MICHAEL TOMASELLO

A NATURAL HISTORY OF HUMAN MORALITY

Printed in the United States of America

First printing

Library of Congress Cataloging-in-Publication Data

Tomasello, Michael.
A natural history of human morality / Michael Tomasello.
pages cm
Includes bibliographical references and index.
ISBN 978-0-674-08864-1 (hardcover : alk. paper) 1. Ethics—History. 2. Ethics, Evolutionary. 3. Philosophical anthropology. I. Title.
BJ1298.T66 2015
170.9—dc23
2015010960

Contents

Preface

A Natural History of Human Morality is a companion volume to my 2014 book *A Natural History of Human Thinking*. The parallel titles are appropriate because in both volumes I propose the same two-step sequence in the evolution of human social life: first, new forms of collaborative activity, and then new forms of cultural organization. In the first volume I attempted to spell out the species-unique kinds of thinking that emerged from these new forms of social life. In the current volume I attempt to explicate how these new forms of social life structured the way that early humans came to engage in moral acts that either subordinated or treated as equal their own interests and the interests of others, even feeling a sense of obligation to do so. This moral attitude or stance did not—and still does not—win out consistently in individuals' actual decision making, of course, but it does make those decisions, whatever their outcome, moral decisions.

I have been gathering my thoughts for this volume over the past five years or so, beginning with a seminar on the evolution of human cooperation held in the fall of 2009 here at the Max Planck Institute for Evolutionary Anthropology, and continuing with a similar seminar on the evolution of human morality held in the winter of 2012–2013. Many interesting and fruitful discussions in those seminars have shaped my thinking on these issues significantly, and I thank all of those who participated. I also had a number of very useful discussions during this same time period with Sebastian Rödl, who helped me with some difficult philosophical concepts.

In addition, a number of people read and provided very useful commentary on earlier versions of the manuscript. In particular, one or another draft version was read by Ivan Cabrera, Robert Hepach, Patricia Kanngiesser, Christian Kietzmann, Berislav Marusic, Cathal O'Madagain, and Marco Schmidt. I thank them all for their extremely helpful comments and suggestions. I would like to single out for special thanks Neil Roughley and Jan Engelmann, who engaged with me and the manuscript especially deeply and on multiple occasions. For certain, the manuscript is much more coherent for all of their insights. I also thank Andrew Kinney, Richard Joyce, and an anonymous reviewer from Harvard University Press for their comments on the manuscript as well.

Finally, as with the first volume, my deepest gratitude goes to Rita Svetlova, with whom I have discussed extensively all of the most important ideas in this volume—and others that she helped me to let go of—to the great benefit of the final product. I dedicate this book to her and our children.

A Natural History of Human Morality

their especially cooperative social arrangements. Given these assumptions, our attempt in this book is (1) to specify in as much detail as possible, based mainly on experimental research, how the cooperation of humans differs from that of their nearest primate relatives; and (2) to construct a plausible evolutionary scenario for how such uniquely human cooperation gave rise to human morality.

The starting point is nonhuman primates, especially humans' nearest living relatives, the great apes. As in all social species, great ape individuals living in the same social group depend on one another for survival—they are interdependent (Roberts, 2005)—and so it makes sense for them to help and care for one another. Moreover, as in many primate species, great ape individuals form long-term prosocial relationships with specific other individuals in their group. In some cases these relationships are with kin, but in other cases they are with unrelated groupmates, or "friends" (Seyfarth and Cheney, 2012). Individuals depend on these special relationships to enhance their fitness, and so they invest in them, for example, by preferentially grooming their friends or supporting them in fights. The evolutionary starting point for our natural history of human morality, therefore, is the prosocial behavior that great apes in general show for those with whom they are interdependent, namely, kin and friends.

Tomasello et al. (2012) provide an account of the evolution of uniquely human cooperation that focuses on how, from this great ape starting point, early human individuals became ever more interdependent with one another for cooperative support. The interdependence hypothesis—whose basic framework we adopt here—is that this took place in two key steps, both of which involved new ecological circumstances that forced early humans into new modes of social interaction and organization: first collaboration and then culture. The individuals who did best in these new social circumstances were those who recognized their interdependencies with others and acted accordingly, a kind of cooperative rationality. Although the individuals of many animal species are interdependent in various ways, early humans' interdependencies thus rested on a new and unique set of proximate psychological mechanisms. These new and unique mechanisms enabled individuals to create with others a plural-agent "we," as in what "we" must do to capture a prey or how "we" should defend our group from other groups. The central claim of the current account is that the skills and motivation to construct with others an interdependent, plural-agent "we"—that is, the skills and motivation to participate with others in acts of *shared intentionality* (Bratman, 1992, 2014; Gilbert, 1990, 2014)—are what propelled the human species from strategic cooperation to genuine morality.

The first key step occurred hundreds of thousands of years ago, as a change in ecology forced early humans to forage together with a partner or else starve. This new form of interdependence meant that early humans now extended their sense of sympathy beyond kin and friends to collaborative partners. To coordinate their collaborative activities cognitively, early humans evolved skills and motivations of *joint intentionality,* enabling them to form together with a partner a joint goal and to know things together with a partner in their personal common ground (Tomasello, 2014). On the individual level, each partner had her own role to play in a particular collaborative activity (e.g., hunting antelopes), and over time there developed a common-ground understanding of the ideal way that each role had to be played for joint success. These common-ground role ideals may be thought of as the original socially shared normative standards. These ideal standards were impartial in the sense that that they specified what either partner, whichever of us that might be, must do in the role. Recognizing the impartiality of role standards meant recognizing that self and other were of equivalent status and importance in the collaborative enterprise.

In the context of partner choice, in which all individuals had bargaining leverage, this recognition of self–other equivalence led to a mutual respect among partners. And since it was vital for partners to exclude free riders, there also arose the sense that only collaborative partners (and not free riders) were deserving of the spoils. The combined result was that partners came to consider one another with mutual respect, as equally deserving second-personal agents (see Darwall, 2006). This meant that they had the standing to form with one another a joint commitment to collaborate (see Gilbert, 2003). The content of a joint commitment was that each partner would live up to his role ideal and, further, that both partners had the legitimate authority to call the other to task for less than ideal performance. Early humans' sense of mutual respect and fairness with partners thus derived mainly from a new kind of cooperative rationality in which it made sense to recognize one's dependence on a collaborative partner, to the point of relinquishing at least some control of one's actions to the self-regulating "we" created by a joint commitment. This "we" was a moral force because both partners considered it legitimate, based on the fact that they had created it themselves specifically for purposes of self-regulation, and the fact that both saw their partner as genuinely deserving of their cooperation. Collaborative partners thus felt responsible to one another to strive for joint success, and to shirk this responsibility was, in effect, to renounce one's cooperative identity.

In this way, participation in joint intentional activities—engendering both the recognition of partners as equally deserving second-personal agents and the cooperative rationality of subordinating "me" to "we" in a joint commitment—created an evolutionarily novel form of moral psychology. This novel form of moral psychology was not based on the strategic avoidance of punishment or reputational attacks from "they" but, rather, on a genuine attempt to behave virtuously in accordance with our "we." And so was born a normatively constituted social order in which cooperatively rational agents focused not just on how individuals do act, or on how I want them to act, but, rather, on how they *ought to* act if they are to be one of "us." In the end, the result of all of these new ways of relating to a partner in joint intentional activities added up for early humans to a kind of *natural, second-personal morality.*

The second evolutionary step in this hypothesized natural history—beginning with the emergence of *Homo sapiens sapiens* some 150,000 years ago—was prompted by changing demographics. As modern human groups started becoming larger, they split into smaller bands that were still unified at the tribal level. A tribal-level group—call it a culture—competed with other such groups for resources, and so it operated as one big interdependent "we," such that all group members identified with their group and performed their division-of-labor roles aimed at group survival and welfare. Members of a cultural group thus felt special senses of sympathy and loyalty to their cultural compatriots, and they considered outsiders to be free riders or competitors and so not deserving of cultural benefits. To coordinate their group activities cognitively, and to provide a measure of social control motivationally, modern humans evolved new cognitive skills and motivations of *collective intentionality*—enabling the creation of cultural conventions, norms, and institutions (see Searle, 1995)—based on *cultural* common ground. Conventional cultural practices had role ideals that were fully "objective" in the sense that everyone knew in cultural common ground how anyone who would be one of "us" had to play those roles for collective success. They represented the right and wrong ways to do things.

Unlike early humans, modern humans did not get to create their largest and most important social commitments; they were born into them. Most important, individuals had to self-regulate their actions via the group's social norms, the breach of which evoked censure not only from affected persons but also from third parties. Deviance in a purely conventional practice signaled a weakness of one's sense of cultural identity, but violation of a moral norm—grounded in second-personal morality—signaled a moral breach (see Nichols, 2004). Moral norms were considered legitimate because the individual, first, identified with the culture and so assumed a kind of coauthorship of them and, second, felt that her equally deserving cultural compatriots deserved her cooperation. Members of cultural groups thus felt an obligation to both follow and enforce social norms as part of their moral identity: to remain who one was in the eyes of the moral community, and so in one's own eyes as well, one was obliged to identify with the right and wrong ways of doing things (see Korsgaard, 1996a). One could deviate from these norms and still maintain one's moral identity only by justifying the deviation to others, and so to oneself, in terms of the shared values of the moral community (see Scanlon, 1998).

In this way, participation in cultural life—engendering both the recognition that all in-group compatriots were equally deserving and a sense that the culture's collective commitments were created by "us" for "us"—created a second novel form of moral psychology. It was a kind of scaled-up version of early humans' second-personal morality in that the normative standards were fully "objective," the collective commitments were by and for all in the group, and the sense of obligation was group-mindedly rational in that it flowed from one's moral identity and the felt need to justify one's moral decisions to the moral community, including oneself. In the end, the result of all of these new ways of relating to one another in collectively structured cultural contexts added up for modern humans to a kind of *cultural and group-minded, "objective" morality.*

One outcome of this two-step evolutionary process beyond great apes—first to collaboration and then to culture—is that contemporary human beings are under the sway of at least three distinct moralities. The first is simply the cooperative proclivities of great apes in general, organized around a special sympathy for kin and friends: the first person I save from a burning shelter is my child or spouse, no deliberation needed. The second is a joint morality of collaboration in which I have specific responsibilities to specific individuals in specific circumstances: the next person I save is the firefighting partner with whom I am currently collaborating (and with whom I have a joint commitment) to extinguish the fire. The third is a more impersonal collective morality of cultural norms and institutions in which all members of the cultural group are equally valuable: I save from the calamity all other groupmates equally and impartially (or perhaps all other persons, if my moral community is humanity in general), with perhaps special attention to the most vulnerable among us (e.g., children). The coexistence of these different moralities—moral orientations or stances, if you will—is of course anything but peaceful. Conflicts among them are the source of many of the most perplexing moral dilemmas that humans face—should I steal the drug to save my friend? should I keep my promise if it means harm to unknown others?—that seemingly have no fully satisfactory solutions (Nagel, 1986). The bare fact of such unsolvable incompatibilities in the dictates of morality suggests a complex and not totally uniform natural history in which different cooperative challenges have been met in different ways at different times.

The possibility that humans operate with several different, sometimes incompatible, moralities—and that they are due, at least in part, to processes of natural selection—raises the specter, feared by many thoughtful persons from Darwin's time on, that evolutionary explanations may serve to undermine the whole idea of morality. But this need not be the case. The point is that the ultimate causation involved in evolutionary processes is independent of the actual decision making of individuals seeking to realize their personal goals and values. The textbook case is sex, whose evolutionary raison d'être is procreation but whose proximate motivation is most often other things. The fact that the early humans who were concerned for the welfare of others and who treated others fairly had the most offspring undermines nothing in my own personal moral decision making and identity. I am able to speak the English language only because of my evolutionary, cultural, and personal histories, but that does not determine precisely what I decide to say at any given moment. In all, we should simply marvel at the fact that behaving morally is somehow right for the human species, contributing to humans' unparalleled evolutionary success, as well as to each individual's own sense of personal moral identity.

And so with this apologia, let us tell a story, a natural history, of how human morality came to be, beginning with our great ape ancestors and

tion/
tion

FIGURE 2.1. The two dimensions of social life for complex organisms.

In a Darwinian framework competition requires no special explanation, of course, but cooperation does. Acting in ways that benefit others is a stable evolutionary strategy only under certain conditions. The first task in this chapter, therefore, is to examine how cooperation works in evolution in general, using as an organizing theme the principle of interdependence. We then use this theoretical framework to characterize the nature of cooperation in great ape societies in particular, with the goal of characterizing, as a starting point for our natural history of human morality, the cooperative interactions of the last common ancestor of humans and other great apes some 6 million years ago.

Foundations of Cooperation

Cooperation presents a variety of puzzles for the theory of evolution by means of natural selection. We need not solve them all here. For current purposes all we must do is to identify those evolutionarily stable patterns of cooperation that are relevant for our investigation of the human species. In identifying these patterns we will be especially concerned both with the proximate (psychological) mechanisms—the cognitive, social-motivational, and self-regulatory processes—that enable the individuals of complex social species to cooperate with one another, and with the adaptive conditions under which those psychological processes might have come to be favored by natural selection.

Evolutionarily Stable Patterns of Cooperation

Standard evolutionary theory dictates that cooperation can be maintained as an evolutionarily stable strategy only if it is not overly detrimental to the reproductive fitness of the individuals involved (altruism is often defined by evolutionary biologists—humorously but pointedly—as "that which cannot evolve"). But there are a number of classic interactive categories that describe ways in which individuals may temporarily suppress their own immediate self-interest to cooperate with others, without thereby sacrificing themselves and their progeny out of existence in the long run. Following the theory of multilevel selection, it is most useful to explicate three broad categories, distinguished by the level at which they operate: kin selection operates at the level of the gene; group selection operates at the level of the social group; and mutualism and reciprocity operate at the level of the individual organism. Each of these categories of cooperative behavior may potentially be realized via a wide variety of different proximate mechanisms in different species.

First, perhaps the most basic process in the evolution of cooperation is kin selection. Darwin wondered why social insects, such as ants and bees, so readily sacrifice for one another (to the point that there even exist sterile helpers). In the context of modern genetics, Haldane and Hamilton solved the problem by noting that in social insects individuals living in the same social group share more genes with one another than do groupmates in other animal species. By helping others, individual ants and bees are promoting copies of their own genes; they are, in a sense, helping themselves. Dawkins (1976) pushed this view to the extreme, considering all of evolution from this "gene's eye view."

The proximate mechanisms for kin selection are normally quite simple. One has to be predisposed to do things that help others (without necessarily understanding cognitively that one is doing so), and one must direct this behavior selectively to kin. This selectivity toward kin is most often accomplished via spatial proximity. For example, ants and bees just do things that help others in their immediate environs, and even more cognitively complex organisms, such as humans, most often identify as kin those with whom they have grown up in close physical proximity (Westermarck, 1891). This psychological simplicity means that kin selection was not a likely breeding ground for the many complex cognitive distinctions and judgments underlying human morality. However, it almost certainly was responsible for the basic prosocial emotion of sympathy, originating in the context of parent–offspring bonding and helping one's kin. As we shall see in the account of great ape cooperation, some species then had occasion to extend their sympathy beyond kin to "friends."

A second major process in the evolution of cooperation is, controversially, group selection. The theory of group selection takes not a gene's eye view on the process but, rather, a group's eye view, with some theorists even noting that a multicellular organism is simply a group of cooperating single-cell organisms (Wilson and Wilson, 2008). The basic idea is that if the social groups of a species are internally homogeneous genetically, and these social groups are at the same time well differentiated from one another genetically, then these groups may actually become units of natural selection themselves. Cooperation enters the story because one can imagine that social groups with more cooperators outcompete social groups with more noncooperators. Individual cooperators are thus at a disadvantage within their group relative to noncooperators (who enjoy the benefits but do not pay the costs), but their group flourishes and so they have an advantage over individuals from other groups of the same species. Most theorists agree that group selection is possible in principle but that, in fact, in most cases there is too much gene flow between groups (due to immigration) for group selection to be a powerful force in more than just a few isolated cases.

Once again, the proximate mechanisms for group selection are simple. Again, one must simply be predisposed to do things that help others (without necessarily understanding cognitively that one is doing so), and one must direct this behavior selectively to groupmates, again recognized in most cases by spatial proximity.[1] Whereas group selection of this type may not have played a crucial role in the evolution of human cooperation and morality, a variant called cultural group selection almost certainly did, albeit quite late in the process. Cultural group selection does not primarily concern genetic evolution but, rather, cultural evolution, as individuals in a group conform to one another's behavior via social learning (thus promoting homogeneity of behavior), and even immigrants conform as well (thus solving the immigration problem). A secondary phase of gene–culture coevolution may also ensue such that individuals best able to, for example, socially learn have an adaptive advantage. Cultural group selection will play a key role in the later stages of our account of the evolution of human morality, as groups that are able to foster and encourage cooperation among their constituent members—through such things as social norms and institutions—outcompete neighboring groups that are not as good at these things.

Third, and at the crux of the current account because of their potential influence on psychological mechanisms, are processes of mutualism and reciprocity. These processes both work at the level of the individual organism, and both operate evolutionarily by somehow "paying back" the individual for his cooperation, either in the moment or later.

From an evolutionary point of view, processes of mutualism are easily explained because all cooperating individuals immediately benefit (although there can still be problems of free riding in some cases). For this reason, very little attention is paid to mutualism in the theoretical literature on the evolution of cooperation, human or otherwise (much more is paid to altruism). But, in fact, mutualistic collaboration is responsible for many of the most distinctive characteristics of human cooperative cognition and sociality. It is responsible in the sense that, early in human evolution, the need for specific types of mutualistic collaboration created the adaptive conditions within which an especially intricate and complex set of proximate mechanisms evolved for regulating social coordination and communication: shared intentionality (Tomasello, 2009, 2014). These proximate (psychological) mechanisms have been accorded little research attention in this context, but they are absolutely crucial—or so we will argue—for understanding the evolution of human cooperation and morality.

As for reciprocity, the classic version is so-called reciprocal altruism (Trivers, 1971): I help or defer to you on one occasion, and you reciprocate by helping or deferring to me on the next occasion, so that we both benefit in the long run. But how does this work psychologically? Classic tit-for-tat reciprocity (sometimes characterized as "you scratch my back, and I'll scratch yours") is often thought of implicitly as a kind of social contract in which we agree ahead of time to obligate ourselves to a future course of action. Although no one would seriously propose a social contract for nonhuman animals, without a social contract it is difficult to understand how reciprocity might work. The first problem is that reciprocal altruism has no explanation for the initial act of altruism at all, which has to be, in this account, blind optimism or accident. The second problem is the powerful incentive to defect: once you have given me my benefit, I have no incentive to give you yours in return; I should simply quit while I am ahead. My only incentive is that perhaps it will cause you to give me an additional benefit in return. But why should it? You have the same incentive to defect that I have. Without some kind of agreement, reciprocity has no real rational or emotional power to motivate altruistic behavior on its own. Indirect reciprocity brings reputation into the picture, but in the end it is also plagued by the same two problems of motivating the first act and cheating.[2]

There is no doubt that behavioral patterns of reciprocity occur widely in nature. The issue is the proximate mechanisms underlying them. What is needed, for current purposes at least, is a more psychologically realistic account to replace the implicit contract view of reciprocity. A good start is the typology of de Waal (2000). Most important, he distinguishes what he calls calculated reciprocity from emotional (or attitudinal) reciprocity. Calculated reciprocity is the implicit contract: we each keep track of who has done what for whom and stop cooperating if we are giving more than we are getting. Not surprisingly, this type of reciprocity seems to be very rare in nature. More frequent, especially in mammals with their propensity for forming long-term emotion-based social relationships, is emotional reciprocity. With emotional reciprocity, individuals form emotional bonds with those who help them (perhaps based on the mechanism by which offspring bond to those who succor and protect them), and then they naturally help those with whom they are socially bonded—kin and "friends," as it were. Emotional reciprocity would seem to be widespread in primates and other mammals, at the very least, but it raises the questions of why individuals form friendly social relationships with nonkin in the first place, why they help those friends, and how their friendships affect their reproductive fitness.

Interdependence and Altruism

Virtually all formal theories of the evolution of cooperation (e.g., Nowak and Highfield, 2011) conceptualize the individual as an asocial monad in constant competition with all other members of its species in a struggle to pass along its genes. But this view, while in some sense valid, is seriously incomplete in the case of cognitively and socially complex organisms, not to mention that it has little concern for proximate mechanisms. The main point is that cognitively and socially complex organisms are enmeshed in many and varied social relationships and interdependencies with others, and this means—assuming that these relationships and interdependencies are important to their fitness—that helping or cooperating with those others, reciprocally or otherwise, is not a sacrifice but an investment.

In the stakeholder model, then, one could say that individuals do get "paid back" for their niceness, so one could call it something like "pseudo-reciprocity" (Bshary and Bergmueller, 2008). Fine. But the important point is that *unlike classical reciprocity, the altruist's behavior is not contingent on the recipient responding or being influenced by the help in any way (or on the altruist's anticipation of any such responding or influence).* The recipient will simply continue doing what he always does—alarm calling, mating, being a coalition or hunting partner, or being in the social group—because it is in his interest to do so, and this just happens to benefit the altruist, as a by-product, as it were. A more active way of thinking about the altruist's behavior, then, is as a kind of investment in the recipient; she invests in his well-being since that contributes to her well-being (Kummer, 1979). In this view, emotional reciprocity is most accurately characterized as mutual investments among interdependent friends, who help one another not in order to pay back past acts but in order to invest in the future. In some cases, each individual may be dependent on the other precisely because of the benefits he provides—for example, they share food with one another reciprocally—but from a proximate point of view the altruistic act is not motivated by any specific previous act, only by the goal of maintaining the relationship. We might thus conceptualize the situation as individuals living *symbiotically* (a concept typically applied only to interactions between species), as there is no exchange of favors or anything of the kind, only individuals going about attempting to increase their fitness directly.

This way of looking at things does not have the many problems of reciprocity—specifically, the problem of motivating the first altruistic act or the problem of defecting—because there is no direct contingency of altruistic acts (although, of course, over time a relationship may break down for many reasons). One individual helps another do what she would do in any case, for her own reasons—up to a mathematical point. But this account does, of course, still have a potential problem of free riding because one can lag on the helping: it would be best if someone else helped my alarm-calling groupmate so that I could get the benefits without paying any costs. But as Zahavi (2003) has pointed out, the exact same logic applies to kin selection: it is in my interest to help my sibling because he shares my genes, but my first preference is that someone else help him so that I do not have to bear the costs or risks. And so, of course, interdependence as described by the stakeholder model does not solve all of the problems of cooperation in one fell swoop; rather, using the same logic as kin selection, it changes the cost-benefit analysis significantly.

The interdependence perspective thus integrates mutualism and reciprocity in a natural way, and it motivates reciprocity in a much more stable way than classical accounts. It also places altruism in a new light. Altruism is not an improbable achievement against the individualizing forces of natural selection; rather, it is an integral part of the social lives of all beings that live with others interdependently—up to a (mathematical) point. Everyone helps and gets helped, up to a point, because everyone is important to someone in some way, up to a point. This view also accords well with the prescient views of Kropotkin (1902) on "mutual aid" as playing a crucial role of the everyday lives of social beings who must struggle more against the exigencies of the physical environment (sometimes cooperatively) than against one another.

Partner Control, Partner Choice, and Social Selection

Cooperators do best when they are surrounded by other cooperators. So once the individuals of a species have begun down the cooperative path, they may actively attempt to influence others around them in a cooperative direction. They may do this most directly through acts of so-called partner control—most often the punishment of noncooperators—which may be seen as the kind of opposite of investing positively in cooperators and friends through acts of helping. The problem with punishment is that it is costly, or at least risky, for the punisher, for example, if the individual being punished rebels. A safer alternative, if it is available, is so-called partner choice in which cooperators simply avoid interacting with cheaters. Although there may be simple ways for avoiding cheaters in some cases, in more socially complex organisms partner choice often requires sophisticated judgments, based on past experience, about which of several individuals will make the best partner. In all, partner control and partner choice represent important complements to the process of investing positively in those on whom one depends. In these cases, individuals actively attempt to influence those on whom they depend, or might depend, either by coercing bad partners into being good ones or by choosing their partners wisely.

Given processes of either partner control or partner choice, over time what emerges is what West-Eberhardt (1979) calls social selection. In Darwin's (1871) modified account of organic evolution, the process of natural selection is complemented by the process of sexual selection. Sexual selection is not a totally new evolutionary process; it is just that in this case the selection is being performed not by the physical environment (as in classical natural selection) but, rather, by the social environment. In sexual selection, individuals of the opposite sex choose potential mates based on characteristics indicating such things as health, strength, and fecundity (e.g., large size, bright coloration, or youth). Those characteristics are thus selectively favored for purposes of mating, and this increases the reproductive fitness of those possessing them.

Social selection is simply a generalization of this process. Individuals in the social group may favor other individuals for all kinds of reasons in addition to sexual attractiveness, and this may affect both the recipient's survival and his reproductive success. Thus, if the individuals of a social group do the most beneficial things for the best alarm callers, then the characteristics of good alarm callers—keen perception, fast reactions, loud calling—will be socially selected for in alarm callers. If individuals of a social group need grooming partners, then individuals who are enthusiastic and good groomers will be selected for, along with their special characteristics. In some interactions, one individual may have more "leverage" than others in the sense that he is a more important partner; for example, dominants might be expected to be more in demand as coalition partners than subordinates, so they can demand more of their potential partners than can subordinates. We may thus think of the kind of socially complex decision making that occurs in partner control and partner choice as a kind of "biological market" (Noe and Hammerstein, 1994).

Although in principle virtually any physical or behavioral characteristic may be subject to social selection, for current purposes cooperation is a special case. If we assume, for example, that in some species collaborating with others to obtain food brings mutualistic benefits to all involved, then one can imagine a biological market based on partner choice and control in which what is being socially selected is the characteristics of good cooperators, for example, tolerance for partners in feeding situations, skills at coordinating and communicating with partners, propensities for helping partners as needed, and tendencies for shunning or punishing free riders—and what is being selected against, of course, are the characteristics of cheaters and incompetents.

Summary

To prepare for our upcoming natural history of human morality, we have briefly considered all three levels of multilevel selection. Kin selection was undoubtedly critical in primate (even mammalian) cooperation, well before humans emerged, in building up the emotional substrate for protecting and caring for offspring, which was then in some cases co-opted for protecting and caring for friends. Cultural group selection—as a special instance of group selection without its attendant problems—very likely played an important role near the end of the process, as modern human cultural groups competed with one another for resources and the most cooperative (or moral) groups won out.

But in our hypothesized natural history the main action will be at the level of the individual organism. This is because, in the current account, acting morally means interacting with others cooperatively by means of and through certain psychological processes. What we have done so far is simply to reconceptualize the evolutionary processes of cooperation that occur at this individual level, most especially mutualism and reciprocity. We have proposed that what is most basic at this level are dependencies (symbioses) among individuals, which may produce mutualistic or reciprocal patterns of cooperation by any of many different proximate mechanisms (e.g., emotional reciprocity in mammals). They also motivate individuals to care for or invest in those on whom they depend (altruism) and to attempt to make their partners as cooperative as possible (partner choice and control). This reconceptualization of individual-level cooperative interactions and relationships lays important theoretical groundwork for explaining the evolutionary emergence of human morality.

Great Ape Cooperation

Given this theoretical framework in which cooperation is based mainly on the principle of interdependence, we may now begin our natural history of human morality proper. We do this by attempting to characterize, as best we can, the social lives of the last common ancestor of humans and other great apes, who lived somewhere in Africa approximately 6 million years ago. We use as contemporary models the social lives of humans' nearest living relatives, the great apes, especially those of humans' very nearest living relatives, chimpanzees and bonobos (although in actuality the vast majority of both field and experimental research is with chimpanzees). We look first at several aspects of their social interactions with conspecifics in the wild, and then at their behavior in experiments directly testing for senses of sympathy and fairness.

Sociality and Competition

Chimpanzees and bonobos live in highly complex social groups typically comprising several dozen or scores of individuals of both genders (so-called multimale, multifemale groups). Daily life is structured by a fission–fusion organization in which small parties of individuals forage together for some time, only to disband soon thereafter in favor of new parties. Males live their whole lives in the same group in the same territory; females emigrate to a neighboring group during early adolescence. During their development, individuals form with others various kinds of long-term social relationships. Most important is, of course, kinship, but also important are relationships with nonkin based on dominance and something like friendship. Much of the complexity of chimpanzee and bonobo social interaction results from the fact that they also recognize and react to these same social relationships as they occur among third parties in the group. Interactions between neighboring groups are almost totally hostile for chimpanzees, whereas for bonobos interactions with foreigners are more peaceful.

Both chimpanzees and bonobos compete with groupmates all day every day. This is not only in the indirect evolutionary sense of competing to pass along genes but also in the more immediate sense of competing face-to-face for food, mates, and other valuable resources. For example, in the case of food, the prototypical situation is that a handful of individuals travel until they find a fruiting tree. Each individual then scrambles up the tree on its own, procures some fruit on its own, and seeks maximum spacing from others to eat. Such scramble competition, in which the winner is the one who gets there first, is often complemented by contest competition, in which the winner is the one who wins the fight or dominance contest, for example, when a dominant in the tree takes what it wants while nearby subordinates defer. Both chimpanzee and bonobo males compete, and even fight, for access to females, though this competition is clearly more intense in chimpanzees. Interestingly, many agonistic encounters in both species are not about immediate access to a resource but, rather, about one individual simply asserting dominance over another—which will then translate into easier access to resources of various kinds in the future. Contests over dominance, therefore, act as a kind of proxy for contests over resources.[4]

Cognitively, chimpanzees and bonobos are built for competition. Thus, not only are they intentional, decision-making agents, who make instrumentally nonkin (Langergraber et al., 2007). This cooperating in order to compete requires individuals to monitor simultaneously two ongoing social relationships, but again, the coalition partners do not do anything special to coordinate their actions other than fighting side by side. And again as in many mammalian species, great ape combatants often actively reconcile with one another after fights, presumably in an attempt to repair the social relationship on which they both depend (de Waal, 1989a). In some cases, dominant males may intervene to break up interactions in which coalitions (which may later challenge him) are beginning to form (de Waal, 1982).

Coalitionary support is crucial in chimpanzee and bonobo dominance contests; it pays to have good and powerful friends. Therefore, individual chimpanzees and bonobos cultivate friends, quite often through reciprocal coalitionary support (de Waal and Luttrell, 1988). They also cultivate friends through other affiliative behaviors, such as grooming and food sharing. Thus, much evidence suggests that grooming in chimpanzees is preferentially directed to potential coalition partners (see Mueller and Mitani, 2005, for a review), and over time individuals that have been preferentially groomed by a partner preferentially groom that partner as well (Gomes et al., 2009). In addition, Mitani and Watts (2001) have shown that male chimpanzees preferentially "share" (i.e., for the most part, tolerate the taking of) meat and other food with their coalition partners. Add to this the finding of de Waal (1989b) that individuals that groom one another also share food with one another preferentially, and the result is a relatively tight set of reciprocal relations among the golden triad of grooming, food sharing, and coalitionary support.[5] Importantly, there is very little evidence that great apes engage in any reciprocity of favors with individuals other than long-term social partners. In the only relevant experimental study, Melis et al. (2008) found that randomly paired chimpanzees did not preferentially help an individual that had just helped them over one that had not. They concluded that, despite clear evidence of long-term reciprocities, "models of immediate reciprocation and detailed accounting of recent exchanges (e.g., tit for tat) may not play a large role in guiding the social decisions of chimpanzees" (p. 951).

Interrelations within the golden triad are often interpreted as instances of reciprocity, and of course, on a purely descriptive level, they are. But de Waal (2000), as noted above, identifies several different proximate mechanisms that might underlie these observed behavioral patterns. And what de Waal and