Concepts and implications of altruism bias and pathological altruism

Barbara A. Oakley[1]

Department of Industrial and Systems Engineering, Oakland University, Rochester, MI 48309

Edited by John C. Avise, University of California, Irvine, CA, and approved April 9, 2013 (received for review February 14, 2013)

The profound benefits of altruism in modern society are self-evident. However, the potential hurtful aspects of altruism have gone largely unrecognized in scientific inquiry. This is despite the fact that virtually all forms of altruism are associated with tradeoffs—some of enormous importance and sensitivity—and notwithstanding that examples of pathologies of altruism abound. Presented here are the mechanistic bases and potential ramifications of pathological altruism, that is, altruism in which attempts to promote the welfare of others instead result in unanticipated harm. A basic conceptual approach toward the quantification of altruism bias is presented. Guardian systems and their over arching importance in the evolution of cooperation are also discussed. Concepts of pathological altruism, altruism bias, and guardian systems may help open many new, potentially useful lines of inquiry and provide a framework to begin moving toward a more mature, scientifically informed understanding of altruism and cooperative behavior.

cooperation | empathy | codependency | narcissism | philanthropy

> Reality must take precedence over public relations, for nature cannot be fooled.
>
> —Richard Feynman

Our eyes can be powerless against visual illusions, with our underlying neural machinery leading us to predictably erroneous conclusions about the size or shape of an object (1). In a similar fashion, our empathic feelings for others, coupled with a desire to be liked, parochial feelings for our in-group, emotional contagion, motivated reasoning, selective exposure, confirmation bias, discounting, allegiance bias, the *Einstellung* ("set") effect, and even an egocentric belief that we know what is best for others, can lead us into powerful and often irrational illusions of helping (2). In other words, people's own good intentions, coupled with a variety of cognitive biases, can sometimes blind them to the deleterious consequences of their actions. This dynamic of pathological altruism involves subjectively prosocial acts that are objectively antisocial. (Naturally, there are many objective perspectives. One seemingly objective observer's verdict of antisocial terrorism can be another's verdict of prosocial altruism, with the words "objective," "antisocial," "prosocial," "terrorism," and even "altruism" itself varying in meaning depending on the perspective of the putatively objective observer.)

At the core of pathological altruism are actions or reactions based on incomplete access to, or inability to process, the wide range of information necessary to make prudent decisions that align with cultural values associated with altruistic behavior. Various psychological, religious, philosophical, biological, or ideological biases could lead a person or group to misinterpret, selectively discount, or overly emphasize certain aspects of relevant information. Thus, pathologically altruistic behavior can emerge from a mix of accidental, subconscious, or deliberate causes. ["Altruism," in the context of this paper, is used to signify well-meaning behavior intended to promote the welfare of another; thus altruistic behavior may be motivated by concern for the other, egoistic concerns for the self, or both (e.g., "it makes me feel good to help them") (3). "Pathological" is used in the sense of being excessive or abnormal, without implying any clinical diagnosis.]

Pathological altruism can be conceived as behavior in which attempts to promote the welfare of another, or others, results instead in harm that an external observer would conclude was reasonably foreseeable. More precisely, this paper defines pathological altruism as an observable behavior or personal tendency in which the explicit or implicit subjective motivation is intentionally to promote the welfare of another, but instead of overall beneficial outcomes the altruism instead has unreasonable (from the relative perspective of an outside observer) negative consequences to the other or even to the self. This definition does not suggest that there are absolutes but instead suggests that, within a particular context, pathological altruism is the situation in which intended outcomes and actual outcomes (within the framework of how the relative values of "negative" and "positive" are conceptualized), do not mesh.

A working definition of a pathological altruist then might be a person who sincerely engages in what he or she intends to be altruistic acts but who (in a fashion that can be reasonably anticipated) harms the very person or group he or she is trying to help; or a person who, in the course of helping one person or group, inflicts reasonably foreseeable harm to others beyond the person or group being helped; or a person who in reasonably anticipatory way becomes a victim of his or her own altruistic actions (2). The attempted altruism, in other words, results in objectively foreseeable and unreasonable harm to the self, to the target of the altruism, or to others beyond the target. Examples at an interpersonal level include the codependent wife murdered by the husband she has refused to leave, or the overly attentive "helicopter" father who threatens to sue instructors that give well-deserved bad grades, or the mother who attempts to protect her son by refusing to vaccinate him and who consequently fuels a loss of herd immunity underpinning a local whooping cough epidemic in which an infant dies. Very different personalities can become entangled in pathologies of altruism, ranging from the sensitive hyperempath, to the normal person, to the utterly self-absorbed narcissist. These differing personalities share genuinely good intentions that play out in detrimental ways.

Sometimes there is a blurry line as to whether a problematic outcome for an altruistic action is reasonably foreseeable. This ambiguity can make it difficult to distinguish between altruism and pathological altruism. For example, let's say that, while altruistically helping a friend move to another apartment, you accidentally dropped and broke an expensive statue. Were your actions pathologically altruistic? In the conceptions of pathological altruism outlined here, no. Your altruism would not have been pathologically altruistic, because the bad outcome—the

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "In the Light of Evolution VII: The Human Mental Machinery," held January 10–12, 2013, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS Web site at www.nasonline.org/evolution_vii.

Author contributions: B.A.O. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: oakley@oakland.edu.

dropped statue—arose as a very unlikely and difficult-to-predict outcome of your good intentions. In a different scenario, however, let's say your brother becomes addicted to painkillers. When he goes through withdrawal, you get more painkillers to help him feel better, and you cover for him when his work supervisor calls. You genuinely want to help your brother, but the reality is that you are enabling his addiction. In this case, your well-meaning altruism is pathological.

These examples help clarify the concept of pathological altruism, but similar situations could be more ambiguous. What if you had dropped your friend's expensive statue after you had consumed a bottle of wine? Or what if your painkiller-addicted brother was waiting to be enrolled in a treatment program? We yearn for the definitive in conceptual definitions, but the reality is that there always will be a residual uncertainty.

Motives are also important. Well-meaning intentions can lead either to altruism or to pathological altruism. Self-servingly malevolent intentions, on the other hand, often have little or nothing to do with altruism, even though such malevolence can easily be cloaked with pretensions of altruism. A con artist soliciting for a "charity" that he uses to personally enrich himself would not be a pathological altruist.

Both altruism and empathy have rightly received an extraordinary amount of research attention. This focus has permitted better characterization of these qualities and how they might have evolved. However, it has also served to reify their value without realistic consideration about when those qualities contain the potential for significant harm.

Part of the reason that pathologies of altruism have not been studied extensively or integrated into the public discourse appears to be fear that such knowledge might be used to discount the importance of altruism. Indeed, there has been a long history in science of avoiding paradigm-shifting approaches, such as Darwinian evolution and acknowledgment of the influence of biological factors on personality, arising in part from fears that such knowledge somehow would diminish human altruistic motivations. Such fears always have proven unfounded. However, these doubts have minimized scientists' ability to see the widespread, vitally important nature of pathologies of altruism. As psychologist Jonathan Haidt notes, "Morality binds and blinds" (4).

Relevant here are the remarks of historian of science Thomas Kuhn, who observed that when a paradigm shift occurs, scientists see data for the first time (5). Such is the case with pathologies of altruism, which are not the commonly supposed rare aberrations, "but rather a behavior that overwhelmingly occurs in human social intercourse" (6). It therefore is realistic to encourage exploration of a new, scientifically based paradigm acknowledging that, even given differing semantic parsings, subjectively altruistic feelings sometimes can be objectively problematic and even ultimately antisocial.

The bottom line is that the heartfelt, emotional basis of our good intentions can mislead us about what is truly helpful for others. Altruistic intentions must be run through the sieve of rational analysis; all too often, the best long-term action to help others, at both personal and public scales, is not immediately or intuitively obvious, not what temporarily makes us feel good, and not what is being promoted by other individuals, with their own potentially self-serving interests. Indeed, truly altruistic actions may sometimes appear cruel or harmful, the equivalent of saying "no" to the student who demands a higher grade or to the addict who needs another hit. However, the social consequences of appearing cruel in a culture that places high value on kindness, empathy, and altruism can lead us to misplaced "helpful" behavior and result in self-deception regarding the consequences of our actions (7, 8).

Pathological altruism can operate not only at the individual level but in many different aspects and levels of society, and between societies. Recognizing that feelings of altruism do not necessarily constitute objective altruism provides a new way of framing and understanding altruism. This previously unrecognized perspective in turn may open many new, potentially useful lines of inquiry and provide a framework to begin moving toward a more mature, scientifically informed understanding of altruism and cooperative behavior. The thesis of pathological altruism emphasizes the value of true altruism, self-sacrifice, and other forms of prosociality in human life. At the same time, it acknowledges the potential harm from cognitive blindness that arises whenever groups treat a concept as sacred (4).

The public as a whole would benefit from knowledge that what might feel subjectively altruistic may have negative unintended consequences that both worsen the situation that was meant to be improved and impact other areas negatively. Even the government can work more efficiently when voters and legislators realize that attempts to help others come with very real costs and can have tradeoffs that worsen the very concerns that were meant to be alleviated.

Along these lines, then, this paper suggests that pathologies of altruism and of empathic caring should receive concentrated research focus. Specific recommendations are outlined as well. As an underlying motivation, we should remember that in the nineteenth and twentieth centuries, there was an unparalleled improvement in public health as the entire discipline of medicine came under scientific scrutiny. Medical therapies that at one time were thought to be "obviously" beneficial, such as bloodletting and blistering, were finally subjected to review that found them wanting. In a similar vein, if we are truly to help others, this new century at last forms the time for scientists to subject altruistic modern social engineering and activism efforts, as well as academic disciplines that hinge on "helping," and finally, altruism itself, to far more disciplined scientific scrutiny. It is time for dispassionate exploration of how altruism and empathy themselves can inadvertently bias our efforts to create truly cooperative modern, complex societies.

Evolutionary Considerations

In one sense, pathological altruism can be thought of as a pattern of nurturing or beneficial behavior with evolutionarily unsuccessful consequences. Evidence for antecedents of such behavior can be seen in the animal world; examples include the unwitting hosts of brood-parasitism, as with the wood thrush who devotes substantial resources to raising the offspring of cowbirds. Such antecedent behavior is manifest at even a genetic and molecular level. For example, beneficial replication processes within a cell can be co-opted by viruses (9). Consequent cell lysis or exocytosis allows the new viral bodies to spread the contagion.

Molecular perspectives, in fact, can inform how we perceive altruism and cooperative behavior. A stable molecular bond has the property that the bound state is a lower-energy configuration than the unbound state. A physical system tends toward the configuration that minimizes potential energy. Such "cooperative" behavior often needs an initial activation energy—that is, it comes at cost—but the resulting state resides more naturally and easily at the lower energy level for the newly formed single, integrated, cooperating entity. (This entity may or may not have replicative abilities.)

In these situations, pathological altruism or its antecedents might be thought of as arising in two ways. First, it can arise when other entities—systems that are not, or are no longer, integrated into the first cooperating entity—are able to tap into the lowered energy states and possible replicative abilities produced by the first cooperating entity. Tapping into those lowered energy states may weaken or destroy the first entity. (Initially, such secondary entities may be part of the first entity even as they begin their dissociation, as with precancerous cells. It also is worth noting that cooperative "entities" may be composed of different species, as with wrasses that swim with impunity into the mouths of groupers to feed off parasites, or with human intestinal flora.)

Second, pathological altruism or its antecedents can arise when the lowered energy state of the first system allows the system to

grow to such a size that it increases the potential for disintegration or destruction from noncooperative mechanisms affiliated with the entity. An example can be found in nuclear fission, where longer-range electrostatic repulsion between protons overcomes the attractive, albeit short-range, nuclear force between nucleons. In more complex cellular processes, the surface area-to-volume ratio limits the cell size. Doubling the size of the cell, for example, requires eight times more nutrients and would have eight times more waste, even though the surface area increased only by a factor of four.

We see these same cooperative versus noncooperative balances playing out on a larger, social scale. For example, the Amazonian Yąnomamö villagers preferred to live in small villages of around 40 people, which seemed to provide an optimal reduction in energy costs affiliated with daily needs for food and safety versus internal strife. However, villages of larger size provided more safety against other, potentially hostile villages. In other words, larger villages could, in some environments, be better at minimizing overall energy costs. Thus, some villages grew to more than 100 inhabitants in size. However, internal repulsive forces increased in the form of disputes that arose as the number of inhabitants in a village increased. Larger villages eventually fissioned, thus beginning the process anew (10). At a much higher level of social complexity, there was an initial economic boom as the European Union was first established. This boom has become tempered as internal nominally altruistic and cooperative efforts—the type of efforts that work fairly effectively in less complex social systems—are ultimately proving disputatious and disruptive.

As entities move to higher levels of complexity, the yin and yang of lowered energy states resulting from cooperation, versus noncooperative internal and external forces and effects, can cause boom-and-bust behavior on evolutionary timescales. How entities resolve these issues of cooperation versus noncooperation is a factor in determining whether entities self-destruct, proceed through cycles of growth versus decline, or are able to move successfully to still higher levels of complexity. Whenever higher levels of complexity are achieved, new issues of cooperation versus noncooperation develop, and the cycle begins anew.

One issue is clear. As entities become more complex, they generally develop evolving "guardian" type feedback mechanisms that allow not only the detection and mitigation of the effects of noncooperative mechanisms ("defectors") but also adaptation to changes in those noncooperative mechanisms. Without such flexible guardian systems, entities fall prey to other entities or to their own inherent noncooperative features. On a cellular level, we see that guardian immune systems have evolved from the rudimentary enzyme systems of unicellular organisms, which protect against bacteriophage infections, to the extraordinarily sophisticated immunological defense mechanisms seen in vertebrates.

Similarly, social systems of cooperative behavior must devise effective immunological guardian functions against efforts to siphon away the energetic advantages of cooperative behavior. Such immune guardian functions also must serve to mitigate disruptive internal forces and effects. (Of course, on a biological level, we see from the many varieties of autoimmune disease that immune-type guardian systems, even when designed with care, can create their own host of difficulties and can be hijacked by noncooperative elements, as with leishmaniasis or AIDS. Similar issues would appear to hold true for complex social systems.)

Thus, to the five mechanisms that have been posited for the evolution of cooperation—kin selection, direct reciprocity, indirect reciprocity, network reciprocity, and group selection (11)— must be added a sixth, guardian function. For cooperative behavior to continue in complex biological or sociological entities, that is, for entities not to fall prey to ever-present, ever-evolving defectors, some form of evolving active guardian function must be present that detects when debilitating or destructive advantage is being taken of cooperative or altruistic behavior. The guardian system must not only detect but also disable such noncooperative behavior or render the entity immune to the pernicious effects. Without such detection and mitigation mechanisms, we see modeled evolutionary entities that are wiped out by defectors (12).

Virtually all the mechanisms for the evolution of cooperation have some degree of overlap. Direct reciprocity, for example, perforce plays a role in indirect reciprocity. In a similar fashion, guardian functions overlap with the other five evolutionary cooperative mechanisms. Reciprocal strategies,such as tit-for-tat, for example, inherently contain what might be thought of as rudimentary and passive guardian functions: If you defect, I will defect. Differences in guardian function between groups could reinforce group selection mechanisms. Guardian functions also could relate to the reputational effects of indirect reciprocity in enhancing cooperation: I may report anyone who does not support the leader, because my family can suffer if I don't. By separating out guardian functions, which address the potential for support or damage to cooperative processes, vitally important mechanisms can be understood and more carefully modeled. Moreover, counterintuitive findings in complex cooperative social systems, such as the importance of selfish behavior and the tradeoffs of religious and ideological mechanisms in inducing and enforcing cooperation, can be clarified (13, 14). For example, poorly designed guardian functions that do not adequately account for Machiavellian leadership and behavior, might play an important role in the failure of social structures. In another example, strong guardian functions that might protect against some internal threats could simultaneously create stifling rigidity that renders the society less able to cope with other challenges. Over previous decades, medical science has come to appreciate the over arching importance of immune systems (themselves examples of guardian systems) in biology. Similarly, awareness of pathological altruism allows those analyzing the evolution of cooperation to appreciate the importance of the full panoply of guardian systems at the many different levels of complexity.

Implications

Let us step back briefly to explore how pathologies of altruism arise at an individual level. Naturally, the small percentage of toddlers and young children who show little concern for others seem predisposed for antisocial behavior as they mature (15). On the other hand, children who manifest altruistic behavior are generally well-adjusted. However, there is a small group of pathologically altruistic children who rate high on altruistic behavior but low on self-actualizing behavior such as showing pleasure at success or doing something on their own. For such children, a psychological cost can arise even at an early age, as shown by high scores in emotional symptoms, including unhappiness, worries, fear, nervousness, and somatization (16).

As neuroscience and genetics are beginning to elucidate the biological as well as cultural basis of altruistic and empathic behavior, it has become clear that individuals vary in their innate underpinnings involving empathy and altruism (17). Therefore an educational, religious, and societal "one size fits all" approach to enculturation that uniformly affirms the importance of altruistic caring, without a tempered acknowledgment of the tradeoffs, may inadvertently be harmful for some children in the long run. (In other words, social attempts to blindly encourage altruism become themselves a perfect example of pathological altruism.) Without insight into the undesirable effects arising from empathy and altruistic intentions, children and adults with an existing hypersensitivity toward others find it more difficult to detect and react appropriately to manipulation or to situations in which natural feelings of empathy could lead to undesirable outcomes.

Indeed, it seems that caring for others, helpful as it sometimes may be to those receiving or demanding that care, can have pernicious long-term consequences for the care giver, including guilt, burnout, depression, and stress disorders (18, 19). Stress resulting from empathic caring has been shown to produce errors

in medical treatment (20). Feelings of empathic caring also appear to lie at the core of dependent personality disorder, codependent behavior, and even anorexia (2). Caring, empathic, helicopter parents can, with the best of intentions, inflict lasting damage on their children (21).

Empathy is not a uniformly positive attribute. It is associated with emotional contagion; hindsight bias; motivated reasoning; caring only for those we like or who comprise our in-group (parochial altruism); jumping to conclusions; and inappropriate feelings of guilt in noncooperators who refuse to follow orders to hurt others (22–29). Oxytocin, the "goody-goody hormone" that underlies maternal bonding and many aspects of empathy, also increases both envy and gloating (30). Empathy also can be used by the self-serving, including psychopaths, to deduce how to further their own ends (31). Being emotionally close to someone who is selfish or dishonest has been found to lead people to becoming more selfish and dishonest themselves (32). Allegiance bias causes forensic scientists to call their findings for the team they believe has hired them (33). [Indeed, the reliability of all types of forensic science evidence, including ostensibly objective techniques such as DNA typing and fingerprint analysis, has been called to question (34).] Judges, almost all of whom are lawyers, favor the legal system in their decisions; this bias has far-reaching and deleterious effects on American law (35).

Quietly going along with the flow—refusing to blow the whistle on objectively criminal behavior, for example—also sometimes may be a form of pathological altruism that grows from our feelings of empathy. In other words, the altruism and empathy we feel often isn't really about the person or group ostensibly being helped but instead often are about us. Sometimes they relate to the pain we might feel at being ostracized or shunned for thinking or acting differently. Or they relate to building our reputation—we wish to be publicly perceived as being altruistic, whether or not our efforts are truly altruistic, so that we can receive the reputational benefits of indirect reciprocity. (Juries are notoriously magnanimous with other peoples' money.) Some would say that, once egoism is involved, the result is no longer altruism, so there is no such thing as pathological altruism. However, such an interpretation would also mean there is no altruism, because egoistic reward circuitry appears to be an important determinant of altruistic behavior.

As the work of Nobel laureate Daniel Kahneman, Jonathan Haidt, and others has shown, humans possess both intuitive fast and rational slow cognitive processes (4, 36, 37). Intuitions come first; reasoning follows to support that intuition (38, 39). Empathy is driven by fast processes. We often make snap judgments as a result of empathy and superficial notions of altruism [related to the "moral heuristics" described by Sunstein (40)]. Then, as both Kahneman and Haidt have explored in depth, we are experts at justifying emotionally based decisions with back-filled rationality. *Einstellung*, the inability to see another solution once an initial solution is prefixed in mind (41), means that a superficially helpful approach can become reified, further reinforced by motivated reasoning, selective exposure, belief perseverance, and growing overconfidence (42), along with moral heuristics such as those involving omission bias and outrage (40).

However, surprisingly, an individual can be oblivious to the consequences of these interwoven effects as a consequence of "bias blind spot" (43). In this fashion, an initial snap, commonsense judgment about what seems right in helping others can gel quickly into formidable certitude without consideration of important relevant facts. As noted by Mercier and Sperber, "there is considerable evidence that when reasoning is applied to the conclusions of intuitive inference, it tends to rationalize them rather than to correct them ... reasoning pushes people not towards the best decisions but towards decisions that are easier to justify" (42). Intelligence is no safeguard regarding these confirmation bias-related issues. Highly intelligent people, for example, do not reason more even-handedly and thoroughly; they simply are able to present more arguments supporting their own beliefs (44). As Columbia's Mark Lilla has pointed out "Distinguished professors, gifted poets, and influential journalists summoned their talents to convince all who would listen that modern tyrants were liberators and that their unconscionable crimes were noble, when seen in the proper perspective. Whoever takes it upon himself to write an honest intellectual history of twentieth-century Europe will need a strong stomach" (45). In fact, combating extreme confirmation bias has been called one of psychology's most pressing research priorities (46).

Sometimes it is appropriate to turn off or distance oneself from feelings of empathy, and it appears such emotional distancing can be learned (47, 48). In fact, it is clear that turning off empathy—becoming dispassionate—is normal in certain conditions, such as a surgeon performing surgery. Indeed, many hospitals have policies forbidding surgeons from operating on family members, a circumstance in which it would be more difficult to maintain a dispassionate stance.

In psychology, lack of awareness of limitations and tradeoffs regarding empathy has spilled over into the therapeutic process itself. Older therapists remember sayings such as "empathy defeats therapy" (49), but such attitudes have fallen away as psychologists increasingly have placed a premium on empathic care during the therapeutic process. In a related vein, within the field of nursing, the importance of empathy and compassion for patients is emphasized so unrelentingly that it would be reasonable to explore the possibility of a causal relationship between the unilateral focus on caring and the severe issue of burnout among nurses (50). Health care workers are not taught about the potential hazards of excessive or misplaced empathy; consequently, a gradual dehumanization process unfolds (51). An unconditional support of empathy and altruism makes matters so difficult for some members of general society that a counterculture of popular literature and support groups involving codependency has arisen. However, such approaches suffer from a lack of scientific merit or rigor (52).

It is clear that, without the support of science, it is impossible to steer societal mores toward a more nuanced understanding of altruism and empathy that ultimately can benefit everyone.

Extended Implications

There are broader implications related to these issues, particularly regarding the policy aspects of the scientific enterprise. Good government is a foundation of large-scale societies; government programs are designed to minimize a variety of social problems. Although virtually every program has its critics, well-designed programs can be effective in bettering people's lives with few negative tradeoffs. From a scientifically-based perspective, however, some programs are deeply problematic, often as a result of superficial notions on the part of program designers or implementers about what is genuinely beneficial for others, coupled with a lack of accountability for ensuing programmatic failures (53). In these pathologically altruistic enterprises, confirmation bias, discounting, motivated reasoning, and egocentric certitude that our approach is the best—in short, the usual biases that underlie pathologies of altruism—appear to play important roles.

For example, teen pregnancy has received substantive focus in recent years. Teenagers in the United States become pregnant, contract sexually transmitted diseases, and have abortions at much higher rates than teenagers in most other industrialized countries. However, the most effective, scientifically proven approaches to reducing teen pregnancy are often ignored. As psychologist Timothy Wilson noted in summarizing the many problematic efforts in this area: "The fact that policy makers learned so little from past research—at huge human and financial cost—is made even more mind-boggling by being such a familiar story. Too often, policy makers follow common sense instead of scientific data when deciding how to solve social and behavioral problems" (54). Policy-makers and policy-supporters,

in other words, are shaped by cohesive cognitive biases regarding their intentions to help others.

In yet another area, ostensibly well-meaning governmental policy promoted home ownership, a beneficial goal that stabilizes families and communities. The government-sponsored enterprises Freddie Mac and Fannie Mae allowed less-than-qualified individuals to receive housing loans and encouraged more-qualified borrowers to overextend themselves. Typical risk–reward considerations were marginalized because of implicit government support (55). The government used these agencies to promote social goals without acknowledging the risk or cost. When economic conditions faltered, many lost their homes or found themselves with properties worth far less than they originally had paid. Government policy then shifted to the cost of this "altruism" to the public, to pay off the too-big-to-fail banks then holding securitized subprime loans. For those who care about helping the needy in this country, or those who object to corporate bail outs, these trillion-dollar costs bring into high relief the immediate need for scientifically informed planning and evidence-based reevaluation. What is of primary concern here is that altruistic intentions played a critical role in the development and unfolding of the housing bubble in the United States, which in turn had enormous impact on the US economy. This recent history emphasizes the importance of studying not only altruism but also its biases and the consequences of those biases.

In foreign aid, $2 trillion dollars have been provided to Africa over the past 50 years. As chronicled by economist and former World Bank consultant Dambisa Moyo, a native of Zambia, such aid has resulted in measurably worsened outcomes in a broad variety of areas, supporting despotism and increasing corruption and a sense of dependency in Africans (56). In some cases, the money has been directly responsible for extraordinary damage (57, 58). Experienced foreign aid worker Ernesto Sirolli echoes many when he notes that much Western aid arises from narcissistic paternalism and patronization (59). We see here yet another situation where preconceived altruistic notions render it more difficult to focus on and react to indications supplied by data.

Viewing altruistic behavior as a source of both potentially positive and potentially negative influences may provide a framework for understanding better a variety of complex challenges. For example, one of the most important national issues of our time, as outlined in the National Academy Press publication *Choosing the Nation's Fiscal Future*, is the looming federal deficit (60). Ralph Cicerone, President of the National Academy of Sciences, and Jennifer Dorn, President and Chief Executive Officer of the National Academy of Public Administration, jointly wrote: "Much is at stake. If we as a nation do not grapple promptly and wisely with the changes needed to put the federal budget on a sustainable course, all of us will find that the public goals we most value are at risk."

How can such budgetary policies arise and continue? Arguably, their establishment and growth is cultivated by broadly Judeo-Christian cultural values and educational processes related to empathy and altruism. [Cultures can conceptualize empathy, altruism, and associated values in different ways (4, 61).] In this cultural perspective, empathy and altruistic intentions often are viewed as monolithically positive, nearly sacred qualities with negligible tradeoffs, whether or not the empathy is genuinely beneficial or the outcome of the altruistic intentions is truly altruistic. "It's the thought that counts," as the saying goes when discounting negative consequences of altruism.

A supportive bias for claimed altruistic efforts appears to have contributed not only to a plethora of economic woes but also to a continuing record of difficulties in the social sciences, where programs, theories, and therapies with altruistic intent—particularly those which coincide with preconceived "obviously beneficial" notions of helping—do not appear to receive the same careful scientific scrutiny as less obviously well-intentioned programs (54, 62, 63). This lack of critical appraisal has been seen in vitally important areas such as the mitigation of posttraumatic stress disorder, the reduction of family violence, the elimination of racial prejudice, the reduction of sex differences in mathematics, and the lessening of adolescent behavior problems and drug use (64–71). In one example, a therapy called "Critical Incident Stress Debriefing" was broadly implemented throughout the United States to reduce posttraumatic stress disorder, even though this costly program simply did not work and, in fact, sometimes worsened the very stress it was meant to resolve (67).

Well-meaning but unscientific approaches toward altruistic helping can have the unwitting effect of ensuring that the benefits of science and the scientific method are kept away from those most in need of help. In the final analysis, it is clear that when altruistic efforts in science are presented as being beyond reproach, it becomes all too easy to silence rational criticism (62, 70, 72–78). Few wish to run the gauntlet of criticizing poorly conducted, highly subjective "science" which is purported to help, or indeed, of daring to question the basis of problematic scientific paradigms that arise in part from good intentions. Edward O. Wilson ran into just such a well-meaning buzz saw with the publication of his *Sociobiology*, as did Judith Rich Harris with *The Nurture Assumption* and Napoleon Chagnon with his studies of rates of violence among the Amazonian Yąnomamö (10, 79, 80).

To object to a scientific theory is one thing, but to object to a scientific theory that connects however tenuously to feelings of morality is quite another. Once morality plays a role, even at the most subliminal level, the formidable cognitive biases of altruism and its pathologies can swing into play. Perhaps for that reason different academic disciplines and specific topics within those disciplines show differing requirements for rigor. In disciplines related to helping people (which can encompass a surprisingly broad swathe of even hard-science topics), scientists' differing treatment of research findings that elicit altruism bias can skew the findings of seemingly objective science (81). As Robert Trivers has noted: "It seems manifest that the greater the social content of a discipline, especially human, the greater will be the biases due to self-deception and the greater the retardation of the field compared with less social disciplines" (82).

One of the most valuable characteristics of science is that, despite the obvious imperfection of biases in ostensibly objective scientists, it provides a potential mechanism for overcoming those biases. At the same time, altruism bias may be one of the most pernicious, hard-to-eradicate biases in science, because it involves even-handed examination of what groups of seemingly objective rational scientists subliminally have come to regard as sacred. [Biases and belief systems can have a sense of the sacred even when not formalized as religions (4).]

As noted previously, many government programs are indeed beneficial, and some are invaluable in allowing the population as a whole to live meaningful lives supported by a safety net for life's inevitable difficulties. However, the National Academy Press publication *Choosing the Nation's Fiscal Future* documents that the federal deficit is clearly heading for a crisis. In other words, as a result of manifold individual decisions, many of which were based on very real intentions to help others, everyone is at risk for serious harm. Such crises may arise, not as a tragedy of the commons, but rather, as a tragedy of altruism.

In the small social groups which characterized most of human history, altruism bias and pathologies of altruism would have had few means for extending broad influence. In modern times, with the mass outreach potential of a few well-intentioned individuals or influential groups, who often have little or no ultimate accountability for programmatic failures or other detrimental effects, pathologies of altruism can assume enormous importance. It is reasonable to help shift the scientific and cultural paradigm and set the stage so that it becomes culturally acceptable, even expected, that one should attempt to quantify objectively purported claims of altruism. This paradigm shift is particularly important with regards to the budgetary tradeoffs and planning that form important aspects

of effective government that promotes cooperative behavior. The reality is, as made clear in the joint statement by the presidents of the National Academy of Sciences and the National Academy of Public Administration, that unless these types of considerations are made expeditiously, extraordinary cuts must be made in even the most genuinely beneficial programs (60). A voting public encouraged to follow a short-term, superficial, "feel good," emotionally-based heuristic for helping others is a voting public that much more easily can make poor long-term decisions.

Toward a Conceptual Framework

As scientists and engineers know well, "all models are wrong, but some are useful" (83). Embedded in any model is perspective, that is, the framework perceived by the developer. In the past, altruism (or cooperation) generally has been conceived and modeled as lying on a continuum between nonexistent and existent, much like the concept of eusociality (in which the opposite of eusociality is asociality; that is, no tendencies for grouping or socializing at all (84–87). ("Asocial" may also be considered, in some conceptions, to be "selfish" or "egoistic.") More recently, altruism has been conceived on a positive-to-negative continuum where negative altruism involves malevolent intentions, Machiavellianism, and psychopathy (88).

However, altruism can be framed in a third way, as a positive-to-negative continuum where negative altruism is altruism with antithetical consequences, i.e., pathological altruism. Viewing altruism in this way provides insights that relate to both individual personality traits and to large-scale modeling. There are tradeoffs to virtually all forms of altruism, and considering altruism as possessing both positive and negative aspects allows one to take more careful consideration of who is helped (the beneficiary) and who is harmed (the victim). Sometimes the same individual or group may be both helped and harmed. Parochial altruism—the combination of in-group altruism and out-group hostility—is positive altruism within one group but negative altruism for another. High taxes, for example, may be considered as positive altruism for one group and as negative altruism for another.

It should be noted that these conceptions formulate the problem primarily in terms of the altruism provider and stress the liability arising from, among other things, empathy and identification. It also is possible to formulate altruism as a dynamic process controlled in part by the altruism seeker (89). Moreover, the entitlements pressed for by the altruism-seeker may be either objectively helpful (for example, a scholarship sought by a hardworking student) or harmful (for example, alcohol sought by an alcoholic). In other cases, the altruism-seeker may desire seemingly infinitesimal acts of altruism that ultimately play a role in widespread long-term negative outcomes, as seen with grade inflation and social promotion. Jean Twenge and her research group have pointed toward substantive increases in narcissism in the population over the past decades, "Trends in positive self-views are correlated with grade inflation . . . but are not explained by changes in objective performance" (90, 91). It also may be that the actual help needed by those seeking or expecting help, as with Munchausen by Internet (in which Internet users feign a variety of ills to draw attention), involves something entirely different from what is sought.

Studies suggest that those involved in altruistic transactions benefit differentially from them, and egoism can play surprising roles. For example, sensitive children may have better personal outcomes if they behave egoistically in some instances. However, as shown with Twenge's work, other children appear to have unrealistic expectations when egoistic considerations are encouraged. The question thus arises: When and for whom is egoistic behavior beneficial or harmful? What is the relationship of egoism to altruism and—most importantly for our purposes—to pathologies of altruism? Further, how can we study these issues scientifically without our own judgments and moral righteousness intruding, guiding answers toward what we are certain will be helpful for others to hear rather than toward what the data actually reveal?

We can find clues as to how to proceed by examining prospect theory, where outcomes are assigned differing values depending on whether there are gains or losses. Losses hurt more than "feel good" gains. With altruism bias, it appears that people assign varying values to outcomes based on their underlying moral assessment. An example of such altruism bias was seen in subjects who were given a posthypnotic suggestion to feel a flash of disgust (an intimate part of moral judgment) when hearing a particular arbitrary word. Moral judgment—that sense of whether something is or is not helpful for others—could be made more severe by the presence of the arbitrary word (92). Researchers were surprised to find that even in a control situation where there was no apparent moral issue, the arbitrary words caused some subjects to make more negative moral judgments; later, the subjects confabulated stories to explain their behavior. Many factors have been shown to influence moral judgment at a subconscious level (4).

It appears that when a person attempts rationally (using the "slow" system) to calculate the utility of something that he or she already has judged through "fast" cognitive processes to be morally beneficial, skewed judgments are made, inflating the good outcomes and deflating the bad. Analogously, one can imagine that if malevolence was the goal, as with ill-intentions by a parochial in-group toward an out-group, benefits would be deflated and harms inflated.

A Path Forward

Personal-Scale Studies. Pathologies arising from altruism can be studied on an individual level. For example, many of the errors of judgment cited in the extensive listing in ref. 93 could result in altruism bias, or altruism bias could underlie and help lead to those judgment errors. In this regard, does the brain use a simple underlying "thumbs up" moral heuristic that leads "rational" thought processes to a foregone conclusion, as with the allegiance effect? Can such a heuristic be seen as a characteristic signature in medical imaging? Do individuals vary in their ability to influence their underlying moral heuristics? Are some individuals addicts of their feelings of self-righteousness? What varying effect does culture have on different individual's moral heuristics? On a side note, it appears that altruism bias, like many such biases, is a Jamesian fringe phenomenon of consciousness, much like the feeling of familiarity. It seems to grow from or relate to that little studied sense of rightness, of certitude—a tip-of-the-tongue feeling built on a web of biases, influences, and perceptions that one thing is beneficial, whereas another is not (94–96). Self-righteousness and pathologies of certitude have received almost no research emphasis (94, 95).

Narcissism, one of the most strongly heritable of all personality traits (97), has been similarly neglected. Narcissism is comorbid with many of the most troublesome personality disorders and dysfunctions, including psychopathy, borderline personality disorder, and bipolar disorder. So it comes as a surprise to learn that there are almost no hard-science imaging studies focusing on narcissism, although many other syndromes, as well as the positive aspects of empathy, have received keen research focus (98–100). Narcissism, in other words, deserves priority in imaging research.

Similarly, the vital topic of codependency has received almost no hard-science research focus, leaving "research" to those with limited or no scientific research qualifications (52). An indication of the popular need for and interest in this area is that a single book, *Codependent No More*, has sold more than five million copies over several decades. It is reasonable to wonder if the lack of scientific research involving codependency may relate to the fact that there is a strong academic bias against studying possible negative outcomes of empathy. Codependency, like narcissism, would thus be an important area of research in the elucidation of pathologies of altruism.

Broad-Scale Studies. At a larger scale, almost any data-driven model or projection in any discipline or government enterprise that even indirectly impacts an area of fairness or morality, or which contains significant potential for disciplinary bias, can be examined to see how well it actually has performed in the context of unfolding real-world data. Unexpected performance of the model or projection could be an indicator of altruism bias, and the bias could be quantified as to when, where, why, how, and to what extent it occurred. For example, a better understanding of altruism bias in data analysis and program development and implementation may provide insights regarding a great variety of phenomena, including the artificially inflated values of economic bubbles or various inadequate statistical measures (for example, those involving unemployment and economic growth) that can falsely boost the effects of well-meaning efforts. Concepts of pathological altruism thus can serve a normative purpose, helping us create better policies. Knowledge of how altruism bias distorts objective scientific inquiry can and should be considered a confounding factor when developing formal models.

It should be noted, however, that those possessing altruism bias would be most strongly biased to object to the very concept of altruism bias (101). Research has shown the near impossibility of reaching biased individuals using rational approaches, no matter their level of education or intelligence; such attempts can be likened to squaring the circle (44, 46).

In another vein, researchers from outside a given discipline, and who are thus less vested in the theories of that domain themselves, could initiate studies to determine whether insufficient statistics, exaggerated claims, drawing the wrong conclusions from other papers, or using data selectively to confirm hypotheses might differ among studies that relate to disciplinary biases or moral issues (many hard-science topics ultimately impact issues of deep moral concern) versus those that do not. Within scientific disciplines, the appearance of group-norm–enforcing signed petitions could be used as indicators of the potential for pathologies of altruism; such petitions might communicate important, albeit unintended, information about the health of a discipline. Are entire disciplines shaped by papers that are not submitted because of legitimate fears of rejection? As Santiago Ramón y Cajal, the father of modern neuroscience, perceptively wrote: "...the good will of scientists is usually so paradoxical that they are more pleased by the defence of an obvious error which has become wide-spread than by the establishment of a new fact." (102) These thoughts were echoed recently in a predictably controversial paper by John Ioannidis pointing toward the shockingly high publication rate of false research findings. Iaonnidis noted: "...for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias" (103). Can disciplinary biases be quantified, perhaps in studies put forth by interdisciplinary groups (including nonacademics) from largely outside the discipline in question? Group-think within disciplines, particularly in regards to differing editorial standards of proof required for studies that do not hold to a discipline's underlying moral paradigm, would be a particularly rich, important, and provocative area of study.

Lilienfeld points toward psychological treatments that "may produce harm in relatives or friends of clients in addition to, or instead of, clients themselves. For example, some treatments that are otherwise innocuous or even effective with clients could produce a heightened risk of false abuse allegations against family members" (67). Is it possible that some social advocacy and social justice efforts result in the same types of pernicious effects on a societal scale so that efforts to build cooperation instead inhibit it? We often do not know, because well-meaning advocates have made raising those questions a taboo. Framing issues in the form of pathologies of altruism and altruism bias forms a mechanism for breaking through the taboo and making dispassionate studies of when helping is truly helping and when it is contributing inadvertent harm.

Forensic studies of allegiance bias (33) could profitably inform academic disciplines as to how to examine the effects of altruism bias both within and outside academia, and indeed, in regards to greater academia itself. In the later regard, it seems academia is reaching multiple crises, often arising from well-meaning efforts; such crises include administrative bloat, college tuitions that have vastly outpaced inflation, and students who are left academically adrift (104).

Potential Steps to Address Altruism Bias in Academic Disciplines and the Scientific Enterprise. There are active steps that could be taken to prevent the potential for altruism bias within the scientific enterprise. In all-important journal review processes, for example, mixed panels of reviewers (e.g., cognitive psychologists and neuroscientists reviewing social psychological papers) could become standard practice (105). Doctoral programs can place heavier emphasis on the scientific method and careful use of statistics so that graduate students, who are themselves future journal reviewers, can learn to spot problematic submissions more easily and perhaps be less likely to conduct problematic research themselves. The many aspects of altruism bias and the problems as well as benefits of empathy can be much more broadly discussed and emphasized in textbooks, beginning even in high school and the early years of college. Disciplines heavily involved in social advocacy, whose primary goal involves truly benefitting others, should be among the first to take interest in incorporating these concepts and approaches into research and training programs, editorial efforts, and textbooks.

Conclusions

Science has put extraordinary emphasis on studying the helpful aspects of altruism, and this emphasis has helped reify altruism's benefits among the general population. However, if science is truly to serve as an ultimately altruistic enterprise, then science must examine not only the good but also the harm that can arise from our feelings of altruism and empathetic caring for others. In support of this idea, it is important to note that during the twentieth century, tens of millions individuals were killed under despotic regimes that rose to power through appeals to altruism (106–110). The study of pathological altruism, in other words, is not a minor, inconsequential offshoot of the study of altruism but instead is a topic of overwhelming scientific and public importance.

ACKNOWLEDGMENTS. The author thanks Philip Oakley and Joseph Carroll for extended discussions and Amy Alkon, Gary Barber, Augustine Brannigan, Bernard Berofsky, Kenny Felder, David C. Geary, Joachim Krueger, Hugo Mercier, Karol Pessin, Sally Satel, John Traphagan, and Carolyn Zahn-Waxler for cogent critical suggestions. The suggestions of the anonymous editor and reviewers also were very beneficial.

1. Shepard RN (1990) *Mind Sights* (W.H. Freeman & Co., New York).
2. Oakley B, et al. eds (2012) *Pathological Altruism* (Oxford University Press, New York).
3. Batson CD (2012) The empathy-altruism hypothesis. *Empathy: From Bench to Bedside*, ed Decety J (MIT Press, Cambridge, MA), pp 41–54.
4. Haidt J (2012) *The Righteous Mind* (Pantheon Books, New York).
5. Kuhn T (1970) *The Structure of Scientific Revolutions* (Univ of Chicago Press, Chicago) 2nd Ed.
6. Ayala F (2012) Foreword. *Pathological Altruism*, eds Oakley B, et al. (Oxford Univ Press, New York), pp 49-74.
7. Fine C (2006) *A Mind of Its Own* (W. W. Norton, New York).
8. Tavris C, et al. (2007) *Mistakes Were Made (But Not By Me)* (Houghton Mifflin Harcourt, New York).
9. Szathmáry E, Maynard Smith J (1997) From replicators to reproducers: The first major transitions leading to life. *J Theor Biol* 187(4):555–571.
10. Chagnon NA (2013) *Noble Savages* (Simon & Schuster, New York).
11. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314(5805): 1560–1563.
12. Nowak MA (2006) *Evolutionary Dynamics* (Belknap, Cambridge, MA).

13. Eldakar OT, Wilson DS (2008) Selfishness as second-order altruism. *Proc Natl Acad Sci USA* 105(19):6982–6986.
14. Wilson DS (2003) *Darwin's Cathedral* (Univ of Chicago Press, Chicago).
15. Rhee SH, et al. (2013) Early concern and disregard for others as predictors of antisocial behavior. *J Child Psychol Psychiatry* 54(2):157–166.
16. Oakley B, Knafo A, McGrath M (2012) Pathological altruism - an introduction. *Pathological Altruism*, eds Oakley B, et al. (Oxford Univ Press, New York).
17. Churchland PS (2011) *Braintrust* (Princeton Univ Press, Princeton, NJ).
18. Najjar N, Davis LW, Beck-Coon K, Carney Doebbeling C (2009) Compassion fatigue: A review of the research to date and relevance to cancer-care providers. *J Health Psychol* 14(2):267–277.
19. Eisenberg N, Eggum ND (2009) Empathic responding. *The Social Neuroscience of Empathy*, eds Decety J, Ickes WJ (MIT Press, Cambridge, MA), pp 71-83.
20. West CP, et al. (2006) Association of perceived medical errors with resident distress and empathy: A prospective longitudinal study. *JAMA* 296(9):1071–1078.
21. Locke J, Campbell MA, Kavanagh DJ (2012) Can a parent do too much for their child? *Aust J Guid Couns* 22(2):249–265.
22. Breithaupt F (2012) Empathy does provide rational support for decisions. But is it the right decision? *Emot Rev* 4(1):96–97.
23. Cikara M, Fiske ST (2011) Bounded empathy: Neural responses to outgroup targets' (mis)fortunes. *J Cogn Neurosci* 23(12):3791–3803.
24. Gutsell JN, Inzlicht M (2012) Intergroup differences in the sharing of emotive states: Neural evidence of an empathy gap. *Soc Cogn Affect Neurosci* 7(5):596–603.
25. Brosnan M, Ashwin C, Gamble T (2011) Greater empathizing and reduced systemizing in people who show a jumping to conclusions bias in the general population. *Psychosis* 5(1):1–11.
26. Frantz CMP, Janoff-Bulman R (2000) Considering both sides: The limits of perspective taking. *Basic Appl Soc Psych* 22(1):31–42.
27. Shiller RJ (2007) Understanding recent trends in house prices and home ownership. (Cowles Foundation for Research in Economics, Yale University, New Haven, CT).
28. Spitzer M, Fischbacher U, Herrnberger B, Grön G, Fehr E (2007) The neural signature of social norm compliance. *Neuron* 56(1):185–196.
29. Chang I (1998) *The Rape of Nanking* (Penguin, New York).
30. Shamay-Tsoory SG, et al. (2009) Intranasal administration of oxytocin increases envy and schadenfreude (gloating). *Biol Psychiatry* 66(9):864–870.
31. Dutton K (2012) *The Wisdom of Psychopaths* (Farrar, Straus and Giroux, New York).
32. Gino F, et al. (2012) Vicarious dishonesty. *Organ Behav Hum Decis Process* 119:15–26.
33. Murrie DC, Boccaccini MT, Guarnera LA, Rufino K, Are forensic experts biased by the side that retained them? *Psychol Sci*, in press.
34. National Research Council (2009) *Strengthening Forensic Science in the United States* (National Academies, Washington DC).
35. Barton BH (2010) *The Lawyer-Judge Bias in the American Legal System* (Cambridge Univ Press New York).
36. Kahneman D (2011) *Thinking, Fast and Slow* (Farrar Straus & Giroux, New York).
37. Cushman F, et al. (2010) Our multi-system moral psychology. *The Moral Psychology Handbook*, ed Doris JM (Oxford Univ Press, New York), pp 47–71.
38. Haidt J (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychol Rev* 108(4):814–834.
39. Zajonc RB (1980) Feeling and thinking. *Am Psychol* 35(2):151.
40. Sunstein CR (2005) Moral heuristics. *Behav Brain Sci* 28(4):531–542, discussion 542–573.
41. Bilalić M, McLeod P, Gobet F (2008) Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition* 108(3):652–661.
42. Mercier H, Sperber D (2011) Why do humans reason? Arguments for an argumentative theory. *Behav Brain Sci* 34(2):57–74, discussion 74–111.
43. Pronin E, Kugler MB (2007) Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *J Exp Soc Psychol* 43(4):565–578.
44. Perkins DN, Farady M, Bushey B (1991) Everyday reasoning and the roots of intelligence. *Informal Reasoning and Education*, eds Voss JF, Perkins DN, Segal JW (Lawrence Erlbaum, Hillsdale, NJ), pp 84–105.
45. Lilla M (2003) The Reckless Mind: Intellectuals in Politics. (New York Review of Books, New York).
46. Lilienfeld SO, Ammirati R, Landfield K (2009) Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspect Psychol Sci* 4(4):390–398.
47. Chiesa A, Serretti A (2009) Mindfulness-based stress reduction for stress management in healthy people: A review and meta-analysis. *J Altern Complement Med* 15(5):593–600.
48. Decety J, Yang CY, Cheng Y (2010) Physicians down-regulate their pain empathy response: An event-related brain potential study. *Neuroimage* 50(4):1676–1682.
49. Friedman EH (2008) *The Myth of the Shiksa* (Church Publishing, New York).
50. Kowalski C, et al. (2010) Burnout in nurses - the relationship between social capital in hospitals and emotional exhaustion. *J Clin Nurs* 19(11-12):1654–1663.
51. Haque OS, Waytz A (2012) Dehumanization in medicine: Causes, solutions, and functions. *Perspect Psychol Sci* 7(2):176–186.
52. McGrath M, Oakley B (2012) Codependency and pathological altruism. *Pathological Altruism*, eds Oakley B, et al. (Oxford Univ Press, New York), pp 49–74.
53. Sowell T (2012) *Intellectuals and Society* (Basic Books, New York) Revised Ed.
54. Wilson T (2011) *Redirect* (Little, Brown and Company, New York,).
55. Acharya VV, Richardson M, Van Nieuwerburgh S, White LJ (2011) *Guaranteed to Fail* (Princeton Univ Press).
56. Moyo D (2009) *Dead Aid* (Farrar, Straus and Giroux, New York).
57. Easterly W (2006) *The White Man's Burden* (Penguin Press, New York).
58. Polman L (2010) *The Crisis Caravan* (Metropolitan Books, New York) trans Walters L.
59. Sirolli E (1999) *Ripples from the Zambezi* (New Society, Gabriola Island, BC).
60. National Research Council and National Academy of Public Administration (2010) *Choosing the Nation's Fiscal Future* (National Academies, Washington, DC).
61. Traphagan JW (2012) Altruism, pathology, and culture. *Pathological Altruism*, eds Oakley B, et al. (Oxford Univ Press, New York), pp 272–287.
62. Wright R, Cummings N, eds (2005) *Destructive Trends in Mental Health* (Brunner-Routledge, New York, NY).
63. Cole S, ed (2001) *What's Wrong with Sociology?* (Transaction Publishers, New Brunswick, NJ).
64. Kalev A, Dobbin F, Kelly E (2006) Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *Am Sociol Rev* 71(4):589–617.
65. Petrosino A, Turpin-Petrosino C, Buehler J (2003) Scared Straight and other juvenile awareness programs for preventing juvenile delinquency. *Ann Am Acad Pol Soc Sci* 589(1):41–62.
66. Petrosino A, Turpin-Petrosino C, Finckenauer JO (2000) Well-meaning programs can have harmful effects! Lessons from experiments of programs such as Scared Straight. *Crime Delinq* 46(3):354–379.
67. Lilienfeld SO (2007) Psychological treatments that cause harm. *Perspect Psychol Sci* 2(1):53–70.
68. Eidelson R, Soldz S (2012) Does comprehensive soldier fitness work: CSF research fails the test. *Working Paper Number 1, May 2012* (Coalition for an Ethical Psychology, Bala Cynwyd, PA).
69. Sander R, Taylor S, Jr. (2012) *Mismatch* (Basic Books, New York).
70. Mills LG (2008) *Violent Partners* (Basic Books, New York).
71. Stoet G, Geary DC (2012) Can stereotype threat explain the gender gap in mathematics performance and achievement? *Rev Gen Psychol* 16(1):93–102.
72. Straus MA (2007) Processes explaining the concealment and distortion of evidence on gender symmetry in partner violence. *Eur J Crim Policy Res* 13:227–232.
73. Straus MA (2008) Bucking the tide in family violence research. *Trauma Violence Abuse* 9(4):191–213.
74. Straus MA (2009) Current controversies and prevalence concerning female offenders of intimate partner violence. *J Aggress Maltreat Trauma* 18(6):552–571.
75. Sunstein CR (2009) *Going to Extremes* (Oxford Univ Press, New York).
76. Satel S (2000) *PC, M.D.* (Basic Books, New York).
77. Sommers CH (1995) *Who Stole Feminism?* (Simon & Schuster, New York).
78. Sommers CH (2001) *The War Against Boys* (Simon & Schuster, New York).
79. Pinker S (2002) *The Blank Slate* (Viking, New York).
80. Dreger A (2011) Darkness's descent on the American Anthropological Association. A cautionary tale. *Hum Nat* 22(3):225–246.
81. Margolis H (1993) *Paradigms and Barriers* (Univ of Chicago Press, Chicago).
82. Trivers R (2011) *Folly of Fools* (Basic Books, New York).
83. Box GEP, Draper NR (1987) *Empirical Model-Building and Response Surfaces* (John Wiley & Sons, New York).
84. Trivers RL (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46(1):35–57.
85. Darlington PJ, Jr. (1978) Altruism: Its characteristics and evolution. *Proc Natl Acad Sci USA* 75(1):385–389.
86. Sober E, Wilson DS (1998) *Unto Others* (Harvard Univ Press, Cambridge, MA).
87. Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425(6960):785–791.
88. Baron-Cohen S (2012) *The Science of Evil* (Basic Books, New York).
89. Halabi S, Nadler A (2009) Receiving help. *The Psychology of Prosocial Behavior*, eds Stürmer S, Snyder M (Wiley-Blackwell, New York), pp 121–138.
90. Twenge JM, Campbell WK, Gentile B (2012) Generational increases in agentic self-evaluations among American college students, 1966–2009. *Self Ident* 11(4):409–427.
91. Twenge JM, *et al.* (2008) Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. *J Pers* 76(4):875–902; discussion 903–828.
92. Wheatley T, Haidt J (2005) Hypnotic disgust makes moral judgments more severe. *Psychol Sci* 16(10):780–784.
93. Krueger JI, Funder DC (2004) Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behav Brain Sci* 27(3):313–327, discussion 328–376.
94. Brin D (2012) Self-addiction and self-righteousness. *Pathological Altruism*, eds Oakley B, et al. (Oxford Univ Press, New York).
95. Burton R (2012) Pathological certitude. *Pathological Altruism*, eds Oakley B, et al. (Oxford Univ Press, New York).
96. Mangan B (2001) Sensation's ghost: The non-sensory 'fringe' of consciousness. *Psyche* 7(18). Available at www.theassc.org/files/assc/2509.pdf.
97. Torgersen S, et al. (2000) A twin study of personality disorders. *Compr Psychiatry* 41(6):416–425.
98. Allen EA, et al. (2011) A baseline for the multivariate comparison of resting-state networks. *Front Syst Neurosci* 5(2):1–23.
99. Garrity AG, et al. (2007) Aberrant "default mode" functional connectivity in schizophrenia. *Am J Psychiatry* 164(3):450–457.
100. Zaki J, Ochsner K (2012) The neuroscience of empathy: Progress, pitfalls and promise. *Nat Neurosci* 15(5):675–680.
101. Pronin E, Lin DY, Ross L (2002) The bias blind spot. *Pers Soc Psychol Bull* 28(3):369–381.
102. Ramón y Cajal S (1937) *Recollections of My Life* (MIT Press, Cambridge, MA).
103. Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* 2(8):e124.
104. Arum R, Roksa J (2010) *Academically Adrift* (Univ of Chicago Press, Chicago).
105. Stoet G, Geary DC (2013) Mixed blessings. *The Times Higher Education.* March 7, p 36 (lett).
106. Montefiore SS (2004) *Stalin* (Alfred A. Knopf, New York).
107. Chang J, Halliday J (2005) *Mao* (Jonathan Cape, London).
108. Waite RGL (1977) *The Psychopathic God: Adolf Hitler* (Basic Books, New York).
109. Short P (2004) *Pol Pot* (John Murray, London).
110. Li Z (1994) *The Private Life of Chairman Mao* (Random House, New York).