

Altruism & Altruistic Love

*Science, Philosophy, & Religion
in Dialogue*

Edited by

STEPHEN G. POST
LYNN G. UNDERWOOD
JEFFREY P. SCHLOSS
WILLIAM B. HURLBUT

OXFORD
UNIVERSITY PRESS

2002

CONTENTS

Contributors, xi

General Introduction, 3

*Stephen G. Post, Lynn G. Underwood, Jeffrey P. Schloss,
and William B. Hurlbut*

PART I: DEFINITIONS

Introduction to Part I, 15

Stephen G. Post

1. The ABCs of Altruism, 17

Elliott Sober

2. Pythagorean Bodies and the Body of Altruism, 29

Edith Wyschogrod

3. Morality, Altruism, and Love, 40

Jerome Kagan

4. The Tradition of Agape, 51

Stephen G. Post

Conclusion to Part I, 65

Stephen G. Post

PART II: HUMAN MOTIVATION AND ACTION

Introduction to Part II

Lynn G. Underwood, 69

5. The Human Experience of Compassionate Love: Conceptual Mapping
and Data from Selected Studies, 72

Lynn G. Underwood

6. Addressing the Altruism Question Experimentally, 89

C. Daniel Batson

7. Explicating Altruism, 106

Kristen Renwick Monroe

8. Extraordinary Acts of Ordinary People: Faces of Heroism
and Altruism, 123

Samuel P. Oliner

Conclusion to Part II, 140

Lynn G. Underwood

PART III. EVOLUTIONARY BIOLOGY

Introduction to Part III, 145

Jeffrey P. Schloss

9. A Darwinian Naturalist's Perspective on Altruism, 151

Michael Ruse

10. Relating Self, Others, and Sacrifice in the Ordering of Love, 168

Stephen J. Pope

11. The Fall and Rise and Fall and Rise and Fall and Rise of Altruism in Evolutionary Biology, 182

David Sloan Wilson and Elliott Sober

12. Some Obstacles to Altruism, 192

Melvin Konner

13. Emerging Accounts of Altruism: "Love Creation's Final Law"?, 212

Jeffrey P. Schloss

Conclusion to Part III, 243

Jeffrey P. Schloss

PART IV: THE SCIENCE OF ALTRUISM

Introduction to Part IV, 249

William B. Hurlbut

14. Implications for the Neurobiology of Love, 254

Thomas R. Insel

15. A Note on the Neurobiology of Emotions, 264

Antonio R. Damasio

16. Impairment of Interpersonal Social Behavior Caused by Acquired Brain Damage, 272

Hanna Damasio

17. The Communication of Emotions and the Possibility of Empathy in Animals, 284

Stephanie D. Preston and Frans B. M. de Waal

18. Empathy, Evolution, and Altruism, 309

William B. Hurlbut

Conclusion to Part IV, 328

William B. Hurlbut

PART V: RELIGION

Introduction to Part V, 333

Stephen G. Post

19. Science and Religion on the Nature of Love, 335

Don S. Browning

20. Separation, Attachment, and Altruistic Love: The Evolutionary Basis for Medical Caring, 346

Gregory L. Fricchione

21. Compassion out of Wisdom: Buddhist Perspectives from the Past
toward the Human Future, 362

Ruben L. F. Habito

Conclusion to Part V, 376

Stephen G. Post

Concluding Summary: Future Research Needs on Altruism and
Altruistic Love, 379

Stephen G. Post and Lynn G. Underwood

Annotated Bibliography: Research on Personality and Individual
Differences in Altruism, 387

Shelley Dean Kilpatrick and Michael E. McCullough

References, 475

List of Articles by Subject, 485

Index, 491

General Introduction

STEPHEN G. POST, LYNN G. UNDERWOOD,
JEFFREY P. SCHLOSS, & WILLIAM B. HURLBUT

In the context of human behavior, altruism, from the Latin root *alter* (meaning “other”), concerns the place of the other in moral experience, especially when the other is in need. An altruist intends and acts for the other’s sake as an end in itself rather than as a means to public recognition or internal well-being, although such benefits to self need not be resisted. In altruism, solipsism, the view that the self is all that exists or can be known (*solus*, meaning “alone,” and *ipse*, meaning “self”) is transcended: Self no longer perceives self as the only center of worth and no longer perceives worth in others only to the extent that they contribute to egoistic interests. In essence, as one is altruistic, one realizes the independence of others from his or her use and senses that his or her claims to ontological centrality are illusory.

Altruism, especially when it extends beyond biological relations (kin altruism) and beyond “tit-for-tat” calculations grounded in self-interest (reciprocal altruism), is widely lauded and is commonly considered a foundation of moral life, although it need not imply the total eclipse of self-concern or a quest for self-immolation. In its fullest expression, which may include significant self-sacrifice in the aid of strangers or even enemies, altruism is a source of perennial fascination across cultures. Regardless of its duration, intensity, emotional engagement, sacrifice, and extensivity, the common feature of altruism is affirmation of and care for “the other as other” (a phrase that seems ubiquitous in contemporary moral philosophy of “alterity,” especially in the context of phenomenology).

We use both “altruism” and “altruistic love” in the book title. Biologists use altruism to refer to actions that benefit the reproductive success of others at expense to the self, making no reference to motivation or conscious intentions. Social scientists often focus on psychological altruism in efforts to measure the extent to which altruistic acts are genuine. In human beings, and perhaps in some other highly developed species, the possibility of motivational or psychological altruism seems significant. Human altruism can be minimalist and for the most part emotionally uninvolved, such as the everyday respect for others that is expressed in etiquette and adherence to the minimal principle of nonmaleficence. It can become fuller or more idealistic in its expression, as in the case of efforts to actively assist others in genuine need.

What is at the very core of human altruistic love, which is altruism grounded in deep empathic affirmation? Phenomenologist Jules Toner defined love as “affirmative affection” (1968). We all know what it feels like to be valued in this way, and we remember loving persons who conveyed this affective affirmation through tone of voice, facial expression, a hand on the shoulder in time of grief, and a desire to be with us. This affection is affirming, which is to say that the agent of affirmation sees preciousness in us as we are. Altruistic love, which is uniquely human however much certain of its building blocks might be found in nonhuman species, is an intentional affirmation of our very being grounded in biologically given emotional capacities that are elevated by world view (including principles, symbol, and myth) and imitation into the sphere of consistency and abiding loyalty. As such, altruistic love is the most complex and interesting expression of human altruism.

Altruistic love is very closely linked to care (*cura*), which is love in response to the other in need. Care is the form love takes when it is attentive to the other in need. Love implies benevolence, care, compassion, and action. Compassion, for example, is the form love takes in response to suffering; it is a readiness to enter into the other’s suffering. Justice is the concern of love and care because no individual agent can attend to the needs of all and eventually must give attention to the underlying social and economic inequalities that give rise to need. But the essential core of love exists prior to these expressions of love. In the words of Irving Singer, “To love another person is to treat him with great regard, to confer a new and personal value upon him” (1985, p. 11). Yet those who claim to be persons of altruistic love and who fail to act on behalf of others insofar as they are able manifest an obvious contradiction, for love implies action.

Altruistic love is a human ideal across most, if not all, religions. One can argue about the extent to which religions succeed or fail in teaching and implementing a love idealism that goes beyond insular “tribal” interests. This is in part a matter for empirical investigation, although at face value it appears that esteemed religious altruists have accomplished much for all humanity. Desmond Tutu and Abraham Joshua Heschel, prophets of justice; Gandhi and Martin Luther King Jr., nonviolent liberators; Bonhoeffer, opponent of Hitler; Dorothy Day, feeder of hungry laborers; Eugene Rivers, bringer of peace to inner cities; Dame Cicely Saunders, giver of new hope to the dying; Jean Vanier, founder of l’Arche; Dag Hammarskjöld, seeker of peace in the Congo; and the Dalai Lama of Tibetan Bud-

dhism—these are only a few examples of modern leaders whose deep religious convictions about love and justice shaped their contributions to public life, political change, and human progress.

There is no doubt that numerous entirely secular individuals have achieved high levels of other-regard, and some authors in this book (for example, Monroe) do not see religion as a crucial element in the formation of altruistic personalities. Yet history is also filled with the accomplishments of those who perceived a relationship with a Supreme Being or Divine Love and who interpreted their altruism as a direct consequence of this perception. These figures are spiritual cosmopolitans absorbed in other-regarding service. We are sometimes too exclusively interested in what they do rather than in why they do it, because, as Martin Luther King Jr. often commented, the “what” depends on the “why.” They do just “happen” to be deeply spiritual persons. The more we can understand about them, scientifically and otherwise, the better. We do not suggest that altruistic love in its embrace of all humanity absolutely requires a spiritual and religious foundation; yet we cannot ignore the narrative of human history and experience indicating that this foundation is quite often present, even in modernity. For some, this may suggest that altruistic love is a cosmic force of some type in which human beings are able to participate, for example, through “grace.” We are interested in knowing more about “the inspired apostles of love . . . the great moral teachers of humanity . . . founders of all genuine religions . . . the true sages, seers, and prophets of practically all countries, cultures, and periods” (Sorokin, 1954, p. 485).

Where, however, does science come into view with regard to the study of human altruism and to its most intensive and extensive expression in love for humanity? How does the scientific study of altruistic actions in nonhuman species pertain to human altruism, if at all? How can we better understand human altruistic motivations and actions with a focus on all that these involve evolutionarily, genetically, developmentally, neurologically, emotionally, and conceptually? It is in the context of the dialogue between science, philosophy, and spiritual traditions that this book addresses various views of the roles of altruism and egoism.

It must be stated from the outset that there has been a theoretical trend in social scientific research, as well as in evolutionary biology, that generates cause for honest doubt about the very possibility of human altruism in any form. There may be some possible element of credibility enhancement in this hermeneutics of doubt. As Robert H. Frank writes, “The flint-eyed researcher fears no greater humiliation than to have called some action altruistic, only to have a more sophisticated colleague later demonstrate that it was self-serving” (1988, p. 21). Yet there is a body of coherent theory and data that casts some doubt on altruistic motivations, and this demands serious and balanced attention. The authors in this book take disparate positions, and we have not attempted to create any false impressions of unity.

The hermeneutics of doubt is grounded in the difficulty of proving, once and for all, the existence of psychologically altruistic motivations; such doubt is sometimes a matter of implication, as is the case when all living creatures are understood as conduits for the transmission of “selfish genes.” (There are also those, known in moral philosophy as ethical egoists, who believe that even if genuine

human altruistic motivations and acts are possible, they should be forbidden in order to encourage others to stand on their own two feet). Yet the attribution of “selfishness” to genes—a questionable importation of moral language into the domain of DNA—is compatible with authentic altruism on the part of the individual. As Matt Ridley puts it, “Selfish genes sometimes use selfless individuals to achieve their ends” (1996, p. 20).

One might ask whether the burden of proof should be placed on those who believe that human nature contains altruistic capacities alongside egoistic ones or on those who believe that genuinely altruistic capacities do not exist. Our intent in this book is to grapple honestly with current scientific questions about the existence of genuine altruism and to explore the nature of human other-regarding motives and acts.

This book brings together the work of leading biologists, social scientists, philosophers, and religion scholars. Although many philosophers and theologians have waxed eloquent about other-regarding motivations and acts, often with impressive grounding in the history of ideas, intellectual and practical progress now rests in the dialogue with scientific knowledge. On the other hand, many scientists have furthered our understanding of the phenomenon of altruism without nuanced interpretation of the human experience. Our purpose in publishing this volume is to enhance the dialogue on altruism across fields and disciplines in order to make progress in understanding and perhaps eventually in practice. We take to heart Irving Singer’s recommendation in the conclusions to his monumental study of the history of the idea of love; that is, “the most promising opportunity for us in the twentieth century is to be found in a synthesis of scientific and humanistic approaches to human affect” (1987, p. 345). Singer adds that although evolutionary biologists define altruism in a way that “does not necessarily imply sentiments of benevolence or love” (1987, p. 358), their work does provide a heuristic key into altruism and altruistic love. Similarly, Singer finds value in a closer examination of how altruism and altruistic love are related to developmental stages in the human. Our book is consistent with Singer’s signposts for the future, although we have added the study of the neurology of emotion to the mix.

On the Matter of Altruistic Extensivity

The neo-Darwinian emergence in evolutionary biology of kin altruism and reciprocal altruism must be viewed favorably as an antidote to the old social Darwinian paradigm of the ruthless survival of the fittest. The selfish-gene paradigm allows for a great deal of altruism toward those recognized as carrying a significant portion of one’s own genes, and the tit-for-tat model of reciprocal altruism allows for the emergence of a great many social sentiments in human and some nonhuman animals—including cooperation, sympathy, kindness, and trust. Yet the scope of altruism is limited by a set of underlying reproductive and self-interested priorities. A perennial ethical and spiritual challenge in human society is how to broaden the scope of altruism to the extent that all human beings, simply as human existents, are equally regarded (Outka, 1972).

One of the twentieth century's pioneer social scientists of altruism was Pitrim A. Sorokin, whose work, *The Ways and Power of Love* (1954), represents the initial attempt to contextualize concepts of human and divine love within the scientific literature on altruism and egoism. During the 1940s and 1950s, Sorokin directed the Harvard Research Center in Creative Altruism, which attracted the attention of eminent scholars and practitioners of altruism and altruistic love, both secular and religious, from around the world. He studied the creative energy of altruistic love and the techniques of altruistic transformation embedded in cultural and religious rights of passage. Sorokin first used the term *altruistic love* to describe this particular form of human altruistic behavior, and we therefore are compelled to give him some attention here. Sorokin also coined the term *extensivity* in the field of altruistic studies and related it to "intensity" of altruism or "love."

Sorokin developed an ideal typology of altruism based on large-scale interview and questionnaire studies. He described a set of persons whose love is "very intense toward a small in-group (their family, their friends, their clique or sect or faction), but whose love for anybody beyond this little universe is nonexistent. The extensity of their love is thus very low" (1954, p. 19). Sorokin collected 1,000 cases of "American Good Neighbors" and analyzed them along the axes of extensity, intensity, duration, motivational purity, and overt action. Among his many useful conclusions, Sorokin defined the key problematic in altruism as "the tragedy of tribal altruism," or the blunt fact that "in-group altruism" inevitably means "out-group egoism" (p. 459). It was in fact Sorokin who coined the now commonly used term *in-group altruism*.

Sorokin stressed the need to move from tribal egoism to universal altruism and considered in-group altruism, no matter how genuine, as the most serious of continuing human problems. He described the social dynamics of the common in-group persecution of those persons who are able to achieve degrees of extensity beyond group insularity, and he attempted to study those persons in depth in order to better understand the roots of their exemplary lives. His words are impassioned: "An exclusive tribal solidarity—known also as tribal patriotism, tribal loyalty, and tribal altruism—has mercilessly set man against man, and group against group. It has killed more human beings and destroyed more cities and villages than all the epidemics, hurricanes, storms, floods, earthquakes, and volcanic eruptions taken together. It has brought upon mankind more suffering than any other catastrophe" (1954, p. 461).

In other words, as contemporary evolutionary psychologists argue, the human propensity to aid and cooperate with others evolved in the context of intergroup conflict, and thus the complement of prosocial capacities is the tendency to form exclusionary alliances. The aggressive aspects of human nature had already been highlighted by presociobiological ethologists such as Konrad Lorenz. The most significant moral, scientific, and religious challenge that we face as a species is the overcoming of intergroup conflict. Sorokin suggested that human beings should turn their hate and aggression against disease, starvation, poverty, and other assaults on human well-being. If the inevitable correlate of altruism is aggression, then is the capacity for empathy potent enough to overcome the in-group/out-group barrier and to inhibit aggressive tendencies because of

the distress that the empathic observer feels in response to the noxious consequences of aggression (Feshbach & Feshbach, 1986)? Can the symbols that live in us and in which we live bring us to full equality regarding altruism? Is empathy so thoroughly the product of in-group evolution that in-group insularity is ineradicable? The question here is not just whether empathy can be extended, that is, if it is heavily or entirely constrained by in-group evolutionary origins. The further question is how empathy, which by its very nature is ethically neutral, can be used as a tool for nurture rather than for exploitative control. Even within groups, as Joseph Conrad vividly depicts in his novels and as Hitler understood to his political advantage, the ability to sense what others are feeling can be used either to provide sensitive care or to manipulate, even cruelly dominate. Evolutionary debates over the very origin and significance of empathy—not just its group domain but its biological function—are crucial, well grounded, and not yet resolved.

The theme of in-group altruism and out-group aggression is central to another significant work on the science of altruism, Elliott Sober's and David Sloan Wilson's *Unto Others: The Evolution and Psychology of Unselfish Behavior* (1998). This work is provocative in its critique of both social scientific and evolutionary biological assumptions that too easily disparage the human capacity for altruism, with regard to both motivation and action. Yet this work is highly debated by many of the primary contemporary architects of evolutionary theory. Sober and Wilson argue that unless the human capacity for altruism is in fact a reality by scientific analysis, there is a certain futility in recommending it, for otherwise in the final assessment one is only toying with an exalted illusion.

The dominant theory of evolutionary biology looks at acts that increase the fitness (survival and reproduction) of others at expense to the individual. In this biological sense, altruism does not imply any conscious benevolence. The conclusion offered by many sociobiologists is that such altruistic acts are all a matter of either inclusive fitness, such that one's own genes are passed on through relatives who also possess said genes, or of reciprocal altruism, from which the self derives compensatory reproductive benefits. Robert Trivers, George Williams, and John Maynard Smith are all associated with this perspective. Sober and Wilson, however, take a different view. They argue for a form of group altruism in which acts on behalf of other members of the group go beyond kin interests and reciprocity to the group as a whole. The survival of a group against competitors would make forms of altruism beyond individual reproductive and survival interests evolutionarily plausible. Sober and Wilson thus argue for "genuinely altruistic traits" (p. 6). They conclude that "Altruism can be removed from the endangered species list in both biology and the social sciences" (p. 337). The reader should keep this debate and its consequences in mind throughout this book and will find that our authors take different views as they wrestle with it. Group selection remains controversial but has significant defenders beyond Sober and Wilson (Boehm, 1999).

If group selection theory is valid, the problem of aggressive acts against or indifference to outsiders becomes even more significant. As Sober and Wilson put it,

In any event, it is worth saying here that our goal in this book is not to paint a rosy picture of universal benevolence. Group selection does provide a setting in which help-

ing behavior directed at members of one's own group can evolve; however, it equally provides a context in which hurting individuals in other groups can be selectively advantageous. Group selection favors within-group niceness *and* between-group nastiness. (p. 9)

This is important because, if altruism evolves within groups that compete with other groups, aggressive behavior toward outsiders, which is fully substantiated historically and anthropologically, may be difficult to overcome.

We leave further details of the scientific discussion to the specific sections of our book, although the reader by now has a sense of our broad terrain. Intellectuals have long taken sides in the egoism-altruism debate, although in different language games. In the Age of Enlightenment, philosophers contrasted benevolence and sympathy with love of self; theologians contrasted disinterested beneficence and charity with selfishness. Medieval thinkers contrasted *amor concupiscentiae* (self-love) with *castus amor* ("pure" love). Contemporary Protestant thought contrasts *eros* with *agape*, whereas psychologists speak of narcissism and its alternatives. Sociologist Auguste Comte (1798–1857) coined the terms *altruism* and *egoism*, by which he meant an unselfish desire to live for others in contrast to the impulse to benefit and gratify the self. The extensive study of altruism and egoism in the sciences has yet to adequately inform reflection in the humanities, which have generally tended to ignore scientific data in discussing such phenomena as compassion, kindness, and love. Similarly, the sciences have much to benefit from the insights of the humanities into aspects of the phenomena that are possibly beyond the reach of science and yet crucial to the overall human and cultural context.

Our task is to better understand the emergence of the capacities for altruistic actions and for empathic affect and how these contributed to a larger capacity for love. Setting aside all dualism and reductionism, our task is to better understand this transposition and genesis from above, within, and below. We have no interest in Platonic or Cartesian views of the self that bifurcate body and mind/soul, thereby relegating the scientific study of the neurological, biological, and evolutionary features of human altruism and altruistic love to irrelevancy.

Organization of the Book

We begin this book with a section on the definition of altruism and altruistic love as these are considered in the sciences and humanities. Part of the difficulty in engaging in significant dialogue across fields is that different disciplines mean different things by the same word. All fields of study can make greater progress in knowledge by conversation across boundaries, although this requires as much conceptual clarity as possible. Some attention in this section is also given to the phenomenology of human altruism and altruistic love. We begin with attention to human experience, bracketing out all heuristic filters, scientific and otherwise, that might be reductionistic.

The second section takes up the social scientific research into the question of the reality and nature of altruism and altruistic love, considered both quantitatively

and qualitatively. Social scientists have addressed altruism in a variety of contexts. Those most relevant to this book address underlying motivations rather than merely helping actions. This section highlights the disagreement in social science over the authenticity of other-regarding motivation. The social scientific approaches represented here examine rescue behavior that may place the agent at risk. But there is also investigation of daily acts of love, care, and kindness, which value the other and seem motivated primarily by desire for the good of the other.

The third section of this book takes up the powerful debates within evolutionary biology and evolutionary psychology regarding altruism. To what extent are all forms of altruism determined and limited in scope by selfish genes, as some evolutionary biologists would suggest? Is group altruism scientifically valid? If so, what are the implications of this for human altruism? Can human altruism ever extend beyond in-groups to out-groups and even to the religious and philosophical moral ideal of all humanity being equally considered? What scientific explanations are there for genuinely counterproductive altruistic behavior in various nonhuman species, and what do these explanations imply for human beings? Are human acts of radical self-sacrifice for the good of others contrary or consistent with nature and human nature as understood by evolutionary paradigms? How do data from the biological sciences compare with those from the social sciences regarding the existence of altruistic love at various levels of inclusivity?

The fourth section considers the emotional aspects of altruistic love, focusing on the role of empathy in both humans and nonhumans. A substance dualist view of the moral agent has historically focused entirely on properties of disembodied mind or soul, dismissing the biology and evolution of empathy and other affective states that contribute to love. How can empathy be better understood in its biological context? How does variation in neurobiological capacity for empathy affect cognitive functioning and the ability to act in a loving manner? Why are some persons more or less empathic than others? What biological processes are at work that encourage or inhibit the expression of altruistic love? How strong an influence on behavior is empathy, how much is it used for loving purposes, and is it a necessary feature of love in humans? What are the biological processes of attachment and bonding that pertain to empathy? What are the discontinuities between human and nonhuman capacities that allow for the emergence of a regard for the other as other? Do humans really have the capacity to practice altruistic love, and how do culture and symbol influence this capacity? What are the implications of altruistic love for a larger view of nature, human nature, and the universe?

The fifth section takes up altruistic love in its religious contexts but with an eye toward emerging areas for future study. Attentive to the best science and to religion pluralistically considered, this volume attempts to establish a framework for dialogue that will hopefully shape the discussion of altruistic love for some time to come.

In summary, these are some of the central questions that are at least raised, if not fully resolved, in this volume. We hope the volume can serve as a springboard for research into the following questions:

- What are the evolutionary origins and neurological substrates for altruistic behavior?
- What developmental processes foster or hinder altruistic attitudes and behavior in various stages of life from early childhood onward?
- To what extent do human individuals and societies manifest behavior that is motivationally or consequentially altruistic? What psychological, social, and cultural factors influence altruism and caring?
- Do spiritual and religious experiences, beliefs, and practices influence altruism?
- How does the giving and receiving of altruistic love interact with personal well-being and health?
- How can researchers from various disciplines collaborate to enhance this field of study?
- Overall, is it possible to gain new insights that can be utilized to help people and their communities to better appreciate the significance and importance of love and to benefit from its expression as a lived reality?

Finally, let us make a concluding comment about the science-religion dialogue over evolution and altruism. It is interesting to note that all the axes of theoretical tension described in the introduction—freedom versus determinism, individual versus group functionalism, material versus ideational agency, and the essential selfishness versus the loving competence of human nature—were philosophical and theological controversies centuries before evolutionary theory entered the fray. Although it is tempting to ascribe disagreement between scientific and theological perspectives on these issues to disciplinary conflict, it may reflect more the inherent and abiding ambiguities of human nature itself. The various disciplines are engaging common poles of human experience, though with different methods and metaphors . . . the very thing that may provoke mutual refinement.

REFERENCES

- Boehm, C. (1999). *Hierarchy in the forest: The evolution of egalitarian behavior*. Cambridge: Harvard University Press.
- Feshbach, S. F., & Feshbach, N. D. (1986). Aggression and altruism: A personality perspective. In C. Zahn-Waxler, E. M. Cummings, & R. Iannotti (Eds.), *Altruism and aggression: Biological and social origins* (pp. 189–217). Cambridge: Cambridge University Press.
- Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: W. W. Norton.
- Outka, G. (1972). *Agape: An ethical analysis*. New Haven, CT: Yale University Press.
- Ridley, M. (1996). *The origins of virtue: Human instincts and the evolution of cooperation*. London: Penguin.
- Singer, I. (1984). *The nature of love. Vol. 1: Plato to Luther*. 2d ed. Chicago: University of Chicago Press.

- Singer, I. (1987). *The nature of love. Vol. 3: The Modern World*. Chicago: University of Chicago Press.
- Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge: Harvard University Press.
- Sorokin, P. A. (1954). *The ways and power of love: Types, factors, and techniques of moral transformation*. Boston: Beacon Press.
- Toner, J. (1968). *The experience of love*. Washington: Corpus Books.

Introduction to Part I

STEPHEN G. POST

Academic disciplines tend to erect walls of separation, developing independent language games and definitions of terms, and sometimes even contending for supremacy in the world of thought. But progress in altruism and altruistic love is too important to humanity to be slowed by separatism. Nevertheless, disciplines must have their unique methodologies and heuristic keys. So what is needed is a new field of altruistic studies, one in which the walls of separation between disciplines are replaced by low, neighborly fences that encourage conversation, respect, and collaboration. As Robert Frost wrote, “Something there is that doesn’t love a wall.”

A certain acrimony has surrounded the study of altruism, often prompted by the arrogant assumption that a single discipline could provide the one and only hermeneutic. Altruism and altruistic love, however, are sufficiently complex so that, above all, humility is needed. It is unlikely that either of these phenomena will ever be fully understood by the life sciences, the social sciences, or the humanities, even if they are in creative conversation rather than walled off from one another.

In this section, the authors have been asked to do their best to clarify what they mean by altruism and altruistic love. Elliott Sober, a philosopher of science whose focus is on evolutionary biology, provides “The ABCs of Altruism,” defining altruistic behavior as enhancing “the fitness of someone else (the ‘recipient’) at some cost in fitness to the donor.” The honeybee, for example, to benefit the group,

disembowels itself when stinging an intruder. A mindless organism can, in this sense, be an altruist. Sober then turns to definitions of psychological altruism, psychological egoism, and psychological hedonism, a subset of egoism. Here he moves to the human domain of motive and intentionality. His position is pluralistic in the sense that he recognizes that whatever egoistic inclinations humans have, there are altruistic ones as well. Sober realizes that this is difficult to prove once and for all, because it is always possible to assert that even the most altruistic act is driven by a desire to avoid guilt or to feel good about oneself. Sober ends his chapter with the suspicion that the popularity of a purely egoistic image of the human self is shaped by a culture of individualism and competition.

Edith Wyschogrod, an eminent phenomenologist and moral philosopher, provides a broad definition of human altruism as “an action favoring other individuals at the expense of the altruist.” All moral experience, she argues, begins in discovering the other “as an ethical datum that makes a claim upon the self to engage in other-regarding acts.” She devotes attention to the problem of genetic reductionism, that is, interpreting human behavior in purely genetic terms. These are terms that she brilliantly associates with the Pythagorean tradition. In her discourse on discovering the other as other (“alterity”), she directly challenges the idea that ethical life is possible from the vantage point of self-interest, no matter how enlightened to long-term perspectives. Wyschogrod’s “Pythagorean Bodies and the Body of Altruism” provides a highly original and rigorous phenomenological analysis of altruism. The reader benefits not only from her ideas but also from the exercise of reading an important genre of philosophy.

Jerome Kagan, a leading psychologist of human development, defines altruism as the “helping agent’s awareness of the need of another and the intention to be of assistance.” Intentionality, rather than actual outcomes, is central to this definition. Kagan, in his “Morality, Altruism, and Love,” is particularly critical of evolutionary biology, which he believes provides ideological support for a culture of narcissism. He is also highly critical of those who extrapolate from nonhuman animal models to human behavior. The human being is, Kagan asserts, utterly unique, emergent from evolution with the gift of a moral sense that includes a propensity to care for the other, although this propensity can be culturally undermined to degrees.

I have reserved my own chapter in this section, “Altruism and Altruistic Love: The Tradition of Agape,” for an attempt at bringing together the different disciplinary languages of altruism in a manner that takes all the relevant disciplines seriously and points toward the meaning of the phenomenon of altruistic love as the epitome of altruistic intent and behavior in humans. As is the case in Wyschogrod’s chapter, this chapter also engages the dialogue between the sciences, philosophy, and religion.

Across all these chapters, the authors share a common conclusion that altruistic motivations and actions are humanly possible. Although it remains true that altruistic motivations are hard to prove absolutely, on which side of the altruism-egoism debate should the burden of proof be placed?

1

The ABCs of Altruism

ELLIOTT SOBER

In this chapter, I provide a conceptual map of some of the main questions that have been posed about altruism. What biologists mean by the term *altruism* is not at all the same as what psychologists and ordinary people mean by the term. After explaining the difference between evolutionary and psychological altruism and how the latter is related to the concept of love, I focus on the evolutionary concept and describe how it is possible for the competitive process of natural selection to lead to the evolution of altruistic traits. Then I turn to the psychological concept and describe how it is related to, though different from, the concept of morality.

Evolutionary Altruism

A trait is said to be evolutionarily altruistic by virtue of the effects it has on *fitness*. An organism's fitness is its ability to be reproductively successful; an organism's survival is relevant to its fitness only to the extent that survival promotes reproductive success. An altruistic behavior is one that enhances the fitness of someone else (the "recipient") at some cost in fitness to the donor. Thus a mindless organism can be an evolutionary altruist. Darwin thought that the barbed stinger of the honeybee is an altruistic trait—the bee disembowels itself when it stings an intruder to the nest; the stinger keeps pumping venom even after the bee has perished, thus conferring a benefit on the group.

The puzzle about evolutionary altruism is that it appears to be a trait that natural selection will stamp out, not promote. If altruists and selfish individuals live in the same group, altruists will donate fitness benefits to others, whereas selfish individuals will not. Altruists may receive benefits from the donations of other altruists, but so too may selfish individuals be on the receiving end. Altruists thus create a “public good” (a term from economics) at a cost to themselves; this good benefits both altruists and selfish individuals alike. It follows that altruists will be less fit than selfish individuals in the same group. Natural selection is a process that causes fitter traits to increase in frequency and less fit traits to decline. How, then, can natural selection explain the existence of evolutionary altruism?

Darwin’s answer was the hypothesis of *group selection*. Although altruists are less fit than selfish individuals in the same group, groups of altruists will be fitter than groups of selfish individuals. Hives that contain bees with stingers will be more successful than hives whose bees lack stingers. Altruistic traits evolve because they benefit the group and in spite of the fact that they are deleterious for the individuals that have them.

Darwin was a “pluralist” about natural selection; he held that some traits evolve because they are good for the individual, whereas others evolve because they are good for the group. This pluralism became a standard part of the evolutionary biology practiced during the decades in which the Modern Synthesis was created (1930–1960). The idea of group adaptation was often applied uncritically during this period, but the same can be said of the idea of individual adaptation. This situation changed in the 1960s, when group selection was vigorously attacked. Its exile from the subject was hailed as one of twentieth-century biology’s major advances. Since then, the hypothesis of group selection has been making a comeback; many biologists now think that group selection is theoretically well grounded and empirically well supported as the explanation of some (but by no means all) traits, whereas many others continue to reject it. The dust has not yet settled, but my impression is that *multilevel selection theory*—the idea that selection processes can and do occur at all levels of the biological hierarchy—is on its way to becoming a standard feature of evolutionary thinking (Sober and Wilson, 1998).

Psychological Altruism

In contrast with the concept of evolutionary altruism, psychological altruism is a property that applies only to individuals who have minds. A honeybee may be an evolutionary altruist, but it presumably is not a psychological altruist. And if, out of the goodness of my heart, I encourage you not to have children, my behavior may be an instance of psychological altruism, but it is not evolutionarily altruistic (assuming that my action does not enhance your fitness).

Although the concept of psychological altruism gets applied to people and to actions, the best place to begin is to think of psychological altruism as a property of motives or desires or preferences. Eve’s desire that Adam have the apple is *other-directed*; the proposition that she wants to come true (*that Adam has the apple*) mentions another person but not Eve herself. In contrast, Adam’s desire that he

have the apple is *self-directed*. In addition to purely self-directed and purely other-directed desires, there are *mixed* desires, wherein people desire that they and specific others be related in a certain way; had Eve wanted to share the apple with Adam, her desire would have been mixed.

An altruistic desire is an other-directed desire in which what one wants is that another person do well. This may involve your wanting the other person to have what he or she wants, or it may involve wanting the other person to have what you think would be good for him or her.

Well-wishing can be nonpaternalistic or paternalistic. Altruistic desires, understood in this way, obviously exist. The controversy about psychological altruism concerns whether these desires are *ultimate* or merely *instrumental*. When we wish others well, do we ever have this as an end in itself, or do we care about others only because we think that how they do will affect our own welfare? The theory known as *psychological egoism* maintains that all ultimate motives are self-directed. When Eve wants Adam to have the apple, she has this other-directed desire only because she thinks that his having the apple will provide her with a benefit.

Psychological hedonism is one variety of egoistic theory. Hedonism claims that the only ultimate motives that people have are the attainment of pleasure and the avoidance of pain. The only things we care about as ends in themselves are states of our own consciousness. This special form of egoism is the hardest one to refute. It is easy enough to see from human behavior that people don't always try to maximize their access to consumer goods. However, even when someone chooses a job with a lower salary over a job that pays more, the hedonist can say that this choice was motivated by the desire to feel good and to avoid feeling bad. Indeed, hedonists think they can explain the most harrowing acts of self-sacrifice—for example, the soldier in a foxhole who throws himself on a live grenade to save the lives of his comrades. The soldier supposedly does this because he'd sooner not exist at all than live with the knowledge that he had allowed his friends to perish. This hedonistic explanation may sound strained; why not say instead that the soldier cared more about his friends than he did about his own survival? But that the explanation sounds strained does not mean that it must be false.

Because hedonism is difficult to refute, egoism is also difficult to refute. However, that does not mean it is true. Human behavior also is consistent with the view called *motivational pluralism*; this is the claim that people have both self-directed and other-directed ultimate aims.

This theory does not assert that there are human actions that are driven by purely other-directed ultimate desires. Perhaps one consideration that lurks behind everything we do is a concern for self. What pluralism asserts is that some of our ultimate desires are other-directed. Because actions may be caused by several desires acting at once, pluralism is best understood as a claim about the character of our desires, not about the purity of our actions.

It is an interesting fact about our culture that so many people are certain that egoism is true and that so many others are certain that it is false. A Martian anthropologist might find this surprising, in view of the fact that the behavior we observe in everyday life seems to be consistent with both egoism and pluralism. Our convictions evidently have outrun the evidence we have at hand. Is the popularity

of egoism due to the fact that we live in a culture that emphasizes individuality and economic competition? Is the popularity of pluralism due to the fact that people find it comforting to think that benevolence is an irreducible part of human nature? These questions are as fascinating as they are difficult to answer.

Altruistic Love

Love is an emotion, so it, like psychological altruism, belongs to the realm of the mental. The honeybee is not a psychological altruist, nor can it be said to sacrifice its life out of love for its nestmates. Love comes in many varieties; it can be sexual or platonic, and it can be self-centered or altruistic. Altruistic love, of course, entails an altruistic motive or desire or preference on the part of the lover. However, the converse does not hold; the existence of an altruistic desire does not entail the presence of love. When I read about the disasters that happen to strangers far away, I find myself wanting them to be better off. I wish them well, but I do not find in my heart the emotion of love. In fact, I don't detect in myself any emotion at all as I casually peruse the morning newspaper. Perhaps this introspective impression is not correct, but it seems perfectly possible to have an unemotional altruistic desire.

What, then, is the extra ingredient that turns an altruistic desire into altruistic love? This is an instance of a larger question. What is an emotion, and how does it differ from a mere desire? When I nearly have an automobile accident while driving my car, I experience fear. I believe that I am in danger, and I want to be safe. But fear is something more than this belief and this desire. It is a feeling. But what *is* a feeling, and what distinguishes one feeling from another? This is a difficult question; indeed, it isn't even clear what the appropriate vocabulary is in which one should attempt to construct an answer. So let us focus, more modestly, on psychological altruism. Altruistic motivation is part, if not the whole, of altruistic love. We know well enough what psychological altruism is and how it differs from evolutionary altruism.

The Prisoners' Dilemma: Rational Deliberation

Game theory was first invented in mathematical economics (Von Neumann & Morgenstern, 1944). Only later was it brought within evolutionary theory (Maynard Smith, 1982). Perhaps the most famous game analyzed in game theory is the *prisoners' dilemma*. After explaining how this game works when it is formulated as a problem about individual decision making, I show how the problem was reformulated in evolutionary game theory. The solution that is correct in the one context differs from the solution that makes sense in the other (Skyrms, 1994). Understanding this is the key to seeing what it takes for altruism to evolve.

Suppose that the police arrest you and your accomplice in crime. They take you into separate rooms and interrogate you separately. If both of you remain silent, they will have very little evidence against you, and so you can anticipate being put in jail for only a short period of time, say 3 months. However, the police

tell you that if you provide them with full information and your accomplice remains silent, they'll let you go free (0 months in jail). They also say that if your accomplice tells all, but you remain silent, you'll go to jail for 9 months. And if both of you spill the beans, you'll each go to jail for 6 months. What should you do? You have two choices (confess or remain silent), and there are four possible situations you might be in, depending on what you and your accomplice do:

		<i>The other person</i>	
		<i>remains silent</i>	<i>confesses</i>
<i>You</i>	<i>remain silent</i>	-3 months	-9 months
	<i>confess</i>	0 months	-6 months

I use negative numbers to represent the payoffs you would receive in different circumstances because I want bigger numbers to represent better outcomes—you would rather have 0 months in jail than 3, and you would rather have 6 months in jail than 9.

Given these payoffs, you decide to confess—this action is said to “dominate” the other, meaning that it is the better action, no matter what the other person does. Your accomplice is in exactly the same position, so he decides to do the same thing. Thus, when both players are rational, both end up in jail for 6 months. If both had behaved irrationally and chosen to remain silent, each would have been better off, because then each would have gone to jail for only 3 months. In this case, universal rationality leads to a worse outcome than universal irrationality. In many situations in real life, we think that being rational helps us obtain better outcomes. The prisoners’ dilemma seems paradoxical because it shows how this connection between rationality and doing well is not inevitable.

After the prisoners’ dilemma was invented at the RAND Corporation in 1950,¹ it was used to model a number of real-world social dilemmas. One of them was the question of whether the United States should continue to stockpile nuclear weapons in its cold war arms race with the Soviet Union. Because the model builders were giving advice to the U.S. government, I represent the payoffs that the United States would receive in four circumstances:

		<i>USSR</i>	
		<i>disarms</i>	<i>arms</i>
<i>U.S.</i>	<i>disarms</i>	Second Best	Fourth Best
	<i>arms</i>	Best	Third Best

The best outcome for the United States would be for it to arm itself and for the USSR not to; although armaments cost money, this cost would be more than compensated for by the power the United States would obtain. The worst outcome for the United States would be for the United States to disarm and for the USSR not to, because then the USSR would have power over the United States. The second and third best outcomes are ones in which the United States and the

USSR do the same thing. It would be better for the United States if both sides disarmed than if neither did; this would save money and also reduce the chance of a devastating nuclear war. Given these payoffs, the United States should arm itself; the USSR, faced with the same problem, should do the same thing. It would have been better for both if both had disarmed, but this is not what rational deliberation dictates.

The prisoners' dilemma also had an impact on the budding environmentalist movement. Garrett Hardin (1968), in his influential paper "The Tragedy of the Commons," describes a hypothetical community of farmers who use a shared parcel of land ("the commons") for grazing. Each farmer can either put the maximum number of animals out to graze on this commons or put some lesser number there. Each farmer gets a higher income by using the commons to the utmost; however, when all the farmers do this, they overexploit the commons and ruin it. For Hardin, the prisoners' dilemma distills the essence of how free-market capitalism can destroy the environment and make everyone worse off.

With these three examples in mind, let's extract the abstract structure of the prisoners' dilemma game. Remaining silent in the first game is an instance of "cooperating" with your accomplice. Disarming is an instance of "cooperating" with the Soviet Union in the second game. And restraining your use of the commons is a case of "cooperating" with the other farmers in the third. Similarly, confessing to the police is an instance of "defecting" against your accomplice in the first game, arming is a case of "defecting" against the Soviet Union in the second, and using the commons to the maximum is an instance of "defecting" in the third. The terms *cooperate* and *defect* became standard in the game theory literature. They mean exactly the same thing as *altruism* and *selfishness* do in evolutionary game theory. The abstract structure of the payoffs you receive in a prisoners' dilemma, then, is as follows:

		<i>The other person behaves</i>	
		<i>altruistically</i>	<i>selfishly</i>
<i>You behave</i>	<i>altruistically</i>	$x - c + b$	$x - c$
	<i>selfishly</i>	$x + b$	x

When you and the other person behave selfishly, no benefits are donated, and none are received; call the payoff that you receive in this circumstance " x ." If you behave altruistically and the other person is selfish, you donate a benefit that costs you c to produce, and you receive nothing in return. So your payoff is $(x - c)$. If you behave selfishly and the other person behaves altruistically, you receive a benefit, but you don't pay the cost of producing a donation, so your payoff is $(x + b)$. Finally, if you and the other person both behave altruistically, each of you pays the cost of donation, but each of you also receives a donation from the other person; in this case your payoff is $(x - c + b)$. As mentioned earlier, a dominance argument entails that you should be selfish. The same argument leads the other person to choose the same action. If both players behave selfishly, both do worse than they would have if both had behaved altruistically; notice that $x - c + b > x$, if $b > c$.

Payoffs are supposed to reflect your “utilities.” What are these? As the examples illustrate, they could be avoiding jail, attaining power and safety, taking in dollars, or anything, as long as the actors prefer more of the quantity over less. Well, *almost* anything. Suppose you are in the original prisoners’ dilemma situation and that you care about what happens both to you *and* to your accomplice; you care about these equally. Then your utilities will be as follows:

		<i>The other person</i>	
		<i>remains silent</i>	<i>confesses</i>
<i>You</i>	<i>remain silent</i>	$-3-3 = -6$ months	$-9-0 = -9$ months
	<i>confess</i>	$0-9 = -9$ months	$-6-6 = -12$ months

Now a dominance argument tells you to remain silent. If your partner is similarly motivated, both of you will choose to remain silent, and so you will each go to jail for only 3 months. This is *not* a prisoners’ dilemma in the sense in which that term that has become canonical in game theory, because rational deliberation does not lead both parties to a solution that is worse than the one they would have obtained if both had been irrational. Yet the payoffs illustrated describe the situation that the prisoners would face if their utilities were as described.

This tells us something about what “utility” means in the prisoners’ dilemma game. The assumption is that a person’s utilities do not reflect any concern for what happens to the other player. It is not that game theory cannot handle utilities that reflect self-interest and benevolence simultaneously; after all, the previous table represents people who care about self and other in equal measure. Rather, my point is that this notion of utility is banished from formulations of the prisoners’ dilemma.

In fact, the limitation on what utility means in the prisoners’ dilemma is even more severe. Suppose you are a psychological egoist of the hedonistic variety. All you care about ultimately is experiencing pleasure and avoiding pain. You are being interrogated by the police, and you recognize that your pleasure and pain will be affected in two ways by what transpires. First, going to jail will cause you pain. But suppose you will find it equally painful to know that your accomplice is in jail (if he serves n months in jail, this will cause you as much pain as would be delivered by your spending n months in jail yourself). In this case, just as in the case of genuine benevolence, you will choose to remain silent. If your utilities are as described, then you are not in a prisoners’ dilemma, properly so called. Utility in the prisoners’ dilemma has to be “crassly egoistic,” not just egoistic.

The Prisoners’ Dilemma: Evolutionary Game Theory

Although rationality leads to universal selfishness when the prisoners’ dilemma is formulated as a problem about rational deliberation, it turns out that natural selection does not inevitably lead to universal selfishness when the game is transposed to an evolutionary context. Instead of having two players choose how they

will behave, we imagine that there are two types of organism, selfish and altruistic. Here selfishness and altruism are defined in terms of their fitness effects, not psychologically. A large number of organisms divide into pairs, and the individuals in a pair interact with each other. The payoffs they receive are in the currency of reproductive success; those who receive higher payoffs have more babies. Individuals are assumed to reproduce asexually; it also is assumed that offspring exactly resemble their parents. The next generation therefore may display a mix of altruists and selfish individuals different from the mix found in the previous generation. The members of this new generation then form up into pairs and play the game again. They produce the third generation, and so on.

When this process takes place over many generations, what ultimately happens to the frequencies of altruism and selfishness? That depends on how pairs are formed. If individuals pair at random, then selfish individuals will be fitter than altruists, and so altruism will decline in frequency and go extinct. However, if like always pairs with like, then altruists interact exclusively with each other and thus obtain a payoff of $(x - c + b)$, whereas selfish individuals interact exclusively with each other and thus obtain a payoff of x . If $b > c$, altruists will be fitter than selfish individuals, and so altruism will increase in frequency. I have just described two extreme cases—in which individuals associate randomly and in which like always pairs with like; in between these two extremes are different degrees of positive association. The crucial factor for altruism to evolve is that there be some tendency for like to associate with like—how strong this tendency needs to be if altruism is to evolve depends on the values of the costs and benefits involved.²

Cultural and Genetic Evolution

Although biologists modeling the evolution of altruism usually assume that different phenotypes correspond to different genes, this assumption is not needed in an evolutionary model. If parents transmit their traits to their offspring by teaching, altruism can evolve by cultural evolution. Evolution by natural selection requires a mechanism of inheritance, but the core idea here is just that offspring resemble their parents. The idea that learning and culture also can provide a mechanism of inheritance is worth bearing in mind when one considers Darwin's hypothesis about how human morality evolved:

It must not be forgotten that although a high standard of morality gives but a slight or no advantage to each individual man and his children over the other men of the same tribe, yet that an increase in the number of well-endowed men and advancement in the standard of morality will certainly give an immense advantage to one tribe over another. There can be no doubt that a tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to aid one another, and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection. At all times throughout the world tribes have supplanted other tribes; and as morality is one important element in their success, the standard of morality and the number of well-endowed men will thus everywhere tend to rise and increase. (Darwin, 1871, p. 166)

As Darwin notes, human history is filled with cases of group competition. To think of this history within an evolutionary framework, one need not assume that these cultural differences between groups are due to the fact that people in different cultures have different genes. Even if cultural groups were genetically identical, they still could differ in their phenotypes and still could faithfully transmit those phenotypes across the generations. Even if the evolution of altruism in social insects turns out to be an exclusively genetic process, the evolution of altruism within human beings needs to be understood from the dual perspective of genetic and cultural change.

In fact, there are two ways in which cultural evolution can depart from the usual pattern of genetic evolution. To see what these are, consider the following three types of selection process (Sober, 1993):

	<i>Mechanisms of heredity</i>	<i>Definitions of fitness</i>
<i>I</i>	genes	having babies
<i>II</i>	learning	having babies
<i>III</i>	learning	having imitators

In Type I processes, genes provide the mechanism of inheritance, and fitness is measured in terms of biological reproduction. The process of cultural group selection (Boyd & Richerson, 1985) that I just described is of Type II. Learned traits, no less than traits that are genetically transmitted, can affect the survivorship of individual people and how reproductively successful they are. In Type III processes, ideas are transmitted by learning, and they succeed and fail for reasons that have nothing to do with their impact on the reproductive success of the individuals who accept them. Here we are recognizing the possibility to which Dawkins (1976) gave the name “meme.” An example of this type of process is provided by the drastic reduction in family size that took place at the end of the nineteenth century in Europe. The “demographic transition,” as it is called, apparently involved a trait with lower biological fitness replacing a trait with higher biological fitness. Having smaller family size was an attractive idea in spite of its Darwinian disutility. The trait had high cultural fitness, and a Type III process led to its spread. In saying this, I am not explaining *why* reduction in family size was suddenly an idea whose time had come. Rather, I am simply noting that the process spread by a Type III process. It is important to recognize that moralities can change for the same reason. In saying that moralities evolve by a selection process, we are not limiting ourselves to models of the first two types.

Altruism and Morality

Altruism, whether evolutionary or psychological, frequently strikes people as a good thing. Because of this, altruism’s dark side needs to be held clearly in view if we are to understand the moral dimensions of altruism and also its evolutionary and psychological character.

The process of group selection does not eliminate competition from the evolutionary process but merely transposes it up one level. In group selection, groups compete with other groups, just as in individual selection, individuals in a group compete with each other. In *The Origin*, Darwin (1859) says that he uses the term *struggle for existence* to encompass two different types of situations. Two dogs may fight with each other over a bone, but two plants may each struggle against the drought. Individual selection will favor the stronger dog in the first case, but it also will favor the stronger plant in the second. Natural selection does not have to involve individuals actively interfering with each other. In the case of the dogs, one of them gets the bone only if the other does not; but in the case of the plants, how well one plant stands up to the sun does not affect how well the other does.³ Applied to the case of group selection, Darwin's point means that group selection can be a process in which groups actively interfere with each other, but it also can be a process in which groups respond with varying degrees of success to a common environmental problem.

What does this tell us about the evolution of altruism? An individual who is an evolutionary altruist benefits the group at cost to the self. Altruists might donate food to members of their own group, or they might burn the crops of other groups. Either way, they help their own group to do better in the struggle with other groups. Group selection can promote within-group niceness, but it also can promote between-group nastiness. It is an obvious fact about nature that the process of individual selection has left plenty of room for individuals to be nasty to each other. We should expect no less of group selection—it can lead groups to be nasty to each other. Group selection doesn't always lead the lion to lay down with the lamb; it can lead lions to cooperate with each other to bring down lambs.

There is a similar dark side to psychological altruism. How can a noninstrumental concern for the welfare of others be morally bad? The easiest way to see this is to realize that being nice to someone can involve being nasty to third parties. If Alan cheats Beth at cards because he wants to give the money to Carol, we may decide that his dishonesty was altruistically motivated *and* morally wrong. A macabre illustration of this point may be found in the training that Nazi concentration camp guards and physicians received. They were told that the revulsion they might experience in carrying out the Nazi genocide was a cost they should be willing to pay for the good of the German people (Lifton, 1986). Altruistic motivation can underwrite evil (Sober and Wilson, 1998).

There is another sort of separation that we need to recognize between altruism and morality. Altruistic ultimate desires, if they exist, involve a concern that this or that individual do well. Perhaps a parent wants her children to do well for reasons that go beyond her belief that this will redound to her benefit. Indeed, this desire may be a cognitive state that some of our primate relatives occupy. The point I am emphasizing is that altruistic concern for specific others is not the same as the acceptance of a general moral principle. If I want my children to prosper, it is a separate question from whether I think that all children should be nurtured by their parents. Moral principles are general and impersonal in what they say. They do not mention self or specific others but describe what anyone in a generally characterized situation is permitted or obliged to do. Even if altruistic desires ex-

isted before modern humans evolved, I suspect that abstract moral principles are a uniquely human achievement.

Most of us recognize this distinction between the specific desires we have and the general moral principles we accept when we compare catastrophes that happen to those who are near and dear with similar catastrophes that happen far away to people we do not know. Our moral principles tell us that the two catastrophes are equally bad if they caused the same amount of suffering and devastation and death. However, most of us care about the one more strongly than we care about the other. Even if we are not wholly indifferent to the distant event, our desire that the victims receive help is weaker in this case than it is in the one that happens closer to home. Our desires—including our altruistic desires—sometimes reflect differences that our moral principles tell us are irrelevant.

This division between altruistic desires and moral principles recurs when we consider the concept of love. We do sometimes speak of generalized love—love for humanity or love of all living things—but love is perhaps most familiar when it is focused on particular individuals. Most of us have loving relationships with only a small circle of individuals; and even within that circle, our love of some individuals is more intense than our love of others. We recognize the possibility that this circle might be expanded and that the love we feel for those within the circle might be made more intense. But what is far more difficult to contemplate is the idea that we might simultaneously greatly expand the circle *and* love all the individuals in it equally. A perspectival painting can be enlarged to encompass more of the landscape, but it still remains perspectival. Perhaps love, like altruistic desire, obeys a different logic from that of our ethical principles.

If general ethical principles and specific altruistic desires are distinct in the way I have suggested, why do human beings have them both? Or, to give this question a temporal dimension, if specific altruistic desires were part of the human phenotype before general ethical principles came along, why did the latter become part of that phenotype? One possible explanation is the social function that moral principles serve. Moral principles are devices for encouraging group-beneficial desires and behaviors. Human beings evolved in a social environment; the desires we have are influenced by that social environment. The morality espoused and enforced by our elders is a powerful influence on the type of person we grow up to be. Moralities will evolve by group selection when they influence the behaviors of individuals and vary with respect to how much they promote the group-beneficial character of those behaviors. As noted before, it doesn't matter whether moralities are transmitted genetically or culturally for this process to go forward.

The idea just described is that human moralities have fitness consequences, just like the barbed stinger of the honeybee. This can be true even if moralities are transmitted by learning and barbed stingers by genes. However, as the previous discussion of three types of selection process suggests, we need to remember that moralities can change for purely cultural reasons that have nothing to do either with genetic transmission or with the effect of moral beliefs on having babies. The morality accepted within a group (or the mix of moralities that are accepted) can affect the survivorship and reproductive success of individuals, but ideas have a way of jumping from head to head for reasons that are orthogonal with the goal of

having babies. A morality has biological effects, but it also has cultural meanings that can lead it to wax or wane. The phenomenon of human morality needs to be analyzed from both angles.

NOTES

1. The abstract structure of the prisoners' dilemma game was first described by Merrill Flood and Melvin Dresher; Albert Tucker (also at RAND) invented the prisoners' story to illustrate the game and thus gave the game its name (Poundstone, 1992).

2. This is true in the "one-shot" prisoners' dilemma, in which players interact with each other only once. However, the question is more subtle for the "iterated" prisoners' dilemma, in which players in a pair interact repeatedly, each time producing an altruistic or a selfish behavior. As Axelrod's (1984) simulations show, cooperative strategies (such as tit-for-tat) can evolve even when pairs are formed at random. However, when this is true, there still is a correlation between the *behaviors* produced by individuals, even though the *strategies* used by the two players are uncorrelated. For an example of what this means, see the discussion in Sober (1993, pp. 115–117).

3. The dogs are playing a "zero-sum game," whereas the plants are not (Sober, 1984, p. 17).

REFERENCES

- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Boyd, R., & Richerson, P. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Darwin, C. (1859). *On the origin of species*. London: John Murray.
- Darwin, C. (1871). *The descent of man and selection in connection with sex*. London: John Murray.
- Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162, 1243–1248.
- Lifton, R. (1986). *The Nazi doctors: Medical killing and the psychology of genocide*. New York: Basic Books.
- Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.
- Poundstone, W. (1992). *Prisoner's dilemma: John von Neumann, game theory, and the puzzle about the bomb*. New York: Doubleday.
- Skyrms, B. (1994). Darwin meets the logic of decision: Correlation in evolutionary game theory. *Philosophy of Science*, 61, 503–528.
- Sober, E. (1984). *The nature of selection: Evolutionary theory in philosophical focus*. Cambridge: MIT Press.
- Sober, E. (1993). *Philosophy of biology*. Boulder, CO: Westview Press.
- Sober, E., & Wilson, D. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge: Harvard University Press.
- Von Neumann, J., & Morgenstern, O. (1944). *The theory of games in economic behavior*. New York: Wiley.