the act of assaulting a person is given a sanitized label than when it is called aggression.

In an insightful analysis of the language of nonresponsibility, Gambino (1973) identified the different varieties of euphemisms. One form, palliative expressions, is widely used to make the reprehensible respectable. Through the

power of hygienic words, even killing a human being loses much of its repugnancy. Soldiers "waste" people rather than kill them, intelligence operatives "terminate them with extreme prejudice" (Safire, 1979). When mercenaries speak of "fulfilling a contract," murder is transformed by admirable words into the honorable discharge of duty. Terrorists label themselves as "freedom fighters." Bombing attacks become "clean, surgical strikes," invoking imagery of the restorative handiwork of the operating room, and the civilians they kill are linguistically converted to "collateral damage" (Hilgartner, Bell, & O'Connor, 1982).

Sanitizing euphemisms, of course, also perform heavy duty in less loathsome but unpleasant activities that people are called upon to conduct from time to time. In the language of some government agencies, people are not fired, they are "selected out," as though they were receiving preferential treatment. In teaching business students how to lie in competitive transactions, the instructor speaks euphemistically of "strategic misrepresentation" (Safire, 1979). The television industry produces and markets some of the most brutal forms of human cruelty under the sanitized labels of "action and adventure" programming (Baldwin & Lewis, 1972). The acid rain that is killing our lakes and forests loses much of its virulence in its euphemistic form as "atmospheric deposition of anthropogenically derived acidic substances" (Hechinger, 1985). The nuclear power industry has created its own specialized set of euphemisms for the injurious effects of nuclear mishaps; an explosion becomes an "energetic disassembly," a reactor accident is a "normal aberration," and plutonium contamination is merely "infiltration" (*San Francisco Chronicle*, 1979).

The agentless passive form serves as a linguistic device for creating the appearance that culpable acts are the work of nameless forces, rather than people (Bolinger, 1982). It is as though people are moved mechanically but are not really the agents of their own acts. Gambino (1973) further documented how the specialized jargon of a legitimate enterprise can be misused to lend an aura of respectability to an illegitimate one. In the Watergate vocabulary, criminal conspiracy became a "game plan," and the conspirators were "team players"—a status calling for the qualities and behavior befitting the best sportsmen. The disinhibitory power of language can be boosted further by colorful metaphors that change the nature of culpable activities.

Advantageous Comparison

Whenever events occur or are presented contiguously, the first one colors how the second one is perceived and judged. By exploiting the contrast principle, moral judgments of conduct can be influenced by expedient structuring of the comparison. Thus, self-deplored acts can be made righteous by contrasting them with flagrant inhumanities. The more outrageous the contrasted actions, the more

likely it is that one's own destructive conduct will appear trifling or even benevolent. Promoters of the Vietnamese war and their supporters, for example, minimized the slaying of countless people by portraying it as a way of checking massive communist enslavement. Given this invidious comparison, perpetrators of the warfare remained unperturbed by the fact that the intended beneficiaries were being killed at an alarming rate. Domestic protesters, on the other hand, characterized their own violence against educational and political institutions as trifling, or even laudable, by comparing it with the carnage perpetrated by their country's military forces in foreign lands. Terrorists minimize their killings as the only defensive weapon they have to curb the widespread cruelties inflicted on their people. In the eyes of their supporters, risky attacks directed at the apparatus of oppression are acts of selflessness and martyrdom. Those who are the objects of terrorist attacks, in turn, characterize their own retaliatory violence as trifling, or even laudable, by comparing it with the carnage and terror perpetrated by terrorists. In social conflicts, injurious behavior usually escalates, with each side lauding its own behavior but morally condemning that of their adversaries as heinous.

Historical advantageous comparisons are also invoked as justifications for injurious behavior. For example, advocates of terrorist tactics strive to legitimize them by noting that the democracies of England, France, and the United States were born out of violence against oppressive rule. Apologists for the lawlessness and deceitfulness in the Iran-Contra weapons sales were quick to invoke transgressions by past political administrations as vindications.

Social comparison is similarly used to show that the social labeling of acts may depend more on the ideological allegiances of the labelers than on the acts themselves. Airline hijackings were applauded as heroic deeds when East Europeans and Cubans initiated this practice, but condemned as terrorist acts when the airliners of Western nations and friendly countries were commandeered. The degree of psychopathology ascribed to hijackers varied depending on the direction of the rerouted flights. Moral condemnations of politically motivated terrorism are easily blunted by social comparison because, in international contests for political power, it is hard to find nations that categorically condemn terrorism. Rather, they usually back some terrorists and oppose others.

Cognitive restructuring of behavior through moral justifications and palliative characterizations is the most effective psychological mechanism for disengagement of moral self-sanctions. This is because moral restructuring not only eliminates self-deterrents but engages self-approval in the service of destructive exploits. What was once morally condemnable becomes a source of self-valuation. After destructive means become invested with high moral purpose, functionaries work hard to become proficient at them and take pride in their destructive accomplishments.

Obscuring Personal Agency

Displacement of Responsibility

Self-sanctions are activated most strongly when personal agency for detrimental effects is unambiguous. Another set of disengagement practices operates by obscuring or distorting the relationship between actions and the effects they cause. People will behave in ways they normally repudiate if a legitimate authority accepts responsibility for the consequences of the conduct (Diener et al., 1975; Milgram, 1974). Under conditions of displaced responsibility, people view their actions as springing from the dictates of authorities rather than from their own personal responsibility. Since they feel they are not the actual agent of their actions, they are spared self-prohibiting reactions. Displacement of responsibility not only weakens restraints over one's own detrimental actions but also diminishes social concern over the well-being of those mistreated by others (Tilker, 1970).

Most of the research on attributional analysis of moral judgment is concerned with whether people view their behavior as determined by external circumstances or hold themselves responsible for it (Ross & DiTecco, 1975; Rule & Nesdale, 1976). Perceptions of causal responsibility are reduced if the harmful consequences of actions are viewed as unintended or unforeseeable, or if the actions arose from the dictates of the situation. Within the attributional framework, these factors are usually studied as mitigators of moral judgment rather than as disengagers of moral self-sanctions.

Exemption from self-devaluation for heinous deeds by displacement of responsibility has been most gruesomely revealed in socially sanctioned mass executions. Nazi prison commandants and their staffs divested themselves of personal responsibility for their unprecedented inhumanities (Andrus, 1969). They claimed they were simply carrying out orders. Impersonal obedience to horrific orders was similarly evident in military atrocities, such as the My Lai massacre (Kelman, 1973). In an effort to deter institutionally sanctioned atrocities, the Nuremberg Accords were established, declaring that obedience to inhumane orders, even from the highest authorities, does not relieve subordinates of the responsibility of their actions. However, since victors are disinclined to try themselves as criminals, such decrees have limited deterrence effects without an international judiciary system empowered to impose penalties on victors and losers alike.

In formal studies of disengagement of self-sanctions through displacement of responsibility, authorities explicitly authorize injurious actions and hold themselves fully accountable for the harm caused by the activity. However, in the sanctioning practices of everyday life, responsibility for detrimental conduct is rarely assumed so explicitly, because only obtuse authorities would leave them-

selves accusable of authorizing heinous acts. They are concerned not only with adverse social consequences to themselves should advocated courses of action miscarry, but with their own loss of self-regard as a result of sanctioning human atrocities in ways that leave blood on their hands. Therefore, authorities usually invite and support detrimental conduct in insidious ways that minimize personal responsibility for what is happening. Moreover, the intended purpose of sanctioned destructiveness is usually disguised so that neither issuers nor perpetrators regard their actions as censurable. When reprehensible practices are publicized, they are officially dismissed as only isolated incidents arising through misunderstanding of what had, in fact, been authorized. Efforts are made to limit any blame to subordinates, who are portrayed as misguided or overzealous. Investigators who go searching for "smoking guns" display naïveté about the surreptitious manner in which culpable behavior is sanctioned and executed. Generally one finds mazy devices of nonresponsibility rather than smoking guns.

Kramer (1990) describes the great lengths to which Shit'ite clerics go to provide moral justifications for violent acts that seem to breach Islamic law, such as suicidal bombings and hostage taking. These efforts are designed not only to persuade themselves of the morality of their actions but to preserve their integrity in the eyes of other nations. The Islamic religious code permits neither suicide nor terrorizing innocent people. On one hand, the clerics justify such acts by invoking situational imperatives and utilitarian reasons, namely that tyrannical circumstances drive oppressed people to unconventional means to combat aggressors who wield massive destructive power. On the other hand, they reconstrue terrorist acts as conventional means, in which dying in a suicidal bombing for a moral cause is no different than dying at the hands of an enemy soldier. Hostages typically get relabeled as spies. When the linguistic solution defies credibility, personal moral responsibility is disengaged by construing terroristic acts as dictated by their foe's tyranny. Because of the shaky moral logic and disputable reconstruals involved, clerics sanction terrorism by indirection, they vindicate successful ventures retrospectively, and they disclaim endorsements of terroristic operations beforehand.

Displacement of responsibility also operates in situations in which hostages are taken. Terrorists warn officials of targeted regimes that if they take retaliatory action they will be held accountable for the lives of the hostages. At different steps in negotiations for the hostages' release, terrorists continue to displace responsibility for the safety of hostages on the reactions of the regime they are fighting. If the captivity drags on, terrorists blame the suffering and injuries that they inflict on the hostages on the regime for failing to make what they regard as warranted concessions to right social wrongs.

A number of social factors affect the ease with which responsibility for one's actions can be surrendered to others. High justification and social consensus about the morality of an enterprise aid in the relinquishment of personal

control. The legitimacy of the authorizers is another important determinant. The greater the legitimation and closeness of the authority issuing injurious commands, the higher is the level of obedience. The higher the authorities, the more legitimacy, respect, and coercive power they command, and the more willing are people to defer to them. Modeled disobedience, which challenges the legitimacy of the activities, if not the authorizers themselves, reduces the willingness of observers to carry out the actions called for by the orders of a superior (Meeus & Raaijmakers, 1986; Milgram, 1974; Powers & Geen, 1972). It is difficult to continue to disown personal agency in the face of evident harm following directly from one's actions. People are therefore less willing to obey authoritarian orders for injurious behavior when they see firsthand how they are hurting others (Milgram, 1974; Tilker, 1970).

Obedient functionaries do not cast off all responsibility for their behavior as though they were mindless extensions of others. If this were the case, they would act like automatons, only when told to. In fact, they are much more conscientious and self-directed in the performance of their duties. It requires a strong sense of responsibility to be a good functionary. In situations involving obedience to authority, people carry out orders partly to honor the obligations they have undertaken (Mantell & Panzarella, 1976). One must, therefore, distinguish between two levels of responsibility—duty to one's superiors, and accountability for the effects of one's actions. The self-system operates most efficiently in the service of authority when followers assume personal responsibility for being dutiful executors while relinquishing personal responsibility for the harm caused by their behavior. Followers who disowned responsibility without being bound by a sense of duty would be quite unreliable.

Diffusion of Responsibility

The deterrent power of self-sanctions is weakened when the link between conduct and its consequences is obscured by diffusing responsibility for culpable behavior. This is achieved in several ways. Responsibility can be diffused by division of labor. Most enterprises require the services of many people, each performing fragmentary jobs that seem harmless in themselves. The fractional contribution is easily isolated from the eventual function, especially when participants exercise little personal judgment in carrying out a subfunction that is related by remote, complex links to the end result. After activities become routinized into programmed subfunctions, attention shifts from the import of what one is doing to the details of one's fractional job—e.g., doing a good job assembling bombers on a production line (Kelman, 1973).

Group decision making is another common bureaucratic practice that enables otherwise considerate people to behave inhumanely, because no single individual feels responsible for policies arrived at collectively. Where everyone

is responsible, no one is really responsible. Social organizations go to great lengths to devise sophisticated mechanisms for obscuring responsibility for decisions that will affect others adversely. Collective action is still another diffusion expedient for weakening self-restraints. Any harm done by a group can always be ascribed, in large part, to the behavior of other members. People, therefore, act more harshly when responsibility is obfuscated by a collective instrumentality than when they hold themselves personally accountable for what they do (Bandura, Underwood, & Fromson, 1975; Diener, 1977; Zimbardo, 1969).

People often behave in harmful ways, not because responsibility is diffused by formal organizational arrangements, but because they all routinely engage in activities that contribute to negative effects. Their automobiles pollute the air they breathe, and they degrade their environment to produce the vast amounts of energy and products they consume. As a result of collective action, good environmentalists can also be polluters but blame others for degrading the environment. The more detrimental the collectively produced effects, the less people feel personally responsible for them (Shippee & Christian, 1978).

Disregard or Distortion of Harmful Consequences

Additional ways of weakening self-detering reactions operate through disregard or misrepresentation of the consequences of action. When people choose to pursue activities harmful to others for personal gain, or because of social inducements, they avoid facing the harm they cause or they minimize it. They readily recall prior information given them about the potential benefits of the behavior but are less able to remember its harmful effects (Brock & Buss, 1962, 1964). People are especially prone to minimize injurious effects when they act alone, and thus cannot easily escape responsibility (Mynatt & Herman, 1975). In addition to selective inattention and cognitive distortion of effects, the misrepresentation may involve active efforts to discredit evidence of the harm they cause. As long as the detrimental results of one's conduct are ignored, minimized, distorted, or disbelieved, there is little reason for self-censure to be activated.

It is relatively easy to hurt others when their suffering is not visible, and when causal actions are physically and temporally remote from their effects. Our death technologies have become highly lethal and depersonalized. Mechanized weapon systems and explosive devices that can kill many people but be set off by someone at a distance illustrate such depersonalized action. Even high personal responsibility is a weak restrainer when aggressors do not know the harm they inflict on their victims (Tilker, 1970). In contrast, when people can see and hear the suffering they cause, vicariously aroused distress and self-censure serve as self-restraining influences. For example, in his studies of commanded aggression, Milgram (1974) found diminishing obedience as the victim's pain became more evident and personalized.

Most organizations involve hierarchical chains of command in which superiors formulate plans and intermediaries transmit them to executors, who then carry them out. The father removed individuals are from the end results, the weaker is the restraining power of the foreseeable destructive effects. Kilham and Mann (1974) suggest that the disengagement of personal control is easiest for the intermediaries in a hierarchical system—they neither bear responsibility for major decisions nor are they a direct party to their execution. In performing the transmitter role they model dutiful behavior and further legitimize their superiors and their social policies and practices. Consistent with these speculations, intermediaries are much more obedient to destructive commands than are those who have to carry them out and face the results (Kilham & Mann, 1974).

Blaming and Dehumanizing Victims

Dehumanization

The final set of disengagement practices operates on the recipients of detrimental acts. The strength of self-evaluative reactions to injurious conduct partly depends on how the perpetrators view the people toward whom the behavior is directed. To perceive another as human activates empathetic or vicarious emotional reactions through perceived similarity (Bandura, 1990c). The joys and suffering of similar persons are more vicariously arousing than are those of strangers or of individuals who have been divested of human qualities. Personalizing the injurious effects experienced by others also makes their suffering much more salient. It is difficult to mistreat humanized persons without risking personal distress and self-censure. However, vicarious emotional activation is cognitively mediated rather than automatically elicited by the experiences of others. Ascriptions of insensateness to victims weakens vicarious self-arousal of distress to their suffering. Empathy is not aroused if recipients of maltreatment are seen as "unfeeling." Subhumans are not only regarded as lacking sensitivities, but as being influenceable only by severe methods.

Self-sanctions against cruel conduct can be disengaged or blunted by divesting people of human qualities. Once dehumanized, they are not longer viewed as persons with feelings, hopes, and concerns, but as subhuman objects. They are portrayed as mindless "savages," "gooks," "satanic fiends," or other despicable wretches. If dispossessing antagonists of humanness does not sufficiently blunt self-reproof, it can be eliminated by attributing bestial qualities to them. They become "degenerates," "pigs," and other bestial creatures. It is easier to brutalize victims when they are referred to as "worms" (Gibson & Haritos-Fatouros, 1986). The process of dehumanization is an essential ingredient in the perpetration of inhumanities. A Nazi camp commandant chillingly explained that the extreme lengths to which they went to degrade victims whom they were

going to kill anyway were not a matter of purposeless cruelty (Levi, 1987). Rather, the victims had to be degraded to the level of subhuman objects so that those who operated the gas chambers would be less burdened by distress. Similarly, over the years, slaves, women, manual laborers, and religious and racial minorities have been treated as chattel or as subhuman objects (Ball-Rokeach, 1972).

When persons are given punitive power, they treat dehumanized individuals much more punitively than ones who have been invested with human qualities. Dehumanization fosters different self-exonerative patterns of thought (Bandura, Underwood, & Fromson, 1975). When dealing with individuals who have been deprived of their humanness, people seldom condemn punitive conduct, and they create justifications for it. However, people disavow punitive actions and rarely create justifications for their use toward individuals depicted in humanized terms.

When several disengagement factors are combined, they potentiate each other rather than simply producing additive effects. Thus, combining diffused responsibility with dehumanization greatly escalates the level of punitiveness, whereas personalization of responsibility, along with humanization, have a powerful self-detering effect.

Many conditions of contemporary life are conducive to impersonalization and dehumanization (Bernard, Ottenberg, & Redl, 1965). Bureaucratization, automation, urbanization, and high geographical mobility lead people to relate to each other in anonymous, impersonal ways. In addition, social practices that divide people into ingroup and outgroup members produce human estrangement that fosters dehumanization. Strangers can be more easily cast as insensate than can personal acquaintances.

Under certain conditions, the exercise of institutional power changes the users in ways that are conducive to dehumanization. This happens most often when persons in positions of authority have coercive power over others and adequate safeguards for constraining the behavior of power holders are lacking. Power holders come to devalue those over whom they wield control (Kipnis, 1974). In a simulated prison experiment (Haney, Banks, & Zimbardo, 1973), even college students, who had been randomly chosen to serve as either inmates or guards, when given unilateral power as guards, began to treat their charges in degrading, tyrannical ways. Thus, role assignment that authorized use of coercive power overrode personal characteristics in promoting punitive conduct. Systematic tests of relative influences similarly show that social influences conducive to punitiveness exert considerably greater sway over aggressive conduct than do people's personal characteristics (Larsen, Coleman, Forges, & Johnson, 1971).

The overall findings from research on the different mechanisms of moral disengagement corroborate the historical chronicle of human atrocities: It re-

quires conducive social conditions, rather than monstrous people, to produce heinous deeds. Given appropriate social conditions, decent, ordinary people can be led to do extraordinarily cruel things.

The power of humanization. Psychological research tends to focus extensively on how easy it is to bring out the worst in people through dehumanization and other self-exonerative means. The sensational negative findings receive the greatest attention. Thus, for example, the aspect of Milgram's research on obedient aggression that is widely cited is the evidence that good people can be talked into performing cruel deeds. However, to get people to carry out punitive acts, the overseer had to be physically present, and to repeatedly order them to act cruelly as they voiced their concerns and increasing objections. Orders to escalate punitiveness to more intense levels were largely ignored or subverted when remotely issued by verbal command. As Helm and Morelli (1979) note, this is hardly an example of blind obedience triggered by an authoritative mandate. Moreover, though it is less often noted, the evidence is equally striking that most people steadfastly refuse to behave punitively, even in response to strong authoritarian commands, if the situation is personalized by having them see the victim or requiring them to inflict pain directly rather than remotely.

The frequent emphasis on obedient aggression is understandable considering the prevalence and harmfulness of people's inhumanities to one another. However, there is considerable theoretical and social significance in the power of humanization to counteract cruel conduct. Studies examining this process reveal that it is difficult for individuals to behave cruelly toward others when they are humanized or even personalized a bit (Bandura et al., 1975). Even under conditions where punitive sanctions are the only means available and they are highly functional in producing desired results, those exercising that power usually cannot carry through punitive behavior toward humanized individuals. The affirmation of common humanity can bring out the best in others. In contrast, even in situations where punitive sanctions are dysfunctional because they usually fail to produce results, punitiveness is precipitously escalated toward dehumanized individuals. The failure of degraded individuals to change in response to punitive treatment is taken as further evidence of their culpability, justifying intensified punitiveness toward them.

The moderating influence of humanization is strikingly revealed in situations involving great threat of violence. Most abductors find it difficult to harm their hostages after they have gotten to know them personally. With growing acquaintance, it becomes increasingly difficult to take a human life cold-bloodedly. Humanization, of course, is a two-way process. Captives may also develop some sympathy for their captors as they get to know them. This phenomenon was graphically illustrated in a Stockholm incident in which people who were held hostage for six days by bank robbers began to sympathize with their criminal

captors and sided with them against the police (Lang, 1974). This hostage incident included several features that are especially conducive to development of human affinity (Bandura, 1990b). Most people support the death penalty in the abstract, but the more they know about particular convicted individuals, the less they favor executing them (Ellsworth, 1978). As Ellsworth explains it, in the absence of personal information, people conjure up an image of the most heinous criminal, an image that disposes them to favor punishment by death.

Attribution of Blame

Imputing blame to one's antagonists or to environmental circumstances is still another expedient that can serve self-exonerative purposes. In this process, people view themselves as faultless victims and see their detrimental conduct as compelled by forcible provocation. Detrimental interactions usually involve a series of reciprocally escalating actions, in which the antagonists are rarely faultless. One can always select from the chain of events an instance of the adversary's defensive behavior and consider it as the original instigation. One's own injurious conduct thus becomes a justifiable defensive reaction to belligerent provocations. Those who are victimized are not entirely faultless because, by their behavior, they usually contribute at least partly to their own plight. Victims can therefore be blamed for bringing suffering on themselves. Self-exoneration is similarly achievable by viewing one's destructive conduct as forced by circumstances rather than as a personal decision. By blaming others or circumstances, not only are one's own actions made excusable, but one can even feel self-righteous in the process.

Observers of victimization can be disinhibited in much the same way as perpetrators are by the tendency to infer culpability from misfortune. Seeing victims suffer maltreatment for which they are held partly responsible leads observers to derogate them (Lerner & Miller, 1978). The devaluation and indignation aroused by ascribed culpability, in turn, provide moral justification for even greater maltreatment. That attribution of blame can give rise to devaluation and moral justification illustrates how the various disengagement mechanisms are often interrelated and work together in weakening internal control.

Imputing blame operates as a prominent disengagement mechanism in sexually assaultive behavior toward women. Rapists and males who acknowledge a proclivity to rape subscribe to myths about rape embodying the various mechanisms by which moral self-censure can be disengaged (Feild, 1978; Malamuth, 1981). These beliefs hold rape victims responsible for their own victimization because they have supposedly invited rape by sexually provocative appearance and behavior, or by resisting sexual assault only weakly. Men blame rape victims more than women do. Trivialization and distortion of consequences to rape victims is another disengagement mechanism that comes into play. Men who are

inclined to assault sexually believe women secretly enjoy being raped. Anticipatory self-censure is eliminated when the traumatic effects of sexual assault are twisted into pleasurable ones for the victim. Such self-disinhibiting patterns of thinking predict proclivity to rape, whereas sexual attitudes, frustration, and quality of sex life do not (Briere & Malamuth, 1983).

Cross-cultural studies reveal that aggressive sexuality is an expression of the cultural ideology of male dominance (Sanday, 1981). Rape is prevalent in societies where violence is a way of life, male supremacy reigns, aggressive sexuality is valued as a sign of manliness, and women are treated as property. Rape is rare in societies that repudiate interpersonal aggression, endorse sexual equality, and treat women respectfully. Cultural ideologies that attach prestige to male dominance and aggressive sexuality weaken self-censure for sexual abuse of women. Cultural practices that belittle the role of women, and a flourishing pornography industry that dehumanizes them, further contribute to the self-disinhibition of aggression toward women (Malamuth & Donnerstein, 1984; Zillman & Bryant, 1984).

Justified abuse can have more devastating human consequences than acknowledged cruelty. Maltreatment that is not clothed in righteousness makes the perpetrator rather than the victim blameworthy. But when blame is convincingly ascribed to victims, they may eventually come to believe the degrading characterizations of themselves (Hallie, 1971). Moreover, ascriptions of blame are usually accompanied by discriminatory social practices that create the very failings that serve as excuses for maltreatment. Vindicated inhumanity is thus more likely to instill self-contempt in victims than inhumanity that does not attempt to justify itself.

Moral Disengagement Is Gradual

The disengagement devices discussed above will not instantaneously transform a considerate person into an unprincipled, callous one. Rather, the change is usually achieved through gradual weakening of self-sanctions, during which people may not fully recognize the changes they are undergoing. Initially, individuals are prompted to perform questionable acts that they can tolerate with little self-censure. After their discomfort and self-reproof have been diminished through repeated performances, the level of reprehensibility progressively increases until, eventually, acts originally regarded as abhorrent can be performed without much distress. Escalative self-disengagement is accelerated if inhumane behavior is construed as serving moral purposes and the people being subjected to maltreatment are divested of human qualities (Bandura et al., 1975; Gibson & Haritos-Fatouros, 1986).

Analyses of moral disengagement mechanisms usually draw heavily on examples from military and political violence. This tends to convey the impres-

sion that selective disengagement of self-sanctions occurs only under extraordinary circumstances. The truth is quite the contrary. Such mechanisms operate in everyday situations in which decent people routinely perform activities having injurious human effects, to further their own interests or for profit. Self-exonerations are needed to neutralize self-sanctions and to preserve self-esteem. For example, institutionalized racial and sexual discrimination, practices that take a heavy toll on their victims, require social justification, attributions of blame, dehumanization, impersonalized agencies to carry them out, and inattention to the injurious effects they cause. As another example, different industries, each with its public-spirited vindications, may cause harmful effects on a large scale, either by the nature of their products or the environmental contaminants they produce.

Concluding Remarks

The massive threats to human welfare stem mainly from deliberate acts of principle rather than from unrestrained acts of impulse. It is the principled resort to destructiveness that is of greatest social concern, but ironically it is the most ignored in psychological analyses of inhumanities. Given the existence of so many psychological devices for disengagement of moral control, societies cannot rely entirely on individuals, however righteous their standards, to provide safeguards against destructive ventures. Civilized conduct requires, in addition to humane personal codes, social systems that uphold compassionate behavior and renounce cruelty. Monolithic political systems that exercise concentrated control over the major vehicles of social influence can wield greater justificatory power than pluralistic systems that represent diverse perspectives, interests, and concerns. Political diversity and toleration of public expression of skepticism create conditions that allow the emergence of challenges to suspect moral appeals. To function more humanely, societies must establish effective social safeguards against the misuse of institutional justificatory power for exploitive and destructive purposes.

References

- Andrus, B. C. (1969). *The infamous of Nuremberg*. London: Fravin.
- Baldwin, T. F., & Lewis, C. (1972). Violence in television: The industry looks at itself. In G. A. Comstock & E. A. Rubinstein (Eds.), *Television and social behavior: Vol. 1. Media content and control* (pp. 290-373). Washington, DC: U.S. Government Printing Office.
- Ball-Rokeach, S. J. (1972). The legitimation of violence. In J. F. Short, Jr. & M. E. Wolfgang (Eds.), *Collective violence* (pp. 100-111). Chicago: Aldine-Atherton.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1990a). Social cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development: Theory, research and applications* (Vol. 1). Hillsdale, NJ: Erlbaum.

- Bandura, A. (1990b). Mechanisms of moral disengagement in terrorism. In W. Reich (Ed.), *Origins of terrorism: Psychologies, ideologies, states of mind* (pp. 161–191). New York: Cambridge University Press.
- Bandura, A. (1990c). Social cognitive theory of social referencing. In S. Feinman (Ed.), *Social referencing and social construction of reality*. New York: Plenum.
- Bandura, A., Underwood, B., & Fromson, M. E. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality*, *9*, 253–269.
- Bassiouni, M. C. (1981). Terrorism, law enforcement, and the mass media: Perspectives, problems, proposals. *Journal of Criminal Law & Criminology*, *72*, 1–51.
- Bernard, V., Ottenberg, P., & Redl, F. (1965). Dehumanization: A composite psychological defense in relation to modern war. In M. Schwebel (Ed.), *Behavioral science and human survival* (pp. 64–82). Palo Alto, CA: Science & Behavior Books.
- Bolinger, D. (1982). *Language: The loaded weapon*. London: Longman.
- Briere, J., & Malamuth, N. M. (1983). Self-reported likelihood of sexually aggressive behavior: Attitudinal versus sexual explanations. *Journal of Research in Personality*, *17*, 315–323.
- Brock, T. C., & Buss, A. H. (1962). Dissonance, aggression, and evaluation of pain. *Journal of Abnormal and Social Psychology*, *65*, 197–202.
- Brock, T. C., & Buss, A. H. (1964). Effects of justification for aggression and communication with the victim on postaggression dissonance. *Journal of Abnormal and Social Psychology*, *68*, 403–412.
- Carmichael, D. J. C. (1982). Of beasts, gods, and civilized men: The justification of terrorism and of counterterrorist measures. *Terrorism*, *6*, 1–26.
- Churchill, R. P. (1983). Nuclear arms as a philosophical and moral issue. *Annals of the American Academy of Political and Social Science*, *469*, 46–57.
- Diener, E. (1977). Deindividuation: Causes and consequences. *Social Behavior and Personality*, *5*, 143–156.
- Diener, E., Dineen, J., Endresen, K., Beaman, A. L., & Fraser, S. C. (1975). Effects of altered responsibility, cognitive set, and modeling on physical aggression and deindividuation. *Journal of Personality and Social Psychology*, *31*, 328–337.
- Ellsworth, P. (1978). *Attitudes toward capital punishment: From application to theory*. Unpublished manuscript, Yale University.
- Feild, H. S. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology*, *36*, 156–179.
- Gambino, R. (1973, November–December). Watergate lingo: A language of non-responsibility. *Freedom at Issue*, pp. 7–9, 15–17.
- Gibson, J. T., & Haritos-Fatouros, M. (1986, November). The education of a torturer. *Psychology Today*, pp. 50–58.
- Hallie, P. P. (1971). Justification and rebellion. In N. Sanford & C. Comstock (Eds.), *Sanctions for evil* (pp. 247–263). San Francisco: Jossey-Bass.
- Haney, C., Banks, C., & Zimbardo, P. (1973). Interpersonal dynamics in a simulated prison. *International Journal of Criminology and Penology*, *1*, 69–97.
- Hechinger, F. M. (1985, February 24). Down with doublespeak. *San Francisco Chronicle* (This World section), p. 19.
- Helm, C., & Morelli, M. (1979). Stanley Milgram and the obedience experiment: Authority, legitimacy, and human action. *Political Theory*, *7*, 321–346.
- Hilgartner, S., Bell, R. C., & O'Connor, R. (1982). *Nukespeak: Nuclear language, visions, and mindset*. San Francisco: Sierra Club Books.
- Johnson, J. T. (1984). *Can modern war be just?* New Haven, CT: Yale University Press.
- Kavka, G. S. (1988). *Moral paradoxes of nuclear deterrence*. New York: Cambridge University Press.
- Kelman, H. C. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues*, *29*(4), 25–61.
- Keniston, K. (1970). Student activism, moral development, and morality. *American Journal of Orthopsychiatry*, *40*, 577–592.

- Kilham, W., & Mann, L. (1974). Level of destructive obedience as a function of transmitter and executant roles in the Milgram obedience paradigm. *Journal of Personality and Social Psychology*, 29, 696-702.
- Kipnis, D. (1974). The powerholders. In J. T. Tedeschi (Ed.), *Perspectives on social power* (pp. 82-122). Chicago: Aldine.
- Kramer, M. (1990). Hezbollah: The moral logic of extraordinary means. In W. Reich (Ed.), *Origins of terrorism: Psychologies, ideologies, states of mind*. New York: Cambridge University Press.
- Lackey, D. P. (1984). *Moral principles and nuclear weapons*. Totowa, NJ: Rowman & Allanheld.
- Lackey, D. P. (1985). Immoral risks: A deontological critique of nuclear deterrence. *Social Philosophy & Policy*, 3, 154-175.
- Lang, D. A. (1974). A reporter at large: The bank drama (Swedish hostages). *New Yorker*, 50(40), 56-126.
- Larsen, K. S., Coleman, D., Forges, J., & Johnson, R. (1971). Is the subject's personality or the experimental situation a better predictor of a subject's willingness to administer shock to a victim? *Journal of Personality and Social Psychology*, 22, 287-295.
- Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, 85, 1030-1051.
- Levi, P. (1987). *The drowned and the saved*. New York: Summit.
- Malamuth, N. M. (1981). Rape proclivity among males. *Journal of Social Issues*, 37(1), 138-157.
- Malamuth, N. M., & Donnerstein, E. (1984). *Pornography and sexual aggression*. New York: Academic Press.
- Mantell, D. M., & Panzarella, R. (1976). Obedience and responsibility. *British Journal of Social and Clinical Psychology*, 15, 239-246.
- Meeus, W. H. J., & Raaijmakers, Q. A. W. (1986). Administrative obedience: Carrying out orders to use psychological-administrative violence. *European Journal of Social Psychology*, 16, 311-324.
- Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.
- Mynatt, C., & Herman, S. J. (1975). Responsibility attribution in groups and individuals: A direct test of the diffusion of responsibility hypothesis. *Journal of Personality and Social Psychology*, 32, 1111-1118.
- Powers, P. C., & Geen, R. G. (1972). Effects of the behavior and the perceived arousal of a model on instrumental aggression. *Journal of Personality and Social Psychology*, 23, 175-183.
- Rapoport, D. C., & Alexander, Y. (Eds.). (1982). *The morality of terrorism: Religious and secular justification*. Elmsford, NY: Pergamon.
- Ross, M., & DiTecco, D. (1975). An attributional analysis of moral judgments. *Journal of Social Issues*, 31(3), 91-109.
- Rule, B. G., & Nesdale, A. R. (1976). Moral judgments of aggressive behavior. In R. G. Geen & E. C. O'Neal (Eds.), *Perspectives on aggression* (pp. 37-60). New York: Academic Press.
- Safire, W. (1979, May 13). The fine art of euphemism. *San Francisco Chronicle*, p. 13.
- Sanday, P. R. (1981). The socio-cultural context of rape: A cross-cultural study. *Journal of Social Issues*, 37(1), 5-27.
- Sanford, N., & Comstock, C. (1971). *Sanctions for evil*. San Francisco: Jossey-Bass.
- San Francisco Chronicle*. (1979, November 22). Award-winning nuclear jargon. p. 24.
- Shippee, G. E., & Christian, M. (1978). *Ego-defensive attributions for an offensive problem: Attributions of responsibility and air pollution severity*. Paper presented at meeting of the Western Psychological Association, San Francisco.
- Tilker, H. A. (1970). Socially responsible behavior as a function of observer responsibility and victim feedback. *Journal of Personality and Social Psychology*, 14, 95-100.
- Zillman, D., & Bryant, J. (1984). Effects of massive exposure to pornography. In N. M. Malamuth & E. Donnerstein (Eds.), *Pornography and sexual aggression* (pp. 115-138). New York: Academic Press.
- Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W. J. Arnold & D. Levine (Eds.) *Nebraska symposium on motivation* (pp. 237-309). Lincoln: University of Nebraska Press.

ALBERT BANDURA is David Starr Jordan Professor of Social Sciences in Psychology at Stanford University. He is a proponent of social cognitive theory, which accords a central role to cognitive, vicarious, self-regulatory, and self-reflective processes in sociocognitive functioning. His recent book, *Social Foundations of Thought and Action: A Social Cognitive Theory*, provides the conceptual framework and analyzes the large body of knowledge bearing on this theory. Bandura is past president of the American Psychological Association and a recipient of its Distinguished Scientific Contributions Award.

Copyright of Journal of Social Issues is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.