

THE SOCIAL  
PSYCHOLOGY OF  
GOOD  
AND  
EVIL

---

*Edited by* Arthur G. Miller

---



THE GUILFORD PRESS  
New York London

© 2004 The Guilford Press  
A Division of Guilford Publications, Inc.  
72 Spring Street, New York, NY 10012

All rights reserved

No part of this book may be reproduced, translated, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher.

Printed in the United States of America

This book is printed on acid-free paper.

Last digit is print number: 9 8 7 6 5 4 3 2 1

**Library of Congress Cataloging-in-Publication Data**

The social psychology of good and evil / edited by Arthur G. Miller.

p. cm.

Includes bibliographical references and index.

ISBN 1-57230-989-X (hardcover: alk. paper)

1. Aggressiveness. 2. Violence. 3. Helping behavior. 4. Good and evil.
  5. Social psychology—Research. I. Miller, Arthur G., 1940–
- HM11116.S63 2004  
303.3'72—dc22

2003020008

## ABOUT THE EDITOR

**Arthur G. Miller, PhD**, is Professor of Psychology at Miami University in Oxford, Ohio. He received his doctorate in social psychology from Indiana University in 1967 and spent 1979–1980 at Princeton University on a National Institute of Mental Health fellowship, studying with Ned Jones. Dr. Miller's professional affiliations include the American Psychological Association, the American Psychological Society, the Society of Experimental Social Psychology, and the Society for Personality and Social Psychology. He has written *The Obedience Experiments: A Case Study of Controversy in Social Science* (1986, Praeger) and edited *In the Eye of the Beholder: Contemporary Issues in Stereotyping* (1982, Praeger). Dr. Miller's primary teaching and research interests include stereotyping and stigma, biases in attribution and social judgment, and judgmental reactions to diverse explanations of evil and violence.

---

**CHAPTER 5**


---

## THE EVOLUTION OF EVIL

JOSHUA D. DUNTLEY  
DAVID M. BUSS

On the evening of July 24, 2002, in the thriving city of Houston, Texas, Clara Harris got into her Mercedes Benz and killed her husband, David Harris, in the parking lot of a hotel. Using her car as the device of murder, she ran into him once. Her anger still not allayed, she circled the lot and ran over him again. Witnesses differ in precisely how many times she backed up and crushed her husband with the 4,000-pound vehicle. One said five times, another four, and a third witness indicated only twice. Video from the hotel security cameras revealed that the correct number was three. Some think that Clara Harris is evil and deserves to rot in jail for the remainder of her life. But some view the homicide as justifiable, or at least understandable.

The circumstance that elicited the homicide was David Harris's affair with Gail Bridges, his former office coworker. Clara Harris discovered the affair through a private detective, whom she had hired when her suspicions were initially aroused. The morning of his death, David Harris swore to Clara that he would end the affair. Later that night, Clara, along with her stepdaughter Lindsey, began to search for David Harris. When they finally tracked him down at a hotel, according to Lindsey, "She said she could kill him and get away with it for what she's been through." Indeed, Clara had gone to great efforts to win her husband back after she discovered his affair. She made herself "real pretty so Dad would want her and not Gail," Lindsey said. During the week before the murder, Clara Harris spent time at a tanning salon, a beauty shop, and a gym. She also consulted a plastic surgeon, inquiring about breast implants.

It might also have aggravated Clara that the hotel was precisely the one where Clara and David had gotten married a decade earlier—on Valentine's Day. When she saw her husband emerge from the hotel elevator with his mistress, the two hand-in-hand, Clara Harris went "ballistic." She ripped the blouse off her rival's body and wrestled her to the ground. Although she clearly intended to do more damage, her husband separated the two women, and Clara was firmly escorted out of the hotel by the clerk on duty. As she left the lobby, David shouted, "It's over! It's over! It's over!"

It was then that Clara Harris became strangely calm, according to her stepdaughter, Lindsey, who accompanied her out of the hotel. Clara silently got into her Mercedes, and her tears stopped flowing. Clara was cool and composed as she suddenly stomped on the accelerator and rammed the car into her husband. She then ran over him again and again. Her stepdaughter tried to exit the vehicle, but had to wait until Clara stopped the car and the damage had finally come to an end. "You killed my Dad," Lindsey said, when the car finally stopped. In light of the circumstances, many in Texas do not judge Clara's horrific deed as evil. Some think that David Harris got exactly what he deserved.

Can good and evil be evaluated from an evolutionary perspective? In this chapter, we consider several related issues:

Have humans evolved adaptations to commit deeds that most would consider "evil"?

Have humans evolved defenses against the perpetration of evil on them?

Do the apparently universal cognitive categories of "good" and "evil" have special functional uses, aiding humans in solving critical, adaptive problems?

### EVIL AS THE INFLECTION OF FITNESS COSTS

Evil has no direct analogue in the formal structure of evolutionary theory. Evolution by natural selection operates by the simple process of differential gene reproduction as a consequence of differences in heritable design. Heritable variants that lead to greater reproduction, compared with competing variants present in the population at the time, become represented in greater numbers in the next generation. Iterated over multiple generations, the process of selection leads to the evolution of adaptations that exist solely because they contributed, either directly or indirectly, to the reproductive success of their bearers. Thus, the process of natural selection is value-free. Whatever qualities lead to increased replicative success are those that evolve.

The evolutionary process of selection produces many products, and humans have little hesitation in labeling some of those products "good" and others "evil." At a first approximation, those we label as "evil" are behaviors or behavioral dispositions that result in a massive imposition of fitness costs on another individual or group. Indeed, as we argue later in the chapter, humans have evolved a specialized psychological categorization system for making these judgments.

The imposition of fitness costs on another individual can vary in magnitude from trivial to catastrophic. At the low end, someone bumping into you in the hall or stepping on your toe might be considered annoying, but probably not evil, unless these acts were repeated to the point of torture. At the high end are events such as robbing, maiming, rape, torture, and murder, with combinations of these usually viewed as embodying evil more fully than any considered alone. Intentional premeditated murder occupies the extreme end of the continuum, but within that broad class, some murders are considered to be more evil than others—murder with malice, murder without provocation, murder of young, defenseless children, murder of adolescent girls, serial murder, mass murder, and genocide. Some homicides, on the other hand, are considered excusable, justifiable, or even altruistic—for example, killing in self-defense, killing to protect a family member from harm, or killing to prevent a helpless stranger from being raped. Of course, as Baumeister (1997) points out, the judgment of the perpetrators and victims will surely differ in how evil these deeds are evaluated to be (a point that we will take up later). The key point here is that the acts we consider evil invariably involve the imposition of massive costs on victims, even though not all massive costs are considered to be evil.

By what metric do we judge acts to be costly? One contention of this chapter is that the deeds we view as evil occupy the extreme end of a continuum of reproductively relevant costs—that is, those that impose a massive fitness cost on the victim will be viewed to be the most evil. Humans, of course, do not think in these terms. We do not think to ourselves: "Gee, the damage done to Sally inflicts a large cost on her fitness, which impairs her relative gene replication . . . hence, it's evil." Rather, we propose that humans have evolved evaluative psychological mechanisms that function to gauge the magnitude of fitness costs inflicted on themselves, their allies, their children, and their extended families—roughly, the degree of evil. We have also evolved evaluative mechanisms to assess the magnitude of fitness benefits that others bestow on us and our vehicular allies—roughly, the degree of good. According to our evolutionary theory of good and evil, the degree of evil and of good judged by a person is partially a function of the person's degree of genetic relatedness to the person upon whom costs are inflicted or benefits are bestowed. Costs inflicted or benefits bestowed upon closer relatives would be more evil or good, respectively, than the

same amounts of costs or benefits accruing to more distant genetic relatives. Degree of evil or good is also a function of "strategic confluence," that is, the degree to which other individuals are allied with us in achieving some goal (Buss, 1996). Thus, extreme fitness costs inflicted on a close friend would be judged as more evil than comparable costs inflicted on a stranger or an enemy. Indeed, massive fitness costs inflicted on an enemy are often judged to be good. In sum, the degree of strategic confluence, including individuals who are either genetic kin or non-kin allies, is predicted to mediate the degree to which an intentionally inflicted fitness cost is judged to be evil.

In order for these evaluative mechanisms to have evolved, however, there must have been evolutionarily recurrent deeds that humans performed that correspond to these psychological categorizations. Thus, before exploring the evolution of human judgments of good and evil and the functions of these psychological mechanisms, we must first explore why people inflict extreme levels of egregious harm on other people.

### HUMAN PSYCHOLOGY AS THE END PRODUCT OF A COMPETITIVE EVOLUTIONARY PROCESS

From an evolutionary perspective, modern humans are the end products of a long line of successful reproducers. Indeed, all humans are evolutionary success stories. Each one of us owes his or her existence to an unbroken chain of ancestors, each of whom did what was necessary to survive and reproduce. If any one of our ancestors or their ancestors had failed at these tasks—for example, by dying before reaching reproductive age, failing to find a mate, failing to best competitors in attracting a mate, or failing to keep their own offspring alive so that they could mate—we would not be here to ponder the momentous issues of good and evil. As end products of this vast chain of events operating over deep evolutionary time, modern humans carry with them the adaptations that led to their ancestors' success and the genes that contribute to the reliable development of these adaptations. These adaptations comprise our universal human nature.

Aside from genetically identical twins and lifelong monogamous mates, the fitness interests of all individuals are, to some degree, unique and diverge from each other. Stated differently, humans are, to some extent, reproductive competitors with other humans to become ancestors. Competition need not be direct and need not involve overt contests. Indeed, competitors need never meet for competition to ensue. Scramble competition, for example, involves striving for the acquisition of limited or better resources in the external environment. Intrasexual competitors can compete with each other in individual courtship displays to attract a particular mate. Parents can compete with other parents merely by investing

in their children's success. Although some of these forms of individual competition, such as investing in children, do not correspond to human intuitions about competition, they do embody competition at the formal level of natural selection, as much as two stags locking horns in direct combat or two humans clawing each other psychologically to get ahead in the status hierarchy.

Since all modern humans are the descendants of ancestors who succeeded countless times in direct and indirect competition, modern humans carry with them the competitive adaptations that led to their ancestors' success, and pass on the genes that contribute to the development of these adaptations to their children. Some of these adaptations function to inflict costs on other humans.

### WHY HUMANS INFLECT HARM ON OTHER HUMANS

At an abstract level, there are two fundamental strategies for besting a competitor in a fitness contest. One strategy involves the acquisition of benefits that aid fitness—for example, scrambling for superior access to resources, displaying more alluring attractant signals to a mate, bestowing on children resources that aid their reproductive success, or aiding one's kin in a manner that increases inclusive fitness (Hamilton, 1964). The other strategy involves inflicting costs on competitors—for example, impairing their access to resources, interfering with their mate attractant signals, or harming a competitor's kin.

In the world of nonhuman animals, both strategies are seen in great abundance. Baby birds compete for their parents' food resources by "begging" with beaks wide open, but they also sometimes push a sibling out of the nest and hence commit siblicide. Male scorpionflies compete for females by securing insects to feed them as part of the nuptial gift, but they also jostle competing males away from the female, inflicting physical costs on their rivals. Among humans, intrasexual strategies of mate competition involve both sending attraction signals (Buss, 1988) and verbally derogating rivals (Buss & Dedden, 1990; Schmitt & Buss, 1996). Although damaging a rival's reputation may not be considered "evil" in the grand scheme of things, from the victim's perspective, the lost social status and consequent failure in mate competition may seem evil. Indeed, these status losses sometimes drive people to kill those they perceive as having harmed them. The emotion of vengeance may have evolved as a defense designed to staunch such costs or deter others from inflicting similar costs in the future.

In summary, we can expect selection to have favored the evolution of some adaptations that function to inflict costs on intrasexual rivals specifically and conspecific competitors generally. These costs vary from small to

large in the currency of fitness damage to the recipient. As the fitness costs grow in magnitude, we become more and more inclined to label the actions as evil. According to our theory of the evolution of evil, humans have adaptations to inflict these costs—adaptations to steal rivals' resources, adaptations to damage rivals' reputations, adaptations to physically injure rivals, and adaptations to steal their mates. Humans also are likely to have evolved adaptations to kill (Buss & Duntle, 2003).

### KILLING AS PROTOTYPICALLY EVIL

Probably no other class of human action is judged to be more evil than premeditated murder, and there may be no other class of actions that inflicts a greater fitness cost on the victim (Buss & Duntle, 2003; Duntle & Buss, 2003). Although no formal theory is needed to tell us that it is bad to be dead, killing is worse for a victim's fitness than is currently recognized by any existing psychological theory, except the homicide adaptation theory (Buss & Duntle, 2003). First, by being killed, the victim forfeits all future reproduction. He loses sexual access to his current partner as well as all future mating opportunities he may have acquired if he remained alive. With his death, he is no longer around to invest in his children. His children's survival and reproduction become imperiled as a result of his untimely death. It is known that the death of a parent can impair the survival of children, in some cultures by as much as 10% (Hill & Hurtado, 1996). If the children live and his mate remarries, the victim's children become stepchildren—the single largest risk factor for physical abuse and child homicide (Daly & Wilson, 1988). Furthermore, the victim's extended kin—his brothers and sisters, aunts and uncles, nephews and nieces, grandparents and grandchildren—all become more vulnerable as a consequence of his death through the loss of his protection and the perception of his family as exploitable. And if all of these fitness costs are not bad enough, his rivals benefit from his death. His mate becomes a potential sexual partner for his rivals. His resources become available for their taking. And his rivals' children now have a competitive edge over the victim's own children. His death, in short, can become his rivals' gain. In summary, killing may inflict more momentous fitness costs on a victim than any other single act—such a consequence constitutes prototypical evil from the perspective of the victim and the victim's friends and kin.

Reversing the perspective from victim to perpetrator yields interesting insights. As a thought experiment, consider that your assignment, should you decide to accept it, is to outreproduce your rivals. You can achieve this goal by various means—for example, besting them in the quest for high-quality food, developing more hygienic practices to better combat parasites and diseases, cultivating strategies that succeed in better attract-

ing desirable mates, or investing more heavily or more skillfully in your children. But one remarkably effective strategy remains by which you could accomplish your mission in dramatic fashion—killing your rivals.

From the perspective of the inclusive fitness of the killer, killing a rival, in principle, can offer a bounty of benefits. By killing a rival, you may gain access to the rival's resources, since the rival is not around to protect them—resources such as land, food, tools, weapons, or shelter. Since rivals are often in competition for the same pool of potential mates, killing a rival can eliminate mating competition. The rival's existing mates become potential new mates for the killer. The killer's current and future children may have less competition in the next generation, thereby enhancing their fitness. The victim's losses, in short, can become a killer's gains.

This brief description of the potential benefits of killing a rival, of course, ignores the costs of killing, and indeed, killing can be a dangerous and costly strategy to carry out. Killers risk being injured or killed while attempting to carry out a murder. Even if "successful," the kin of the victim may extract revenge in the future. In some cultures, killers suffer retribution from the larger group. Furthermore, killing may harm the reputation of the perpetrator, hindering future access to social resources, including mates. The key point is *not* that killing is always, or even often, beneficial to the fitness of the killer. Rather, *killing historically has been potentially beneficial in the currency of reproductive fitness under some delimited circumstances* (i.e., when the risks are low, when costs are unlikely to be incurred, when the potential yield is large in magnitude, or when killing is the least costly strategy available amid an array of costly options). As a result of these benefits to the killer, combined with the costs to victims, as soon as homicide enters a population as a strategy, evolution will immediately select for coevolved defenses, resulting in an "arms race" of antagonistic coevolution.

#### DEPENDING AGAINST EVIL: ANTAGONISTIC COEVOLUTION

There can be little doubt that, from the victim's perspective, their killers or would-be killers would be considered evil. Before we consider the possible evolution of a universal cognitive category of evil, however, it is critical to consider the evolutionary events that would be set into motion once killing entered the human strategic repertoire. Because of the dramatic fitness costs of being killed, selection would act strongly to create defenses against killing—what we have called anti-homicide mechanisms (Duntley & Buss, 1998).

The intensity of selection is generally a function of two critical factors: (1) the fitness consequences, and (2) the frequency of the fitness-relevant

events. There is no doubt that being killed inflicts enormous fitness damage on victims, fulfilling the first criterion. Given the magnitude of damage, the frequency of killing need not be high at all for selection to act consistently and strongly in fashioning anti-homicide defenses. The lifetime odds of being killed in the United States are roughly 1 in 200; they are 1 in 26 for certain groups such as inner-city males (Ghiglieri, 1999). Among more traditional societies, such as the Ghibusi tribe of Africa or the Yanomamo of Venezuela, as many as 30% of all males die at the hand of their fellow humans (Chagnon, 1988; Ghiglieri, 1999). Even among the so-called "peaceful" !Kung San of Botswana, murder rates are higher than in Los Angeles or Detroit. Paleontological evidence, which reveals arrowheads lodged in rib cages and crushing blows to ancient skulls, points to a long human history of killing (Buss & Duntley, 2003).

Although it is impossible to determine with precision the exact frequency of homicide over the long course of human evolutionary history, available evidence suggests that it was likely to be far from uncommon. And given the large fitness impact of being killed, even small rates of killing, such as the 0.5% rate that exists currently in the United States, would easily have met the required criteria for selection to have operated to fashion anti-homicide defenses.

Indeed, humans likely have evolved many different types of anti-homicide defenses (Duntley & Buss, 2003). Stranger anxiety, for example, is an excellent candidate for an evolved anti-homicide strategy. It emerges predictably at 7–8 months of life, is specific to male strangers (who historically have been more dangerous to infants), and appears to be universal across cultures (Heerwagen & Orions, 2002; Marks, 1987; Marks & Nesse, 1994). Other potential anti-homicide defenses include ethnogenesis, fleeing mechanisms, mind-reading abilities specialized for detecting homicidal intent, and many others (Duntley & Buss, 2003).

Because humans risk getting killed in many different circumstances, a single anti-homicide adaptation would have been insufficient to combat all of the dangers. Being killed in infancy is different from being killed by an intrasexual rival in a status dispute as an adult. Being killed in a status dispute is different from being killed by a jealous mate who has suddenly discovered a sexual infidelity. Being killed by an enraged mate is different from being killed in tribal warfare. Given the many, varied, and evolutionarily recurrent circumstances in which the lives of humans have been endangered, selection will have forged a complex armament of defensive, anti-homicide devices.

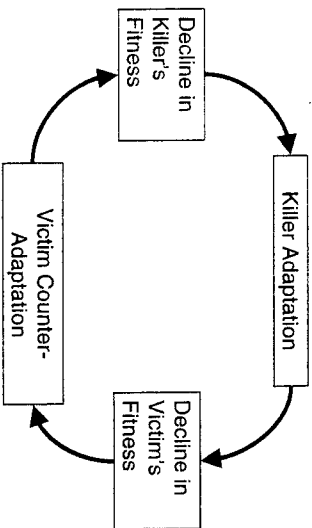
Once anti-homicide defenses begin to evolve, however, killing becomes a more costly strategy to pursue. First, the success rate of the strategy becomes lower as a consequence of anti-homicide defenses, rendering fewer fitness benefits to the killer. Second, attempting to kill can be downright costly to the killer. Killers risk injury from intended victims, and they

risk death, since "killing to prevent being killed," or killing in self-defense, is undoubtedly one of the anti-homicide defenses. Indeed, sometimes entire coalitions of individuals join forces to kill a killer. The upshot is that pursuing a homicidal strategy becomes less evolutionarily profitable as anti-homicide defenses evolve.

These anti-homicide defenses, by making killing less profitable, set in motion another evolutionary process: The coevolution of killer adaptations designed to circumvent the anti-homicide defenses. Selection will favor design features in killers that choose circumstances in which the costs of killing will not be incurred—for example, when the intended victim is particularly vulnerable or weak, when the intended victim lacks kin or coalitional support, or when the victim is caught by surprise. Selection will favor deceptive strategies in killers that include concealing homicidal intent from the victim in order to circumvent the activation of the victim's anti-homicide defenses.

As these more refined killer adaptations begin to evolve, selection will then favor the further evolution of increasingly sophisticated defenses. The consequence is a perpetual antagonistic coevolutionary arms race, as depicted in Figure 5.1.

From the victim's perspective, of course, being the target of a homicide renders the would-be killer evil. From the killer's perspective, how-



**FIGURE 5.1.** Antagonistic coevolution. This figure depicts the hypothesized process of the coevolution of homicide adaptations and anti-homicide defenses. Once an adaptation to kill conspecifics evolves within a population, it produces a decrease in the fitness of the killer's victims. This sets into motion countervailing selection that favors the evolution of defenses that function to prevent falling victim to murder. These defenses, in turn, create selection pressure for the evolution of more refined killer adaptations, such as those that circumvent the defenses of the victim (e.g., by using surprise or deception). To the extent that context-specific killing continues to provide fitness benefits to killers (on average), the coevolutionary arms race between homicidal adaptations and anti-homicide defenses will be perpetual in a way analogous to the coevolution of predators and prey.

ever, eliminating a victim may represent a "good," and the victim's anti-homicide strategies would thereby be viewed as evil.

### THE FUNCTIONS OF COGNITIVE EVALUATIVE MECHANISMS OF "EVIL" AND "GOOD"

It is not implausible that selection has fashioned privileged and universal cognitive categories of "evil" and "good" reserved precisely for monumental fitness-impairing motivations and fitness-bestowing motivations, respectively. The categorization of specific others as evil serves an important function: It targets a key threat to an individual's fitness, serves to encode an array of relevant fitness-damaging events in memory for subsequent strategic retrieval, and motivates action designed to circumvent the fitness threat (see Buss, 1989, for a more general argument about emotions as adaptations designed to deal with strategic interference).

Consider the case of the stepdaughter who is consistently beaten and sexually abused by her stepfather—events known to occur in nontrivial frequencies (Daly & Wilson, 1998). The stepdaughter's categorization of her stepfather as evil serves several important functions: (1) It tags fitness-damaging events for storage, (2) causes emotional arousal in response to fitness-damaging events (such as sexual abuse), and (3) motivates action designed to escape the fitness-damaging events. The fact that stepchildren leave home earlier, on average, than children who reside with both genetic parents may reflect this motivated action (Mitchell, Wister, & Burch, 1989). Historically, evaluative categories such as "evil" would have facilitated avoidance of cost-inflicting individuals.

The evolved category of evil, however, is likely to be a great deal more complex than this. Our ancestors faced many recurrently costly entities. It is likely that the category of evil only applies to some of those entities, and then to different degrees. We propose that an entity is more likely to be perceived as evil when it (1) engages in behaviors that inflict asymmetrically high levels of cost on a person relative to the benefits it receives, and (2) appears to desire to inflict harm.

An entity that inflicts great costs for small gains is more likely to be put into the special category of evil than when there is no asymmetry between costs inflicted and benefits received. Other factors equal, a person who kills for 10 million dollars is viewed as less evil than a person who kills for 10 dollars. And a person who kills someone in self-defense, in order to avoid being killed him- or herself, may not be viewed as evil at all. We propose that psychological mechanisms evolved to recognize other individuals who inflict great costs for relatively little gain. Categorization of such individuals as evil would serve the important adaptive function of helping to avoid becoming a victim of heavy costs.

The *intent* to do harm is an integral part of the definition of evil (Weiner & Simpson, 1989). It is not enough that some entity is costly to one's fitness. The entity must desire, or be perceived to desire, to inflict those costs. Over our evolutionary history, some entities would have been more recurrently associated with cost-inflicting events than others. Based on recurrent evolutionary experiences with cost-inflicting entities, the perception of desire to inflict costs likely evolved to be greater for some entities than others. It makes evolutionary sense that those entities that recurrently exhibited a closer causal relationship between their behavior and the infliction of costs would be perceived as more desirous of inflicting those costs and hence more evil. Snakes, for example, should be considered to be more evil than flowers (Mineka, 1992). And other humans (Buss & Duntley, 2003) can be considered the most evil entities of all. In the United States, homicide is among the top four leading causes of death for people between the ages of 1 and 44 (National Center for Health Statistics, 2002). The other causes of death include disease, accidents, and suicide. Of all causes of death, homicide is the only one with a directly and recurrently identifiable causal agent: another human. Over evolutionary time, this trend would have contributed to the evolution of a perception of other humans as potentially evil. However, the recurrent association of other humans with cost-inflicting behaviors, such as homicide, was not the only factor that must have contributed to the evolution of our perceptions of evil. Uncertainty about the intentions of others also would have played an important role.

There is a degree of uncertainty associated with all of our interactions with other humans. What does the other person hope to gain from this interaction? How much does he or she stand to lose as a result of this interaction? Over the history of our species, our ancestors would have had imperfect knowledge of the important factors associated with determining the likely costs and benefits of an interaction. They would have chosen specifically to engage in interactions that were likely to yield the greatest benefit at the lowest cost. Prior experience with an individual would have been among the most accurate ways of determining whether an interaction would be beneficial or costly, whether the intentions of another individual were good or evil. In the absence of prior experience, how would our ancestors have determined the intentions of others?

Error management theory (Haselton & Buss, 2000) proposes that humans evolved strategic cognitive biases that lead them to avoid high costs and not overlook significant benefits when making decisions under conditions of uncertainty. For example, men, but not women, have been shown to overperceive the sexual interest of a member of the opposite sex who acts friendly toward them (Haselton, 2002). This bias functions to decrease the likelihood that a man will overlook a situation where a significant boost to his fitness may be achieved.

A lifelong apprehensiveness or anxiety about the intentions of strangers, particularly men who are unknown, may be a cognitive bias that evolved to protect us from individuals who may want to inflict costs upon us for their own gain. The ancestral costs of assuming that an unknown male had beneficent intentions when, in fact, he intended to inflict harm would have been great. These costs would have provided selection pressure for the evolution of cognitive biases to assume that unknown individuals, particularly men who are more likely to inflict costs, had hostile intentions. In tribal societies, it is not uncommon that an approaching unknown adult male is intent on inflicting costs (Chagnon, 1983).

Some researchers have argued that our psychology of stereotyping may be an adaptive, energy-saving device (Macrae, Milne, & Bodenhausen, 1994). We propose that stereotyping may be adaptive by helping to make judgments under conditions of uncertainty. In the absence of other sources of information, a stereotype of another individual, though not completely accurate, would be better than no information at all in making decisions about how to interact with that person. Selection would have favored the evolution of stereotyping as a buffer against uncertainty if it provided individuals with an advantage in avoiding hostile conspecifics. Stereotyping would have evolved to help avoid heavy costs even at the price of missing out on some potentially beneficial interactions. The fact that most stereotypes are negative suggests that they are patterned in our psychology by evolved mechanisms designed to avoid some of the potentially heavy costs of interacting with unknown individuals.

#### PRINCIPLED PERSPECTIVE SHIFTS IN EVIL BASED ON FITNESS CONFLICTS

It should be clear from the preceding discussion that categorizations of good and evil hinge heavily on the perspective of the victim versus the perpetrator—a point that has been made by others (e.g., Baumeister, 1997). An evolutionary account, however, renders this intuitively obvious point a great deal more precise. In order to understand why, it is necessary to highlight the different levels of evolutionary analysis, and the precise ways in which perspectives on evil may shift predictably.

Consider a woman who is pregnant. She is young and has many potential reproductive years ahead of her. However, she is unmarried, lacks extended kin in close proximity, and lives in a nutritionally impoverished environment. In these and other circumstances, women often "spontaneously" abort the growing fetus. From the fetus's perspective, however, this is its only shot at life, which sets the stage for mother-fetus conflict (Hang, 1993). What is in the best fitness interests of the fetus (being born) differs fundamentally from what is in the best fitness interests of the incipient



mother (aborting the fetus). A large body of evidence points to the co-evolution of adaptations in both mother and fetus to deal with this conflict. The perpetual arms race produces adaptations that may damage the mother (e.g., by producing hypertension). The key point is that what we often view as a harmonious relationship of self-evident unity of interests—mother and growing child—is actually fraught with evolutionary conflict. In this example, of course, neither the growing fetus nor the mother typically categorize the other as evil. But the disharmony between mother and fetus demonstrates that conflicts permeate all relationships from the moment of conception.

Although selection is generally regarded as operating in the strongest manner at the genic level, it can play out at many levels of analysis. What is good for a particular gene in the currency of fitness may be bad for the other genes in an individual's genotype, resulting in the phenomenon known as *intra-genomic conflict* (e.g., Cosmides & Tooby, 1981). Furthermore, what is good for the genes of an individual can be bad for the genes of other individuals, leading to conflict between conspecifics. What is good for one intrasexual competitor (e.g., ascending in status) might be bad for another (e.g., being supplanted in the status hierarchy). What is good for one woman, such as mating with a highly desirable man, inevitably entails inflicting costs on intrasexual competitors who lose out and fail to obtain this particular mate. What is good for a man (e.g., obtaining sexual access to a woman through deception) might be bad for the woman (e.g., suffering damage to her reputation or other consequences of being deceived). What is good for one kinship group might be bad for another. What is good for one coalition of individuals might be bad for another. And, what is good for one species (e.g., surviving through predation) might be bad for another species (e.g., those preyed upon).

In sum, evolutionary psychology adds theoretical precision to the intuitively grasped perspective differences that characterize victims and perpetrators of horrendous deeds by yielding a precise analysis of the specific forms that conflict takes.

### THE EXPLOITATION OF EVIL

Once psychological mechanisms have evolved to place other individuals into the cognitive category of "evil," these mechanisms can be exploited by others for their own purposes. People often exploit these mechanisms in order to forge alliances designed to inflict costs on competing individuals or groups. Just as President George W. Bush used the label "axis of evil" in an effort to forge alliances with other groups, Osama bin Laden labeled the United States "evil" in order to forge Arab alliances with other groups and to motivate and justify the infliction of massive fitness costs on Americans.

Stated differently, once the cognitive category of evil exists, it can be exploited by individuals or groups to justify the perpetration of massive fitness costs on their enemies. The exploitation of labeling others as evil operates in several ways: (1) it motivates others to join in, amplifying the fitness damage inflicted on enemies; (2) it decreases the overall costs of pursuing a strategy of inflicting such damage because the larger coalition renders success more likely and defeat less likely; and (3) it justifies to nonparticipants the validity of inflicting costs, thus lowering the odds that nonparticipants will ally with the victims.

The universal category of evil can be exploited in another way. Some individuals actively cultivate a reputation as evil to exploit and avoid being exploited by others. In the Iraqi regime of Saddam Hussein, dissenters were often killed, decreasing the likelihood of challenges to Hussein's power. These actions may be regarded as attempting to exploit the evolved psychological mechanisms in others that perceive evil in order to achieve particular ends.

Richard Dawkins notes that religion can be used as a vehicle to promote evil: "My point is not that religion itself is the motivation of wars, murders, and terrorist attacks, but that religion is the principal label, and the most dangerous one, by which 'they' as opposed to 'we' can be identified at all" (Dawkins, 2003, p. 158). Whether one subscribes to a particular belief system or not, religion is a pervasive feature of human experience. Some have pointed out that it appears as though our minds were prepared for religion by evolution (Boyer, 2001). This may be the case, but it is more likely to have occurred in a different sequence: The ideas that comprise the most popular religions exist in the form that they do because they appeal to different psychological adaptations—adaptations that evolved to solve particular ancestral problems. One problem religion helps to solve is the identification of social allies and social enemies (Dawkins, 2003). Another is establishing and maintaining group solidarity. When there is competition between groups for some tangible resource, combined with powerfully reinforced religious beliefs that life has no end, psychological adaptations that embrace religious belief systems may contribute to misguided, cost-inflicting acts toward other groups who are perceived to be evil.

### THE EVOLUTION OF GOOD

Although this chapter focuses primarily on the evolution of evil, a few words can be said about the evolution of good. At one level of analysis, many of the arguments made for the evolution of evil can simply be reversed for the evolution of good. That is, people evaluate "good" when fitness benefits are delivered or received. The magnitude of the fitness bene-

fits is predicted to be highly correlated with judgments of good—an altruistic gift of a house or car would be judged to be more “good” than an altruistic gift of a candy bar. As in the evolution of evil, perspective matters greatly in evaluating good. People who deliver fitness-relevant benefits to oneself, one’s children, other genetic relatives, friends, and coalitional allies are “good.” People who deliver fitness benefits to one’s enemies are “evil.”

Just as humans have evolved adaptations to inflict costs on other humans, we have also evolved adaptations to bestow benefits. Evolutionary psychologists have explored three classes of benefit-bestowing mechanisms: (1) altruism delivered to genetic relatives (e.g., Burnstein, Crandall, & Kitayama, 1994); (2) reciprocal altruism delivered among friends or allies (e.g., Bleske & Buss, 2001); and (3) benefit-delivering mechanisms that do not involve kin altruism or reciprocal altruism and in which the giver does not incur a cost to deliver a benefit (e.g., Tooby & Cosmides, 1996). Parents sometimes sacrifice their own lives so that their children may live, an example of kin altruism. Friends bestow resources on each other, an example of reciprocal altruism. If I give you a ride to school when I am already going there anyway, I deliver a benefit to you without incurring a cost to myself. Evolution by selection has undoubtedly fashioned many benefit-bestowing adaptations in humans.

At another level, the things we tend to single out as especially deserving of the label “good” involve bestowing benefits at a great cost to oneself, *without any apparent return benefit*. Thus, the soldier who throws himself on a grenade to save his buddies, Mother Teresa’s devotion to helping others, the bystander who risks his life to save a stranger from drowning—these are all categorized as good and noble deeds. Indeed, it is precisely when there appear to be *no* return benefits to self, kin, or friends that we are especially prone to label a deed as admirable. The parent who donates \$100,000 to her child’s college education is not deemed as good as the person who makes the same donation to an impoverished stranger’s child. When a beneficent deed benefits self, kin, or allies, we tend to “discount” the amount of good we attribute to the person performing it. Thus, the evolutionary analysis of “good” is not strictly the mirror image of the evolutionary analysis of evil.

People undoubtedly exploit and manipulate perceptions of “good,” just as they exploit and manipulate perceptions of “evil.” Thus, we expect that people will sometimes put “spin” on their delivery of benefits to others, presenting their actions as more altruistic and less self-serving than they actually may be. Conversely, others may attempt to publicly discount apparent acts of good by pointing to some hidden benefit the giver is receiving. Indeed, many acts of apparent self-sacrifice turn out to have hidden benefits to the bestower, complicating the analysis even further. Soldiers wounded in war fighting for “freedom” or “their country” often

receive large boosts in prestige and social reputation through medals of valor. Women find these highly “altruistic” men to be especially attractive, and so the wounded soldier benefits in mating currency. Even the suicide bombers who commit acts regarded by their group as highly “good” receive benefits through “martyr” status and resources bestowed on their families.

In sum, an evolutionary analysis of “good” can be expected to shed much light on how people deliver benefits to others and manipulate the perceptions of others around the delivery of those benefits. Although these brief comments cannot do this complex topic justice, they suggest a few lines along which inquiry might proceed.

#### A COMPARISON OF THEORETICAL PERSPECTIVES ON EVIL

The use of violence against another person or group, particularly if it is perceived as unjustified, is undoubtedly the category of actions most likely to be viewed as evil (Anderson & Carnagey, Chapter 8; Baumeister & Vohs, Chapter 4, this volume). Violence is typically viewed as “not an effective way to get what one wants” (Baumeister & Vohs, Chapter 4, this volume). In response to the question “Why is there evil?”, Baumeister proposes that evil stems fundamentally from a failure of self-control, an inability to stifle aggressive impulses. Our theory of the evolution of evil suggests that these formulations of evil are fundamentally incorrect.

First, let’s consider the contention that violence is not an effective way to get what one wants (Baumeister & Vohs, Chapter 4, this volume). Citing the possibility that violent or aggressive strategies can result in prison sentences or other costs is beside the point from an evolutionary perspective. The key issue is whether selection has favored the contingent use of violence in some circumstances—whether the benefits of violence outweighed the costs, *on average*, relative to other strategic solutions over the sample space of relevant instances in the evolutionary past. Most social scientists, innocent of the logic of the evolutionary process, have intuitions that are wide off the mark when it comes to evaluating whether a particular strategy is “beneficial” or not. Selection, for example, can favor a strategy such as homicide, even if that strategy sometimes results in the strategist getting killed or imprisoned. The logic of this seemingly counter-intuitive argument becomes clear when one realizes that selection operates *not* on whether a particular strategy is effective *in every instance*, but rather on whether the benefits of the strategy outweigh the costs, *on average*, across the entire sample space of instances in which it is deployed (relative to competing designs present in the population at the time). Thus, selection can produce adaptations that result in many instances of failure—bullies who sometimes get beaten up, thieves who sometimes get

