

## Where Did It All Come From?

### *Morality and the Evolution of the Mind*

And this tattooing, had been the work of a departed prophet and seer of his island, who, by those hieroglyphic marks, had written out on his body a complete theory of the heavens and the earth, and a mystical treatise on the art of attaining truth; so that Queequeg in his own proper person was a riddle to unfold; a wondrous work in one volume; but whose mysteries not even himself could read, though his own life heart beat against them; and these mysteries were therefore destined in the end to moulder away with the living parchment whereon they were inscribed, and so be unsolved to the last. And this thought it must have been which suggested to Ahab that wild exclamation of his, when one morning turning away from surveying poor Queequeg – “Oh, devilish tantalization of the gods!”

Herman Melville, *Moby Dick*

#### 7.1 MORALITY AND EVOLUTION

Human beings are part of the natural world and thus owe their makeup, including their cognitive capacities and limitations, to the same forces as any other organism. Therefore, it is natural, at least for the purposes of a scientific theory-building whose methods are not concerned with religious interpretations of human existence, to ask what the theory of evolution can tell us about the nature and content of human morality. As the law is a central element of human beings' normative world, the answer to this question potentially has important consequences for the understanding of the law as well.

Evolutionary thinking has permeated many disciplines. Theorems of evolutionary biology in particular have been put to use – not always felicitously – in many other contexts, from cultural development based on war<sup>1</sup> on to social theory.<sup>2</sup>

<sup>1</sup> Cf. Peter Turchin, *Ultrasociety: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth* (Chaplin, CT: Beresta Books, 2015).

<sup>2</sup> Cf. Luhmann, *Recht der Gesellschaft*; Luhmann, *Gesellschaft der Gesellschaft*.

The key question for the concerns of our inquiry is whether there are hardwired, genetically inherited mental structures that influence or even determine human moral evaluations, intentions and concrete actions. As these moral evaluations clearly impact upon both the making and (according to the most plausible theories of interpretation) the application of law, the influence of such hardwired mental structures on the law may be substantial.

There is a huge variety of research in this area. Some investigations emphasize the importance of the biological foundations of human morality, some the evolution of culture after modern human beings started populating the Earth some 80,000 years ago. Some research challenges traditional assumptions about the content of morality, explicitly including the idea of human rights or the idea of a universalist moral outlook. One example of the former, the mental gizmo thesis, was already discussed in the Chapter 6, although we did not go into detail with respect to the underlying evolutionary theory. From this point of view, evolutionary mechanisms constrain the moralities possible in a way that rules out any natural morality consisting of generous, solely other-regarding altruism or egalitarian justice, because such a morality could not form a behavioral evolutionarily stable strategy for an organism. By contrast, other approaches confirm notions of altruism, fairness and mutual respect not limited to small groups, albeit to varying degrees, in the framework of a naturalistic theory of morality.<sup>3</sup>

As far as the topic of human rights is concerned, there are two dimensions to the problem. The first concerns human rights themselves, whose evolutionary origin forms the object of some such theories, including the mental gizmo thesis. The second dimension concerns the normative principles that we have identified, which underpin intuitions about individual claims and the justification of human rights, including justice, obligatory concern for the well-being of others and respect for their worth. Our review of evolutionary theories needs to deal with both of these dimensions.

<sup>3</sup> Cf. Buchanan and Powell, *The Evolution of Moral Progress*, arguing for evolutionary fixed plasticity that produces a small-group morality under distress cues like out-group threats, but under favorable circumstances allows for the development of inclusivist moralities like human rights; Michael E. McCullough, *The Kindness of Strangers* (London: Oneworld, 2020), arguing that human beings' three key natural endowments are an instinct for reciprocity, the "appetite for helping others in hopes of appearing virtuous" and the instinct for reasoning. These instincts were activated in history to create care for strangers by: (1) god-kings during the Agricultural Revolution generating loyalty by appearing to care for the weak; (2) the Axial Age discovery of the Golden Rule; (3) the sixteenth century discovery of the prudential advantages of helping the weak – for instance, keeping unrest under control; (4) the Enlightenment convictions about equality, dignity and natural rights; (5) nineteenth-century humanitarianism; (6) the post-World War II humanitarianism; and (7) the contemporary drive towards efficient giving, *ibid.* 115 ff.; quote at 264.

## 7.2 PUZZLING ALTRUISM

Evolutionary biology has long been concerned not just with the physical traits of organisms, but also with behavior and the mental structures and inner states that underpin behavior.<sup>4</sup> Evolutionary explanations form a central part of ethology. Sets of behavioral patterns include social behavior, which occurs in many different forms. Some species, like bees and ants, live in highly organized, complex communities. How do such behavioral patterns and the underlying cognitive structures arise? What is their evolutionary origin?

In this context, behavioral traits that benefit organisms other than the actor are of particular interest. Altruistic behavior is a special object of scrutiny for some approaches to evolutionary theory, because it does not seem to fit well into what is regarded as the basic mechanisms of evolution, and into natural selection in particular. How can evolution select for traits that are beneficial not for the bearer of these traits, but for others?

These questions also arise for highly cooperative organisms such as human beings. Human morality is crucial to understanding this cooperative behavior. The mental world of humans includes prescriptive moral rules that affect both the formation of intentions to act and actual action. For humans at least, the evolutionary point of altruism is not the only problem to arise: Justice is another equally puzzling issue. Its principles limit the gains people can reap for themselves. How does this constraint fit into the picture of evolutionary theory? An equivalent question can be asked of the idea of obligatory respect for others. What are the evolutionary roots of this phenomenon?

Many approaches to the issue observe that the gains an organism reaps by cooperating are a central factor in the evolutionary process. It is argued that other-regarding behavior is only an evolutionarily stable strategy if it increases the chances of reproduction by augmenting the reproductive fitness of the cooperating organism. Moreover, every behavioral trait needs to yield such reproductive benefits if it is to continue to form part of the evolutionary process. Otherwise it will be weeded out by the forces of selection.

## 7.3 VARIOUS FORMS OF COOPERATION AND THE PROBLEM OF WHAT MORALITY IS

Before embarking on a review of the main findings of evolutionary theory that have a bearing on the theory of human rights, it is worth mentioning a few points that are sometimes overlooked in discussions of this matter and that will inform our analysis.

<sup>4</sup> Cf. for instance Konrad Lorenz, *Das sogenannte Böse* (Vienna: Dr G. Borotha-Schoeler, 1963); Edward Wilson, *Sociobiology: The New Synthesis* (Cambridge, MA: Harvard University Press, 1975).

First, sociality and cooperation are not necessary Forms of existence for organisms – solitary life forms can be very successful at assuring the reproduction of a species. There is no evolutionary necessity to be social or cooperative; it is simply one of many possible ways of living. Arguments claiming that humanity's specific form of life had to develop – a kind of natural teleology in the form of evolutionary theory – are therefore untenable.

Second, cooperation comes in many shapes and sizes. This is a simple but absolutely crucial observation. A herd of animals staying together to create protection in numbers; active group defense; temporary intragroup coalitions to outcompete other group members or to kill their offspring;<sup>5</sup> the complex functional differentiation of ant colonies or a beehive, including the infertility of some organisms for functional purposes; the mechanisms of cooperation in a human society of hunter-gatherers in 20,000 BCE; or human society in 2023, organized through the institutions of a constitutional state based on human rights, part of the global community under international law – these are all very different forms of cooperation.

Third, human forms of cooperation face the normative standard of legitimacy – bees have no other option than their hives, but how humans should build societies is far from clear. In human history and thought, very different forms of cooperation were used for long stretches of time: A stratified slaveholder society under an authoritarian ruler cooperates, but not in terms of justice, beneficence and mutual respect. In *Politeia*, Plato envisaged a highly cooperative society, where all members devoted their best abilities to the good of the polis – a vision that 2,500 years later was denounced as the seed of totalitarianism.<sup>6</sup>

Fourth, the diverse forms of cooperation presuppose very different cognitive abilities on the part of the organisms cooperating. The mechanisms used by antelopes to cooperate in herds, by bees to assemble in a hive, by ants to form a colony and by human beings to form a constitutional state within the UN system are not the same. These forms of cooperation consequently also come about in very different ways. Herds of antelopes or ant colonies are not created by centuries of reflection, reform and political revolution, while human social institutions are. These developments and the often-radical reorganization of human societies are of major interest, even though humans were living in hunter-gatherer societies, not constitutional states when the survival of the human species was decided, because they may offer clues to human beings' potential, including cognitive capacities that were put to use only in later stages of their history. The fact that human beings are able to form highly complex mathematical theories about the world with explanatory power is certainly of interest for understanding the cognitive machinery of human beings, even though differential equations were (most probably) only rarely discussed in the average cave 60,000 years ago.

<sup>5</sup> Cf. Tomasello, *A Natural History of Human Morality*, 23.

<sup>6</sup> Karl Popper, *The Open Society and Its Enemies*.

Fifth, the mechanisms underlying cooperation are not immediately obvious and are certainly manifold. This is true for human cooperation as well. The way that human societies are integrated into something like an organized structure is a classic and highly contested sociological question. Moral precepts are only one of the mechanisms that may be relevant here. Strategic interaction based on instrumental rationality, as in trade and other tit-for-tat relations, is another. Further elements include force or the threat of violence to command obedience, or the belief in an authority's legitimacy because of its divine origin or because of the leader's particular charisma. Force and belief in the legitimacy of some sort of authority are two very significant factors assuring cooperation in human history, used, for example, to build the pyramids, to construct the Palace of Versailles and to lead millions of Germans to become agents of bloodshed during World War II, often sacrificing their own lives on the way into the moral abyss.

Sixth, it is not obvious that human morality is only or even primarily a tool for cooperation. If one is to even start considering this problem, one needs to know what kind of cooperation is at issue, as we have seen. The effects of morality are not limited to what is functionally necessary for any form of cooperation. Morality can even become a lasting impediment to particular forms of cooperation: A sense of justice is a substantial obstacle to certain highly efficient forms of cooperation that would be useful if only reproductive fitness were at issue – say, an authoritarian command structure. Morality often demands the reorganization of society, thus challenging existing forms of cooperation, not on the terms of what is functionally needed to enhance reproductive fitness, but for the sake of the intrinsic value of justice. The ensuing conflicts have often torn human societies apart. Morality thus does not serve simply any form of cooperation; rather, it *defines legitimate forms of cooperation* – for instance, a just political order. In addition, it provides the tools to make such an order a reality.

The oft-encountered references in evolutionary theory simply to “cooperation” without any further qualifications are therefore inadequate. The explanatory task at stake is much more demanding: What is needed is an explanation of human sociality's very particular forms of cooperation and – as the foundation of the specific forms of human cooperation – of the emergence of the peculiar phenomenon of human morality in the natural history of cognition.

Seventh, care must be taken when determining the *explanandum*, which in our case is human moral precepts and their workings in the human mind. The determination of the *explanandum* is particularly important in the case of morality because what this *explanandum* is is far from clear. There is not only no consensus about how to explain an *explanandum* whose properties are generally accepted (say, apples move perpendicularly to the ground), but the *explanandum* itself is controversial (say, whether humans are psychological egoists or not). Moreover, one encounters another problem in this context: The preferred evolutionary explanatory theory sometimes fallaciously determines what is assumed to be the content of

morality. The following line of argument may serve as an example: Because it is assumed that the mechanisms of evolution can only produce psychological egoists, it is held that human beings must in fact be psychological egoists, because humans are ultimately the product of these evolutionary mechanisms and accordingly can only have the properties that these mechanisms can create.

It is important to avoid this fallacy, which ascertains that what cannot be explained from the point of view of some evolutionary theories cannot in fact exist. This would mean committing a fallacy analogous to the argument that gravitational force cannot exist because gravity cannot be explained within the framework of contact mechanics. The world is full of surprises, and we must remain open to the possibility that human moral cognition contains some of them. Therefore, a proper analysis of the phenomenon to be explained is the precondition of any attempt to formulate an explanatory theory of this phenomenon.

Eighth, it is important to be mindful of the fact that there are quite substantial controversies about the fundamental aspects of evolution – there is not one, let alone “The One” uncontested evolutionary theory, but a plurality of evolutionary approaches that merit serious consideration.

#### 7.4 ANIMAL MORALITY?

A first step towards getting a grip on these matters is to ask whether there is anything species-specific about morality. This is not an easy question to answer. However, there is a substantial body of work on animal cognition that can help to formulate some plausible theses.

The starting points of this research are the great apes, particularly chimpanzees and bonobos, which are human beings’ closest relatives. This does not mean that there is not considerable evolutionary distance between us and them – in fact, the evolutionary development of human beings branched off from that of the great apes some 7 million years ago. Human beings are thus separated from great apes by 14 million years of development.

Studies on great apes have produced evidence of differentiated cognitive abilities. To name just some of these abilities that are important for moral psychology:<sup>7</sup> There is reason to believe that great apes act with instrumental rationality, at least in some instances. Under certain circumstances, they comprehend intentional states such as the goals and desires of others, particularly in the context of competition. They are able to maintain social relationships over long periods among subordinates, dominants and friends within their group. They are able to identify such relations between familiar third parties. They have and express emotions and are emotionally empathetic. They communicate intentionally. Within this framework, they engage in prosocial behavior, including instrumental helping: “When costs are small, and

<sup>7</sup> Cf. on these findings Tomasello, *A Natural History of Human Morality*, 20 ff., 35.

food competition is absent, great apes help others.”<sup>8</sup> They are able to control some of their impulses for immediate self-gratification for prudential reasons, such as avoiding conflict. They are able to inhibit the impulse to forage individually and collaborate with others to obtain resources not available without such cooperation.

As far as morality proper is concerned, some have argued that great apes have some kind of proto-morality, others have denied this and others again have proposed a middle ground.<sup>9</sup> There seems to be evidence that great apes exhibit signs of what is called sympathy towards their kin and friends, including acts of instrumental helping.<sup>10</sup> But various experiments, including adapted forms of the ultimatum game, for instance, produced no evidence that great apes have a sense of fairness.<sup>11</sup>

What seems clear from all of this is that the human moral world is structured on different terms. Humans certainly entertain many warm feelings towards others, from friendly concern and sympathy to love, feelings that motivate them to act in ways beneficial to others. In addition, however, human moral cognition implies principles of altruism and justice that are prescriptive in nature, implying a moral ought, and that are the product of engaging in different forms of reflection on the object of evaluation, sometimes limited to a spontaneous intuition about the justness or moral wrongness of an act, sometimes the result of painstaking soul-searching. Not only acts, but also internal mental states such as intentions form possible objects of such evaluations. These intentions come in more than one shape – direct and oblique intentions, for example. Accordingly, not only does the evaluation of these intentions presuppose an intricate cognitive machinery, but so do these mental states themselves and observers’ ability to identify them.

Things become even more complex when we come to further key notions of human moral cognition, such as the concept of “ought” and moral responsibility. The moral ought cannot be conceptualized without a notion of freedom, because ought is not a determining “must,” but leaves the agent the freedom to act otherwise. It is related to the notion of responsibility, another central category of the moral world of human beings, including their psychology, which presupposes that human beings can understand what is right, that they can make this insight the motivation for action and that they therefore can be held accountable for said action. All of this obviously presupposes a particular cognitive make-up, and this is true whatever one’s

<sup>8</sup> Tomasello, *A Natural History of Human Morality*, 31.

<sup>9</sup> Cf. for instance Frans de Waal, *The Bonobo and the Atheist* (New York: W. W. Norton & Company, 2013), 239 f.: “Who says the bonobo can do whatever he wants? . . . Even if he lacks notions of right and wrong that transcend his personal situation, his values are not altogether different from those underlying human morality. He, too, strives to fit in, obeys social rules, empathizes with others, amends broken relationships, and objects to unfair arrangements. We may not wish to call it morality, but his behaviour isn’t free of prescriptions, either”; Joan Silk et al., “Chimpanzees Are Indifferent to the Welfare of Unrelated Group Members,” *Nature* 437, no. 7063 (2005), 1357; Tomasello, *A Natural History of Human Morality*, 36.

<sup>10</sup> Tomasello, *A Natural History of Human Morality*, 25.

<sup>11</sup> Tomasello, *A Natural History of Human Morality*, 32 ff., 36.

position may be on the question of free will, be it indeterminist, determinist or compatibilistic.

A further point: Morality does not equal sympathy, as illustrated by the moral obligations one has towards those people one does not sympathize with and the shame one may feel if one does not live up to these obligations, despite disliking the patient of the moral action. The same holds for empathy (if distinguished from sympathy). One may understand very well how it feels to be incarcerated and still think that it is not only the legally, but also the morally right thing to do in some cases of crime.

Finally, human moral judgment is connected to a set of particular sentiments that are the consequences of certain moral evaluations. These sentiments include shame, indignation and resentment.

The differences between the cognitive and thus moral makeup of human beings and great apes manifest themselves in the way they are treated in morally relevant contexts:

Despite the fact that they live in cooperative social groups and behave pro-socially towards kin and friends – and are of course worthy targets of our moral concern – chimpanzees themselves are not moral agents. We do not allow them to roam freely in our midst for fear that they will attack our children, steal our food, destroy our property, and generally wreak havoc without regard for anyone else. And if they did all these antisocial things, no one would blame them or hold them responsible.<sup>12</sup>

This kind of differential treatment implies assumptions about the differences between chimpanzees' and human beings' structure of cognition that are the reasons why such differential treatment is justified.

Given these observations, it comes as no surprise that any nonhuman animal can be raised with humans without this animal developing the cognitive abilities of human beings, whereas every human infant develops these abilities, whatever human environment they grow up in, drastic cases of deprivation aside.

There is simply no way around the fact that human cognition and nonhuman animal cognition are different and that there is some kind of biological basis for this difference. The question is only what this biological basis is and how to account for its development in evolutionary terms. This is a first important result.

## 7.5 EVOLUTIONARY PSYCHOLOGY

### 7.5.1 *The Morality of Selfish Genes*

Evolutionary psychology represents one very influential account of the origin and nature of human moral psychology. It draws on certain well-known ultimate

<sup>12</sup> Tomasello, *A Natural History of Human Morality*, 39.

evolutionary mechanisms – kin selection, reciprocity and group selection – to explain the evolution of cooperation.<sup>13</sup>

Inclusive fitness and kin selection are seen to be the solutions to the conundrum of other-benefiting behavior. From this perspective, single “selfish” genes are the objects of the evolutionary process.<sup>14</sup> Their reproduction is assured not only if the individual who is acting survives, but also if other individuals who likewise carry these genes have a sufficiently high chance of reproducing. This is the case if these individuals share a sufficiently great amount of the same genetic material: The behavior that benefits other organisms and that seems to be implausibly altruistic from an evolutionary point of view is in fact a tool to increase the chance of reproduction of the gene responsible for this behavior and will therefore be selected for. The proximate psychological mechanism is other-benefiting; the ultimate evolutionary cause for this mechanism is a higher chance of reproduction for the selfish gene.

### 7.5.2 Reciprocal Altruism

Another central element of evolutionary psychology is reciprocal altruism, which consists of tit-for-tat relations: An individual does something beneficial for another organism in order to be paid back, either now or at a later point. One variant of this is indirect reciprocal altruism, which factors reputation into the equation: Actors do something to increase their good reputation in order to reap benefits from others through their standing, albeit perhaps not in the concrete relation in which they are acting, but in other contexts. A good reputation may pay off in the long run.<sup>15</sup>

Evolutionary psychologists argue that reciprocal altruism is both a proximate psychological mechanism and operates on an evolutionary level. As a proximate psychological mechanism, it inclines an organism to other-benefiting behavior if there is a payback, whether direct or indirect: Organisms help others because they, too, will benefit from offering help – not necessarily now, but perhaps later, when

<sup>13</sup> Cf. for an overview Stephen Stich, John Doris and Erica Roedder, “Altruism,” in *The Moral Psychology Handbook*, ed. John Doris (Oxford: Oxford University Press, 2010). As ultimate evolutionary mechanisms, kin selection, direct reciprocity, indirect reciprocity, network reciprocity and group selection play prominent roles in current theory, cf. e.g. Martin Nowak, “Five Rules for the Evolution of Cooperation,” *Science* 314, no. 5805 (2006). For an example of an attempt to explain human cooperation by evolutionary game theory, Leda Cosmides, Ricardo Andrés Guzmán and John Tooby, “The Evolution of Moral Cognition,” in *The Routledge Handbook of Moral Epistemology*, eds. Aaron Zimmerman, Karen Jones and Mark Timmons (New York: Routledge, 2019), 174–228, 195 ff. Kar, “Psychological Foundations,” 122 ff. understands the psychological capacities of humans to use human rights as evolutionary solutions of “social contract problems.”

<sup>14</sup> Richard Dawkins, *The Selfish Gene* (Oxford: Oxford University Press, 2018).

<sup>15</sup> Cf. Richard D. Alexander, *The Biology of Moral Systems* (New York: Aldine de Gruyter, 1987), 94: compensation by improved reputation, rewards by other group members and success of the group.

they themselves are in need. On the evolutionary level, reciprocal altruism is a possible evolutionarily stable strategy because any seemingly other-benefiting act ultimately increases the reproductive fitness of the actor through the benefits it produces and thus is advantageous from an evolutionary perspective.

There is more than one possible proximate psychological mechanism to make an organism behave according to the principles of reciprocal altruism. One is conscious reciprocal altruism – the actor thinks strategically and offers the patient an advantage in order to get something in exchange. Another is – to use de Waal’s terminology – not calculated, but attitudinal reciprocal altruism: The actor is not consciously acting in a strategic manner, but has developed an other-regarding emotional attitude.<sup>16</sup> Some forms of social relationships of primates or other mammals exemplify this – for example, the grooming of kin and friends.

Supporting mechanisms include partner control through sanctioning, partner choice and social selection: Actors choose those partners who are good collaborators, shunning cheaters. Social selection is the product of partner control and partner choice: It promotes the selection of certain individuals because of the attractiveness of having them as a partner.<sup>17</sup>

There is also the idea of multilevel selective processes: An individual’s adaptiveness depends on the adaptiveness of the way of living of the group to which that individual belongs.<sup>18</sup>

### 7.5.3 *The Morality of Tribes*

At the same time, these mechanisms restrict the possible outcome of evolution: Only a small-group morality makes evolutionary sense because of the “environment of evolutionary adaptation” of modern humans, it is often argued. This is the core claim of what has been called “evoconservatism.”<sup>19</sup> Only such a morality would have been selected for because concern for others must be limited to those who plausibly share a sufficiently large number of genes. Given the living conditions of early human beings, who associated in small groups and were faced by out-group competition for resources, concern for a small, limited group is the only possible product of evolution. By evolutionary necessity, there is a natural tribalism, based on an in- and out-group morality.

The logic of this argument leads to the conclusion that the idea of rights that all human beings, not just group members, enjoy, with correlative duties of all human

<sup>16</sup> Frans de Waal, “Attitudinal Reciprocity in Food Sharing among Brown Capuchin Monkeys,” *Animal Behaviour* 60, no. 2 (2020).

<sup>17</sup> Tomasello, *A Natural History of Human Morality*, 18. On partner choice as an explanation for “otherwise puzzling features of human cooperation,” Pascal Boyer, *Minds Make Societies: How Cognition Explains the World Human Create* (New Haven, CT: Yale University Press, 2018), 173 ff.

<sup>18</sup> Haidt, *The Righteous Mind*; Tomasello, *A Natural History of Human Morality*, 142.

<sup>19</sup> Buchanan and Powell, *The Evolution of Moral Progress*.

beings in turn, is not supported by evolutionary psychology. The same follows for the moral principles that give rise to this idea. As discussed, justice, solidarity and respect are not limited to some in-group, but apply to all human beings. Human rights and principles of justice, altruism and respect in this sense have no ally in human nature from this perspective. They are at best justified cultural artifacts created to tame human nature – or, if one takes a more critical view of their ultimate justification, not even that. Instead, they are a kind of heuristic applied only to in-group members that has a certain rhetorical effect strong enough to make them empirically important factors in human behavior, but that has no substantial legitimacy and does not extend beyond the limits of the group.<sup>20</sup>

Such accounts are important for the theory of human rights, even if one is as firmly committed to the *is/ought* distinction as the present inquiry is. Even if one rightly holds that human beings are capable of defying their natural inclinations, whatever they may be, and could and ought to do so in this area, too, by following the commands of a universalistic morality, the question of whether universal human rights are principles and institutions that are contrary to the nature of human beings requires serious consideration. The significance of this point is underlined by various voices that consider the claim that people are in fact not (and cannot be) naturally motivated to respect the rights of every human being to be an important argument against human rights.<sup>21</sup>

#### 7.5.4 Explanatory Problems

The basic mechanisms of evolution mentioned above are of more limited explanatory power than is often assumed, however: As discussed above, there is substantial evidence that the basic assumption underlying many accounts of the evolutionary origin of morality that human beings lived in small groups of hunter-gatherers competing with other such groups is not on the right track. The respective cultures often spread over geographically huge spaces, individuals travelling large distances, creating groups that were not based on biological kinship relations in any

<sup>20</sup> Cf. on this argument Chapter 6.

<sup>21</sup> Richard Rorty, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1989), 191: “[B]ecause she is a human being’ is a weak, unconvincing explanation of a generous action”; Michael Ignatieff, *Human Rights as Politics and Idolatry* (Princeton, NJ: Princeton University Press, 2001), 79; Michael Ignatieff, “Human Rights, Global Ethics, and the Ordinary Virtues,” *Ethics & International Affairs* 31, no. 1 (2017): 3, 11. “From the perspective of ordinary virtues, the first question we ask of another human being is always: Are you one of us or one of them? From this initial question everything follows, including whether we owe this individual anything. If they are fellow citizens, we may owe them shelter, clothing, a hearing, healthcare, and other forms of assistance. If they are strangers, what we owe them ceases to be a duty and becomes instead a matter of pity, generosity, and compassion”; David Miller, *On Nationality* (New York: Oxford University Press, 1995), 80.

discernable sense, but rather on common cultural bonds.<sup>22</sup> The perception that anyone outside a particular group was regarded as an enemy is not tenable either, given the evidence for complex social ties beyond small groups.<sup>23</sup>

Moreover, kin selection is not a mechanism that could explain the emergence of the particular properties of human morality.<sup>24</sup> The architecture of human morality includes such complex phenomena as duties, rights, specific moral emotions and responsibilities. Humans have the ability to create, apply and enforce moral norms. Kin selection is not sufficient to explain this sophisticated cognitive structure. To serve the ends of kin selection, much simpler psychological proxy mechanisms would suffice – say, some straightforward prereflective attraction to kin, without any intricate moral machinery in place. Furthermore, the commands of morality are not limited to kin, neither as an intuition nor in reflective evaluation. Principles of justice can and often do limit the benefits that kin receive, because others can claim an equal share to kin. Morality and justice thus constitute checks on kin favoritism and are not proximate mechanisms to serve it.

Reciprocal altruism raises similar questions. In particular, first, there is the problem of first altruistic acts: What motivates the initial altruistic act, given that no tit-for-tat pattern has been established yet and thus no rational expectation exists that the beneficent of the altruistic act will reciprocate – “blind optimism or accident”?<sup>25</sup> The second problem is the free-rider problem, the “powerful incentive to defect.”<sup>26</sup> What reason is there not to follow this incentive? The same problems

<sup>22</sup> Cf. for instance Douglas W. Bird et al., “Variability in the Organization and Size of Hunter-Gatherer Groups: Foragers Do Not Live in Small-Scale Societies,” *Journal of Human Evolution* 131 (2019): 96–108, 106: “This general disconnect between traditional views of hunter-gatherer social organization and quantitative ethnographic evidence highlights an important weakness in current paleoanthropological/neurological models of the co-evolutionary relationships between human cognition, pro-sociality, and hunter-gatherer group size and organization. Most well-documented highly mobile hunter-gatherers live and work in groups that are relatively smaller, and more fluid in composition, than farming groups. But the smaller local groups of foragers are not drawn from small discrete communities, and none are dominated by close ties of genetic relationship. Variability in the size of residential and foraging groups is likely shaped in landscapes of anthropogenic construction; and their composition is drawn from expansive, trans-generational networks of relational wealth, bound together in ties of social interaction that extend well beyond a small community of individuals”; Graeber and Wengrow, *Dawn*, 122, 279 ff.

<sup>23</sup> Graeber and Wengrow, *Dawn*, 547 n. 4.

<sup>24</sup> Cf. similarly Tomasello, *A Natural History of Human Morality*, 139 arguing that kin selection cannot produce intricate mechanisms of morality such as joint collective commitments, creating and enforcing social norms, feeling resentment, self-regulating one’s actions by feelings of responsibility, obligation and guilt. Richard Dawkins, *The God Delusion* (London: Bantam Press, 2006), 221, understands care for others who are unrelated and unable to reciprocate (as sexual desire without the possibility of procreation) as “misfirings, Darwinian mistakes: blessed, precious mistakes.”

<sup>25</sup> Tomasello, *A Natural History of Human Morality*, 13.

<sup>26</sup> Tomasello, *A Natural History of Human Morality*, 13.

arise for indirect reciprocity, which takes reputation into account.<sup>27</sup> Why should there be no free-riding here?<sup>28</sup> What is needed to break this deadlock is either an agreement to adhere to certain norms, as has been rightly underlined,<sup>29</sup> which is beyond the cognitive capacities of nonhuman animals, or, it is important to add, a form of altruism that is not conditioned upon reciprocal advantages. Similarly, in the case of emotional, attitudinal reciprocal altruism one can ask: Why do individuals form friendly social relationships with nonkin in the first place? Why do they help these friends?<sup>30</sup>

This aside, it is decisive that human morality is limited neither to calculated nor to attitudinal reciprocal altruism in descriptive terms. As already underlined above and as will be fleshed out in greater detail below, it is a complex reflective capacity to judge corresponding principles of altruism and justice with motivational and emotional consequences, and thus it is neither a calculation of the benefits of nor simply an emotional attitude fostering tit-for-tat advantages.

The paradigmatic case of moral acting is sacrificing some of one's own goods for the benefit of others – agents perhaps even giving their life to save others. The evaluation of such an act as morally good is not restricted to saving kin and friends. On the contrary, agents who save only their kin and friends when they are able to save others as well, or saving others only because they are expecting advantages as a result of their deed, would not count as performing a particularly impressive example of moral action. Similar observations hold for justice: Distributing something justly to receive benefits oneself is not what the idea of justice is about.

It is not a good counterargument to assert that early modern humans' morality was different from what we find today. As we will see, we have only indirect clues to the cognitive abilities of early modern humans. None of them conclusively shows that moral cognition was limited to a small-group morality, speculative just-so stories aside, which are, moreover, based on the most probably wrong assumption, as discussed above, that hunter-gatherers did not live in larger cultural spaces.

Pointing to the distinction between proximate mechanisms and ultimate evolutionary causation does not help to counter this critique either: The nonreciprocal altruism of human beings and their sense of justice, among other principles, do not with sufficient probability produce paybacks for those adhering to them. The fact that being moral regularly does not pay constitutes a central conundrum of moral philosophy: For Kant, for instance, the fact that acting morally often does not lead to particular happiness and not seldom leads to the dungeon and the stake was so obvious that it served as an argument for the existence of God and eternal life, so that the scales of justice could be balanced at least in the afterlife.<sup>31</sup> Consequently,

<sup>27</sup> Tomasello, *A Natural History of Human Morality*, 14.

<sup>28</sup> It is useful to remember the results of the ultimatum and dictator games here – despite continuing selfishness, agents clearly do not exploit any chance to get a free ride, cf. Chapter 6.

<sup>29</sup> Tomasello, *A Natural History of Human Morality*, 13.

<sup>30</sup> Tomasello, *A Natural History of Human Morality*, 14.

<sup>31</sup> Kant, *Kritik der reinen Vernunft* (2. Ed. 1787), 525 ff.

human morality cannot be understood merely as a proximate psychological mechanism for reciprocal altruism as an evolutionarily stable strategy. It does not deliver the advantages that reciprocal altruism presupposes.

A further point of criticism homes in on the assumption that forms the foundation of the evolutionary approach, namely that all of an organism's traits have an adaptive function. This poses overly narrow constraints on the process of evolution, as will be illustrated below.

There are, however, other accounts that paint a different picture of the evolutionary origin of morality. Arguably, the currently most sophisticated theory of the evolution of morality describes a process driven forward by mutualism and interdependence, enabled, however, by a wide-ranging and complex, inborn, species-unique cognitive endowment of human beings. Because of its paradigmatic importance for current debates, a closer look is necessary to see whether this approach by Michael Tomasello solves the riddles of the evolution of human morality.

## 7.6 THE POWER OF JOINT INTENTIONALITY: INTERDEPENDENCE AND COOPERATION

### 7.6.1 *Some Specifics of Human Cooperation*

Tomasello's theory, too, is guided by the idea of cooperation: Its central thesis is that human moral psychology is a phenomenon emerging from the increasingly sophisticated forms of collaboration of hominids and finally human beings. Cooperation is also key to moral ontogeny: Moral development does not depend on general intelligence, but on a set of inborn, species-specific cognitive capacities universally possessed by all human beings independent of cultural background, which enable human beings to develop their rich moral world through cooperative constructive interaction similarly to the way they acquire other parts of their cognitive and social world:<sup>32</sup> "It is precisely this biological preparation – in the form of maturationally expressed capacities – that makes uniquely human sociocultural activities and

<sup>32</sup> Michael Tomasello, *Becoming Human: A Theory of Ontogeny* (Cambridge, MA: Harvard University Press, 2021). The set of inborn capacities includes joint attention/common ground, cooperative/referential communication, role-reversal imitation, dual-level collaboration and basic helping, *ibid.* 312 f. Tomasello underlines that maturation and experience all go hand in hand in complex psychological phenomena. The theory is therefore "Neo-Vygotskian," *ibid.* 6, 34, 297 ff. For these five competencies, however, the maturational (that is, inborn) element predominates, *ibid.* 313. He identifies further species-specific inborn capacities, from pointing, iconic gestures and pretense to laughter and smiling, *ibid.* 54 ff., 95 ff., 106 ff. The main factors driving development are interaction with peers and adults and executive self-regulation, *ibid.* 36. The terminology adapted here takes for granted that the predicates "innate" or "inborn" do not exclude the influence of triggering experiences. They do not imply that the process is "impervious to experience," *ibid.* 83. Even acquiring the ability to walk upright is dependent on the child having the chance to move their limbs, even though the ability itself is innate.

experiences possible in the first place.”<sup>33</sup> It is not just “general biological preparedness for learning and inference, but also specific biological preparedness for uniquely human cooperative interaction and mental coordination with others.”<sup>34</sup> The capacity for joint intentionality is key:<sup>35</sup> It is the “ultimate source of human uniqueness.”<sup>36</sup>

The main criteria for identifying capacities as innate are a lack of cultural diversity of outcome in experimental intercultural designs with small children and very similar maturational age trends across cultures.<sup>37</sup> All children, for example, learn the complex activity of upright walking irrespective of cultural background in a “quite predictable developmental period.”<sup>38</sup> It is thus safe to assume that the capacity of upright walking is innate.

Cooperation between human beings extends to the rules they apply to their own cooperative activities: The “‘cooperativization of self-regulation’ is of the essence of normative sociality and morality.”<sup>39</sup>

The theory (quite correctly) emphasizes how fundamental human cooperation is for the human way of living. While cooperative cultural institutions in particular create a framework for competition, competition is not the true driver of human development: “It is only if one neglects the cultural-institutional context of human behaviour that one can hallucinate the competitive cart as leading the cooperative horse.”<sup>40</sup>

As a first very important step, this approach goes to greater effort than others properly to account for the specifics of forms of human cooperation. This includes identifying the various layers of the cognitive dimensions and preconditions of cooperation. This proves very fruitful because cooperation is a highly underdetermined concept – as illustrated above, cooperation can take very different forms and does not necessarily imply moral norms – it could be based, for instance, solely on force and power.

A battery of experiments provides some clues about the differences between the cooperative behavior of humans and that of great apes. More precisely, this research concerns the behavior of small children and thus does not even focus on the mature forms of cooperation between adults because Tomasello argues that the cognitive capacities of children are similar to the cognitive capacities of early humans. These findings provide at least some good indications of distinctively different cognitive dimensions of forms of cooperation, spelled out in respective behavior. These include the following patterns, for instance:<sup>41</sup>

<sup>33</sup> Tomasello, *Becoming Human*, 7.

<sup>34</sup> Tomasello, *Becoming Human*, 301.

<sup>35</sup> Tomasello, *Becoming Human*, 8, 15 ff., 27, 56, 63, 190, 304.

<sup>36</sup> Tomasello, *Becoming Human*, 86.

<sup>37</sup> Tomasello, *Becoming Human*, 312 f.

<sup>38</sup> Tomasello, *Becoming Human*, 33.

<sup>39</sup> Tomasello, *Becoming Human*, 215.

<sup>40</sup> Tomasello, *A Natural History of Human Morality*, 158.

<sup>41</sup> Tomasello, *A Natural History of Human Morality*, 20 ff.; Tomasello, *Becoming Human*, 45 ff., 193 ff., 219 ff.

Children cooperate with joint intentionality: They pursue joint goals with others, and with joint attention. They reverse roles in collaborative activities and communicate cooperatively to coordinate collaboration. Children divide goods in particular ways: They share goods even if these goods are readily monopolized by one partner. They display a tendency to share goods more equally if these goods are the result of collaborative efforts. They have a preference for collaboration even if the payoffs are identical to those of solo activity. In relation to their partners in the respective collaborative activity, children share less with a free rider. They help others in various forms. They modify their behavior when peers are watching. They show a particular quality of commitment to the collaborative activity: They continue to collaborate until the end, even if they have already received the advantages offered by the collaboration. Unlike great apes, children from the age of three sanction others for violating social norms, and they do so on behalf of others. This indicates that their attitude towards social norms is not just prudential.<sup>42</sup> They are as prepared to help when they are alone as when they are with their mother. They help persons who do not know that they are being helped. Such research bolsters the view that small children are intrinsically motivated to foster the well-being of others. None of the above examples forms part of the cognitive repertoire of great apes:<sup>43</sup> “[G]roup actions of apes are all about individuals achieving their individual ends in group contexts – they are using one another as social tools.”<sup>44</sup>

This provides some hints on how different the cognitive apparatus of great apes and of human beings already is at a relatively early age. It illustrates vividly the point made above: One needs to spell out what kind of cooperation one is talking about and what the cognitive preconditions are that make it possible before one even can start thinking about an evolutionary theory of this form of cooperation.

### 7.6.2 *Sympathy and Fairness Develop in Small Steps*

The evolutionary theory of morality that tries to explain these findings is a theory of the incremental development of morality. This incremental development theory offers an alternative to standard accounts based on kin selection and inclusive fitness, which are regarded as insufficient.<sup>45</sup>

Tomasello’s approach starts by identifying two forms of cooperation: altruistic concern for others and mutualistic collaboration and, correspondingly, a morality of sympathy and of fairness. The former is freely performed, while the latter is accompanied by a sense of obligation and deservingness and by punitive moral attitudes such as resentment or indignation towards others who are unfair. The

<sup>42</sup> Tomasello, *A Natural History of Human Morality*, 101.

<sup>43</sup> Tomasello, *A Natural History of Human Morality*, 76 for a summary of these findings. On the lack of joint intentionality in great apes, Tomasello, *Becoming Human*, 13.

<sup>44</sup> Tomasello, *Becoming Human*, 190. Cf. e.g. on chimpanzee group hunting, *ibid.* 193 ff.

<sup>45</sup> Tomasello, *A Natural History of Human Morality*, 13 ff.

former is pure cooperation, the latter the “cooperativization” of competition to find balanced solutions for conflicting demands. Morality in this sense is unique to human beings.<sup>46</sup>

We proceed from the assumption that human morality is a form of cooperation, specifically, the form that has emerged as humans have adapted to new and species-unique forms of social interaction and organization. Because *Homo sapiens* is an ultracooperative primate, and presumably the only moral one, we further assume that human morality comprises the key set of species-unique proximate mechanisms – psychological processes of cognition, social interaction and self-regulation – that enable human individuals to survive and thrive in their especially cooperative social arrangements.<sup>47</sup>

Note that this passage raises an important problem that will occupy us in more than one context: According to this formulation, there are “new and species-unique forms of social interaction” *prior* to human morality, as morality is a means to adapt to these forms of interaction. The problem is: Do not these “new and species-unique forms of social interaction” *presuppose* the existence of the cognitive mechanisms necessary for morality? Is a certain cognitive apparatus not the *precondition* for certain forms of living, including forms of collaboration, and consequently cannot be interpreted as an *adaptation* to these forms? If the cognitive abilities are *prior* to the “new and species-unique forms of social interaction” they enable, the evolutionary theory needs to account for the emergence of these cognitive abilities independently of the way of living that was only possible *after* these cognitive abilities were acquired.<sup>48</sup>

How, then, did this “key set of species-unique proximate mechanisms,” the “psychological processes of cognition, social interaction and self-regulation” that are the natural mental foundations of human morality, evolve?

Tomasello argues that all began far back in time, with hominid species: First, there were dyadic relations of cooperation in hominids that led via role formation to normative standards. These cooperative relations were established to aid survival by foraging together, which led to sympathy beyond kin and friends, including collaborative partners.<sup>49</sup> In order to enable them to collaborate, hominids evolved joint intentionality – that is, the capacity to form a common goal with a partner and to know things together.<sup>50</sup> At the same time, Tomasello asserts that a context of collaboration including “obligate collaborative foraging with various and robust means of partner choice and partner control” is necessary for the development of

<sup>46</sup> Tomasello, *A Natural History of Human Morality*, 1 f.

<sup>47</sup> Tomasello, *A Natural History of Human Morality*, 3.

<sup>48</sup> The concept of adaptation becomes meaningless if every property of a species is regarded as an adaptation to the kind of life the organism’s capacities enable it to lead.

<sup>49</sup> Tomasello, *A Natural History of Human Morality*, 4.

<sup>50</sup> Tomasello, *A Natural History of Human Morality*, 4.

joint intentionality.<sup>51</sup> This is a concrete example of the problem just identified: According to this account, are forms of collaboration the precondition for the evolution of cognitive abilities such as joint intentionality or, vice versa, are cognitive abilities the preconditions for the possibility of certain forms of collaboration?

This collaboration shaped role models of what it meant to be a good partner, which were the first step to impartial normative prescriptions. This had a major consequence: “Recognizing the impartiality of role standards meant recognizing that self and other were of equivalent status and importance in the collaborative enterprise.”<sup>52</sup> The self–other equivalence led to mutual respect among partners. Each could hold the other accountable for living up to this role ideal. A self-regulating “we” emerged that was considered legitimate because it derived from a joint commitment that the partners had created themselves and because both regarded the other as deserving of respect. This was the birth of an “evolutionarily novel form of psychology.”<sup>53</sup> The ensuing behavior is motivated by genuine moral motivation, not just by the wish to avoid punishment or reputational attacks: “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’.”<sup>54</sup> Tomasello hypothesizes that this natural, second-personal morality emerged 400,000 years ago.<sup>55</sup>

### 7.6.3 *The Path to Objective Group Morality*

A second evolutionary step was taken with the emergence of *Homo sapiens*. The reasons for this step lay in demographic development: Groups of modern human beings formed tribes that built up a culture, “one big interdependent ‘we’” that secured role conformity through identification with the group and sympathy of the group members towards each other in order to ensure survival. The individual learned to use the perspective of “any rational person,” based on the self–other equivalence, and thus “served as interchangeable (agent-independent) cogs in the conventional cultural practices that kept the culture going.”<sup>56</sup> Humans become persons through social recognition.<sup>57</sup> In order to enable the cognitive coordination of their activities, group members evolved “new cognitive skills and motivations of *collective intentionality* – enabling the creation of cultural conventions, norms and institutions.”<sup>58</sup> Group members knew about similarities between them and knew that

<sup>51</sup> Tomasello, *A Natural History of Human Morality*, 40.

<sup>52</sup> Tomasello, *A Natural History of Human Morality*, 4.

<sup>53</sup> Tomasello, *A Natural History of Human Morality*, 5.

<sup>54</sup> Tomasello, *A Natural History of Human Morality*, 5.

<sup>55</sup> Tomasello, *A Natural History of Human Morality*, 86.

<sup>56</sup> Tomasello, *A Natural History of Human Morality*, 96.

<sup>57</sup> Tomasello, *A Natural History of Human Morality*, 106.

<sup>58</sup> Tomasello, *A Natural History of Human Morality*, 5, 92 ff. (emphasis in original).

the others knew, too.<sup>59</sup> Conventional cultural practices were turned into social norms, representing the right and wrong thing to do, and substituting partner control – that is, the sanctioning of cheating – by social control with the aim of creating conformity.<sup>60</sup> Institutions enshrined social norms and created institutional facts.<sup>61</sup> New forms of *cultural agency* with respect to the group's conventions, norms and institutions emerged.<sup>62</sup> Humans found new forms of self-regulating *self-governance* based on collective commitments that created obligations and led to moral judgments about others' moral judgments and to specific moral emotions, such as guilt.<sup>63</sup>

Group identity was based on similarity; solidarity was therefore felt with in-group members who resembled the actor in behavior and appearance, not with “out-group barbarians” in the vicinity.<sup>64</sup> The fact that the preferential treatment of in-group members as compared to out-group members already is evident in children is taken to confirm this point: “In-group favoritism accompanied by outgroup prejudice is one of the best-documented phenomena in all of contemporary social psychology . . . , and it emerges in young children during the late preschool and especially during the school-age period.”<sup>65</sup>

Moral norms were considered legitimate because the members of groups identified with their respective culture, formed a moral identity, regarded themselves as coauthors and felt respect towards the other group members.<sup>66</sup> Deviation could only be justified by recourse to the shared values of the moral community. These processes led to the 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.”<sup>67</sup> For modern human beings, the result of this objectification of conventional norms was a “kind of cultural and group-minded ‘objective’ morality.”<sup>68</sup> As a consequence, human beings harbor three kinds of morality in their breast: the “cooperative proclivities of great apes, the ‘joint morality of collaboration’ and an ‘impersonal collective morality’ of cultural norms and institutions in which all members of the cultural group are equally valuable.”<sup>69</sup> These kinds of moralities may conflict in dilemmas such as: “Shall I steal for a friend?”

<sup>59</sup> Tomasello, *A Natural History of Human Morality*, 93 f.

<sup>60</sup> Tomasello, *A Natural History of Human Morality*, 98 f.

<sup>61</sup> Tomasello, *A Natural History of Human Morality*, 104.

<sup>62</sup> Tomasello, *A Natural History of Human Morality*, 87, 107 ff.

<sup>63</sup> Tomasello, *A Natural History of Human Morality*, 89, 108 ff.

<sup>64</sup> Tomasello, *A Natural History of Human Morality*, 89 f., 122.

<sup>65</sup> Tomasello, *A Natural History of Human Morality*, 91.

<sup>66</sup> Tomasello, *A Natural History of Human Morality*, 86 ff.

<sup>67</sup> Tomasello, *A Natural History of Human Morality*, 6.

<sup>68</sup> Tomasello, *A Natural History of Human Morality*, 6, 123.

<sup>69</sup> Tomasello, *A Natural History of Human Morality*, 7.

Social norms as objectified conventional norms are moral insofar as they conform with the second-personal morality of sympathy and fairness<sup>70</sup> – this “natural morality” forms a baseline of human morality, not only in developmental but also in normative terms. Interestingly, Tomasello’s theory incorporates studies that suggest that young children regard norms frequently connected with sympathy and fairness as applicable to all humans, not just in-group members. This does not seem to be easy to reconcile with a simple in-group/out-group dichotomy and the supposedly dyadic nature of second-personal morality.

Cases giving rise to competing moral claims need to be solved on an individual basis using “creativity.” Moral debates are crucially shaped by the second-personal moral principles of sympathy and fairness, what these principles mean and the answer to the question of who is part of the moral community.<sup>71</sup> Exclusion from the moral community was, it is argued, the reason for the justification of slavery or apartheid. The standards of inclusion and exclusion can change.<sup>72</sup> “Moral entrepreneurs” like Martin Luther King Jr. and Mahatma Gandhi were able to move people by putting suffering and injustice in front of people’s eyes: a “second-person protest writ large.”<sup>73</sup>

In any case, there is a genuine, not just strategic concern for morality,<sup>74</sup> although the ultimate evolutionary cause is to increase reproductive fitness – morality ultimately emerged because “early humans who were concerned for the welfare of others and who treated others fairly had the most offspring.”<sup>75</sup> Moral humans simply had “more babies.”<sup>76</sup>

Children are socialized in the moral community. Second-personal morality is culturally universal, Tomasello asserts, objective morality is not. This can be explained by reference to the ontogenesis of morality: Children first develop a second-personal morality and later the culturally relative social norms.<sup>77</sup> Humans in general have some “built-in responses to morally relevant situations based on intuition and emotion that have evolved to deal with evolutionarily important situations, especially those in which there is no time for a considered decision.”<sup>78</sup>

<sup>70</sup> Tomasello, *A Natural History of Human Morality*, 100, 122, 124, 127.

<sup>71</sup> Tomasello, *A Natural History of Human Morality*, 127.

<sup>72</sup> Tomasello, *A Natural History of Human Morality*, 113 ff., 126 f.

<sup>73</sup> Tomasello, *A Natural History of Human Morality*, 134; Tomasello, *Becoming Human*, 290: “[T]he individual is always free to go beyond the culture’s social norms if necessary.”

<sup>74</sup> Tomasello, *A Natural History of Human Morality*, 108: Children give to a needy person despite seeing others not doing so (not just reputation management); *ibid.* 111: “The proximate psychological mechanisms responsible for human moral action do not involve, essentially, prudential concerns for one’s self-serving interests or strategic calculations of one’s own reputation: they involve moral judgement by a moral self (with the representative authority of the moral community) that endures over time and that judges the self impartially in the same way that it judges others.”

<sup>75</sup> Tomasello, *A Natural History of Human Morality*, 7.

<sup>76</sup> Tomasello, *A Natural History of Human Morality*, 149, quote 163.

<sup>77</sup> Tomasello, *A Natural History of Human Morality*, 117, 155 ff.

<sup>78</sup> Tomasello, *A Natural History of Human Morality*, 118, following Haidt.

Besides the evolution of the principles of morality that are foundational for human rights, rights as a normative category (albeit not human rights) Tomasello explicitly discusses as part of human morality.<sup>79</sup> In addition, the reflection about fairness and “deservingness” as other considerations seems to imply the idea of the rights of moral patients. Law in general is conceptualized as an institutionalized framework of social integration.<sup>80</sup> The idea of equal respect, which is foundational for human rights, is also important, albeit somewhat ambiguous: It is limited to a cultural group but can be extended by moral entrepreneurs to all human beings.

Concern for the well-being of others does not mean reciprocal altruism: More demanding principles are involved. The two problems mentioned above – how to explain the first altruistic act and the free-rider problem – indicate the explanatory limits of an account based on reciprocal altruism. The element crucial to overcoming this difficulty, especially the “undermining effect of cheating,” is interdependence.<sup>81</sup>

Mutualism (as reciprocity) works at the level of the individual. A stakeholder model is endorsed: Mutualism provides benefits for others, depending on the stake the agent has in the survival (because of the benefits the other provides) of their partner in mutualism.<sup>82</sup> Group selection is not held to be a major influence, because the gene flow between different groups prevents the development of a genetic makeup that is sufficiently differentiated between groups.<sup>83</sup>

Tomasello argues, however, for cultural group selection: Those groups with the “conventions, norms and institutions that best promoted cooperation and group cohesion won out, by assimilating or eliminating competitors from other groups.”<sup>84</sup> Such groups formed collective agents. With the sedentary lifestyles established by the Agricultural Revolution, intergroup competition shifted towards a reconciliation of the moralities of different subgroups.<sup>85</sup> Gene–culture coevolution can be a factor in this process.<sup>86</sup> The preconditions for cultural group selection are “species-universal skills and motivations for creating social norms and institutions in the first place,”<sup>87</sup> as general learning mechanisms are not sufficient to create moral mental categories.<sup>88</sup> The evolutionary process of moral development was initiated by the shortage of individually available food, a growing population and group competition.<sup>89</sup>

<sup>79</sup> Tomasello, *A Natural History of Human Morality*, 101.

<sup>80</sup> Tomasello, *A Natural History of Human Morality*, 130 ff.

<sup>81</sup> Tomasello, *A Natural History of Human Morality*, 147.

<sup>82</sup> Tomasello, *A Natural History of Human Morality*, 15.

<sup>83</sup> Tomasello, *A Natural History of Human Morality*, 12 ff.

<sup>84</sup> Tomasello, *A Natural History of Human Morality*, 119.

<sup>85</sup> Tomasello, *A Natural History of Human Morality*, 134.

<sup>86</sup> Tomasello, *A Natural History of Human Morality*, 12.

<sup>87</sup> Tomasello, *A Natural History of Human Morality*, 142.

<sup>88</sup> Tomasello, *A Natural History of Human Morality*, 154.

<sup>89</sup> Tomasello, *A Natural History of Human Morality*, 136.

Interestingly, Tomasello posits that similar mechanisms are involved in this phylogenetic trajectory as are at work in the ontogeny of human beings.<sup>90</sup>

#### 7.6.4 *Self–Other Equivalence as a Spandrel*

Tomasello’s argument contains an important pillar: Self–other equivalence is crucial to the development of morality.<sup>91</sup> It is argued that, at the beginning, this was a purely cognitive judgment that paved the way for the normative “deservingness” of human beings to be treated equally.<sup>92</sup> This is important, as there is a gap between the judgment that two agents fulfill equivalent roles and the judgment that they deserve equal respect, the former a factual, the latter a normative statement.<sup>93</sup>

The cognitive judgment of self–other equivalence thus forms the basis for the emergence of new normative principles. It is “moral-structural” in the sense that it is a cognitive structure that could be recruited for moral functions. It is thus, in Gould and Lewontin’s famous terms, a “spandrel”<sup>94</sup> – an architectural element used for other purposes than the one for which it was built. This is a crucial observation – it opens the door to additional dimensions in evolutionary accounts of cognitive abilities and invites us to take a fresh look at morality.

The theory of interdependence takes the argument very seriously that the evolution of a species motivated to subordinate or equate their own interests to those of others cannot be explained by natural selection. The reference to spandrels sends this “spectre,”<sup>95</sup> Tomasello highlights, to its final resting place.

#### 7.6.5 *Paradigmatic Incrementalism*

This theory is arguably currently the most detailed attempt to explain the evolution of moral cognition, relating its findings to current moral philosophy and social science. It stands for a certain research paradigm in evolutionary theory, one that

<sup>90</sup> Cf. Tomasello, *Becoming Human*, 195 ff. (joint intentionality, second-personal agency and morality); *ibid.* 198 ff. (self–other equivalence and the emergence of normative claims, role ideals, role reversals); *ibid.* 204 ff. (self-regulation by joint commitments, second-personal protest as indicator of respect); *ibid.* 249 ff., 287 (collective intentionality).

<sup>91</sup> Tomasello, *A Natural History of Human Morality*, 150 ff.

<sup>92</sup> Tomasello, *A Natural History of Human Morality*, 150 ff.

<sup>93</sup> Tomasello, *Becoming Human*, 201: “[T]he recognition of self–other equivalence is not by itself a moral motivation or act; it is simply the recognition of an inescapable fact that characterizes the human condition.”

<sup>94</sup> Tomasello, *A Natural History of Human Morality*, 151: “The term structural indicates that, on the evolutionary level, this dimension of human morality was not originally selected to serve this function; it came into existence in the service of other functions. Of course, this would not work if the recognition of self–other equivalence was maladaptive in the social-interactive contexts we are considering, but it could easily be a cost-neutral ‘spandrel’ that cognitively structured such contexts.”

<sup>95</sup> Tomasello, *A Natural History of Human Morality*, 153.

can be called functional incrementalism: Its leading hypothesis is that all cognitive capacities have evolved ultimately to increase reproductive fitness through an incremental process of small evolutionary steps. From this perspective, morality is a functional tool narrowly tailored by natural selection to enable cooperation, albeit enabled by the “spandrel” of self–other equivalence.

The questions that this account raises vividly illustrate some of the challenges encountered by an evolutionary explanation of morality. These questions are thus likewise of paradigmatic importance.

#### 7.6.6 *The Analysis of Morality and the Evidence for Evolutionary Incrementalism*

A first problem to be discussed is the precise analysis of human morality: As already explained, in morality, there are not only sympathy and fairness, sympathy understood as an inclination to benefit others, but there is also (under qualified circumstances) an obligation to foster the well-being of others. It is not only the commands of justice that have obligatory force. Certain commands of human benevolence do as well: There is an obligation not to let somebody die in front of your porch if you can pick up the phone and call an ambulance. Such actions are evaluated as being morally good. Therefore, they possess a specific deontic status. Moreover, morality concerns not just the behavior, but also specific kinds of intentions of agents. Sympathy and fairness thus do not exhaust the content of morality.

A second problem is that there is no evidence confirming that human moral evolution did in fact take the steps described, that there were first dyadic forms of interaction in which role models emerged, that cooperative forms of foraging lead to sympathy and equal respect, to moral self-regulation and a moral ought, that hominids had a sense of legitimacy and that anything like a second-personal morality existed among them. The same holds for the second step, the development of the objective morality of groups, the exclusion of outsiders based on “similarity,” the individual turning into a “cog in the cultural machine” or the group-mindedness of moral orientation. There is no evidence either that early human beings were like young children in terms of their cognitive capacities. There may be a huge, qualitative difference between the maturing cognitive capacities of the young children of modern humans and the mature capacities of hominids.

Tomasello is entirely honest about the complete lack of evidence that the described stages did in fact exist. The methodological problem of sparse evidence<sup>96</sup> is acknowledged, and Tomasello highlights that what is presented is not an account of human cognitive development sufficiently based on paleoanthropological evidence, but a “speculative evolutionary narrative.”<sup>97</sup> It is an “imaginative

<sup>96</sup> Tomasello, *A Natural History of Human Morality*, 105.

<sup>97</sup> Tomasello, *A Natural History of Human Morality*, 143.

reconstruction of historical events with little in the way of artefact or paleoanthropological fact to help.”<sup>98</sup>

This comes as no surprise: There are not many clues generally that help us to understand what kind of cognitive skills hominids possessed. Artifacts such as tools are key, as are some indications of behavioral patterns, like signs of cannibalism or burial rites (or the lack thereof). Many cognitive activities simply leave no traces behind. We know from striking pieces of art that early modern human beings evidently had the cognitive capacity to produce art. We do not know, however, whether they engaged in mathematical thinking, in whatever form, because such thinking leaves no traces in the paleoanthropological record, though some artifacts like the Ishango bone raise intriguing questions. This is also true for moral precepts. Whether group life in some hominids included ideas of “self–other equivalence” or not, whether “legitimacy” played any role or not or whether ideas of a moral “ought” structured their thought along the lines of a second-personal morality or not is anybody’s guess, because there are simply no clues that would help to decide these questions. One cannot brush over this state of affairs. It is very important – not to mention an issue of fundamental scientific ethics – to be prepared to leave questions open if there is no way of answering them given the evidence available.

#### 7.6.7 *Stumbling Blocks on the Way to Second-Personal Morality*

Further important lessons for the understanding of the evolution of morality can be drawn from the details of the “speculative evolutionary narrative.” These concern in particular the conceivable argument that the lack of evidence does not matter as the evolutionary trajectory had to be as described, because this and only this trajectory makes evolutionary sense.

A key question in this respect concerns the transition from one stage of the hypothesized incremental process of cognitive development to the next. There are two dimensions to this process – the development of new skills acquired by individuals and the development of new genetically inherited skills of the species. The first must be based on some kind of learning mechanism and the cognitive ability to create cultural traditions if these skills are passed on to the next generation to learn anew. The second must be based upon genetic change – that is, mutations that inscribe certain cognitive capacities in the human genome.

According to this theory, there is in certain respects a parallel between phylogeny and ontogeny – similar mechanisms are held to drive the development of the species and of the individual forward, such as joint intentionality, role-taking or self–other equivalence.<sup>99</sup> According to the incremental account of evolution, some of the steps described seem to be taken by individuals, although it is not quite clear how many.

<sup>98</sup> Tomasello, *A Natural History of Human Morality*, 154.

<sup>99</sup> Cf. n. 90.

In any case, hominids acquire increasing cognitive capacities during the development of the species until the species-specific competencies that Tomasello identifies as the basis of modern humans' moral ontogeny have evolved.<sup>100</sup>

These different competencies are responses to new ecological challenges. Those individuals possessing the properties making them able to meet these challenges have an adaptive advantage that sets the stage for further developments.<sup>101</sup> The question is, then, whether one of the stages described leads necessarily to the next. If not, if there are other developmental options, what the evolutionary trajectory really looked like and to what ecological challenges hominids and modern humans adapted remain open questions, given that, as Tomasello underlines, no conclusive paleoanthropological evidence bearing on these questions exists. That any such evolutionary necessity exists is far from clear, however. On the contrary, the apparently continuous development from the dyadic interactions of foragers to human moral codes contains yawning gaps. Crucially, in order to bridge these gaps, certain cognitive resources are necessary that are qualitatively different from those required to engage in other forms of interaction assumed to be prior, without there being any discernible evolutionary necessity that these capacities and no others had to evolve.<sup>102</sup>

To illustrate what is at issue: The development starts, it is argued, in dyadic relations of cooperation. If one assumes for the sake of the argument that there are role models for cooperation (although, as explained above, no evidence is presented either for this or for the consecutive steps), these may give rise to prudential rules (how to be a good forager), but not to moral rules (help other group members in need). Role models derived from certain tasks like foraging are functionally circumscribed, moral rules are not. There is no bridge from one to the other. This problem arises in the next stage of the development as well: Cooperative foraging does not necessarily lead to sympathy for one's partner, an instrumental relation to them is also possible – and the reality is that there are many forms of human collaboration, from soccer to lawyering, where one can effectively cooperate (play the decisive pass on the field or in the courtroom) without sympathy (because one wants to win, not because one likes the striker or one's legal teammate).

This problem continues to plague this incremental account: Equal importance in cooperative foraging may not lead to equal mutual respect either. First of all, the roles in cooperative foraging need not be equal but may be very different, according to a division of labor. In modern humans, collaboration does not necessarily

<sup>100</sup> Tomasello, *Becoming Human*, 45 ff., 312 f.

<sup>101</sup> Tomasello, *Becoming Human*, 5, 11.

<sup>102</sup> Tomasello, *Becoming Human*, 11, refers to "environments of evolutionary adaptedness." Such environments do not necessitate specific developments, given that they merely make certain developments evolutionarily advantageous and that there are many different ways to adapt to such an environment. Moreover, it is unclear what these environments exactly looked like, given the lack of paleoanthropological evidence.

lead to equal respect between cooperators – say, the bank teller and the bank’s CEO. The “impartiality of role standards” (if there are any) has the effect that there is an objective yardstick for measuring whether a role-defining task is performed well, but it does not necessarily have the effect that everyone is regarded as being of “equivalent status and importance in collaborative enterprises,” as actors may fail to fulfill these role standards or may simply fulfill a functionally marginal role according to an impartial role standard, leading to a low status – as the low social respect for cleaners in human societies exemplifies (unfortunately, of course). In addition, any such notions of equality could be limited to the collaborative effort and not extend what is supposed to develop, namely equal respect for all. Self–other equivalence in role fulfillment does not help either. Role equivalence (if given) does not entail mutual respect, because the latter encompasses all of a person’s characteristics, not just their ability to fulfill roles. (You can respect somebody as being as good a defender as yourself on your football team and disrespect them in other contexts.)

The assumed legitimacy of such role-derived norms of an “us” indicates another problem. It presupposes (without further argument) that legitimacy was an important concept for hominids. In addition, a kind of social contract, a “joint commitment,” and the idea of the equal status of others are invoked to provide legitimacy – substantial and controversial assumptions about normative reasons. It is implied, not explained, why these reasons were necessarily valid reasons for hominids more than 400,000 years ago – and not just concepts that were alien to the cognitive capacities of hominids or, if comprehensible at all, were as unconvincing as critics of contractarian theory consider them to be today.

Moral “self-regulation” is a demanding cognitive task, including the idea of moral obligations to do or not to do something and the idea of respect for obligations that motivate the agent to perform the obligatory action. The origin of such an “ought” as a central element of self-regulation creates specific riddles. The need to fulfill a role may lead to an instrumental ought (if you want to catch this prey, you ought to approach it without making any noise), but not to a moral ought. Holding others “accountable” – with sanctions, one has to assume – can lead to strategic behavior, conforming to norms when sanctions are likely and not conforming to norms when the chance of being sanctioned is remote, but it will not necessarily lead to moral self-regulation.

How this development leads to joint intentionality is unclear, too. This is no minor point, because joint intentionality is the very precondition for certain forms of collaboration, as the research on the difference between great apes and children shows, not its consequence – another example of the problem flagged above. Again, there are explanatory gaps between the different stages of development. Thus, phylogenetically, there is no evolutionary necessity for one stage to follow the other as described – many other developmental pathways are open, including not developing any of the species-specific capacities of modern humans discussed by

Tomasello, as indeed many species, including great apes, have failed to do.<sup>103</sup> Ontogenetically, similar conclusions must be drawn, because the incremental steps described are, for the reasons just discussed, also not sufficient to explain how individuals ultimately develop the cognitive capacities that are the basis for their normative thinking and culture.

### 7.6.8 *The Objective Morality of Cogs in the Machine*

Similar problems are encountered in the second stage, the development from the natural, second-personal morality to the cultural and group-minded “objective” morality, “scaling up” sympathy and justice. To illustrate this with some examples: The perspective of “any rational person” now taken entails many things, not least a concept of rationality. Where does this concept come from? How role conformity leads to a strong idea of universalization and even rationality remains unexplained. Self–other equivalence presupposes that the other is regarded as a person. This seems irreconcilable with Tomasello’s assertion that personhood depends on recognition by others, not recognition of others as equal on the prior understanding of the shared personhood of human beings.

The problems of the relevance of the concept of legitimacy and the reasons providing legitimacy are pertinent here, too. Morality is not just based on a group-minded rationality. According to some elements of the discussed theory, morality is content neutral in a certain respect: If something turns into a group morality, it is obligatory, regardless of its content. Morality, however, is not content neutral: It is about justice and what we owe to others. Accordingly, these norms are not simply obligatory because they are the norms of “us,” but because of their specific content. There are some good reasons to see the set of possible moralities as being substantially constrained. According to the analysis above, this is particularly plausible if one distinguishes between a conventional morality and a reflective morality that is the product of critical scrutiny of moral precepts, ideologies and prejudices inherited from the past. The understanding of the second-personal morality of sympathy and fairness as a baseline of moral systems actually seems to point in the same direction, as it forms a critical yardstick for justified moral norms. However, second-personal morality is conceptualized in too constrained a manner, as we have seen, and its emergence is not sufficiently accounted for.

<sup>103</sup> At some stages of the theory, reference is made to something that resembles final cause arguments: In order to enable them to collaborate, the theory maintains, hominids evolved joint intentionality. In order to enable the cognitive coordination of their activities, *Homo sapiens* evolved new cognitive skills and became motivated by collective intentionality. Such statements do not sit well with evolutionary theory, because there is no teleology in evolution. Evolution (whatever its mechanisms may be) has no purpose, it simply has results, and human cognition, including its moral domain, may be one of these. Tomasello underlines this, too, cf. Tomasello, *Becoming Human*, 340. His arguments, however, do not seem to adhere always to this methodological stance.

Then there is the related problem of moral innovation: Many deviations from traditional morality came from individuals who relied on sources other than group morality. Socrates is a leading example of this, as are Jesus and Gautama Buddha, not to mention the many crucially important but forgotten moral innovators in human history. In the light of this, there are compelling reasons to conclude that the “objectivity” of morality is not derived from conformity with a group morality, but is a matter of critical thought and insight, often contrary to group perceptions. Moreover, the history of human morality is not one of intragroup harmony and intergroup strife: The moral battles within what may count as such a group (blurry as this concept is), which were often deadly, teach a different lesson. It is not enough to explaining this strife by referring to subgroups. What is central is a person’s capacity for autonomous moral reasoning that makes them more than a “cog in the conventional cultural practices,” empowering them to become the *subject of change*. The reference to the need for creative solutions to norm conflicts underlines this point but cannot easily be reconciled with the thrust of the argument emphasizing that objective morality is a group-minded system of rules. Where does moral creativity come from if morality is group based? What are the mental resources of critical moral thought?

The reflective morality that is important in this context is not just that of dyadic second-personal relations. Altruism and justice are not limited to such relations but include every moral patient. This leads us to the next point: The reference to in-group/out-group differentiations certainly captures one feature of human judgment and behavior. But it does not exhaust the principles of human moral judgment, because there are other patterns of judgment and behavior that clearly transcend and subvert this distinction. Again, the theory concedes this point – for example, by admitting that transcending the in-group/out-group distinction happens in dramatic cases such as slavery or apartheid.

It is important to ask: Why do arguments about the inclusion of human beings in the moral community work, as highlighted in the fight against slavery or apartheid? If there truly is a hardwired in-group morality, this success seems hard to understand. If “similarity” of appearance and manners counts, then how does this account for the power of arguments for including people in the group of those who deserve respect, arguments that appeal not to such conventional similarities but to the common humanity of all people? Why is it not only possible but morally compelling to look at others as more than “out-group barbarians”? The task is to explain the reasons for accepting the equal status of human beings as moral patients independently of in-group/out-group considerations and to understand what this tells us about the moral cognition of human beings.

Evolutionary mechanisms affect an individual’s phenotype. Arguments for group selection therefore need to be taken with great caution. Something like cultural group selection is even less plausible. The eradication of the Maya culture by the conquistadores is not an example of cultural group selection for better cooperators,

but of the victory of violence and greed using superior technological means over human beings who shared the same properties and possibilities for cooperation (on both moral and other terms) as the conquerors. The (at least temporary) influence of the human rights idea on human affairs is the product not of mechanisms discussed by adaptationist theories of cultural evolution,<sup>104</sup> but of dire historical experience and a normative idea with the power to convince.

A last exemplary point: As in the case of joint intentionality, collective intentionality seems to be the precondition of certain forms of culture, not their result.

### 7.6.9 *Is There an Alternative to Incrementalism?*

These criticisms show that the argument stalls on both the phylogenetic and the ontogenetic level: The mechanisms described by the incremental evolution theory cannot plausibly be taken to necessarily produce the cognitive structures that are taken to have developed in the species, and the ontogenetic account appears not to be sufficient to explain the emergence of the cognitive capacities acquired by individuals. The theory therefore in fact formulates a hypothesis about a successive number of mutations that lead to an incremental change in the cognitive abilities of hominids and modern humans, including the acquisition of joint intentionality, self-regulation, obligations mediated through an ought, self–other equivalence, group-mindedness, collective intentionality and so on that are the basis for human ontogeny. It implies that the proximate psychological mechanisms underlying these capabilities became biologically fixed. The ultimate mechanism underpinning the development is that hominids and later humans who cooperate in these increasingly advanced ways have “more babies” and consequently other, differently endowed individuals are “weeded out”<sup>105</sup> by natural selection. The fact that individuals with specific cooperative skills have more babies would not matter if certain proximate cognitive mechanisms were not fixed in the parent generation and bequeathed to their offspring. Otherwise, the next generation would have to learn these cooperative skills anew. The result would be a cultural, not a biological evolution. However, the incremental theory assumes precisely the latter. The account posits that the genetically inherited traits developed in this incremental fashion and proved to be of adaptive value (or to be at least neutral or not sufficiently maladaptive to reduce reproductive fitness), one after the other, and thus persisted in the evolutionary

<sup>104</sup> Such mechanisms are, for instance, that a trait is (randomly) adopted by successful and prestigious group members and then is widely adopted by others because it serves human needs, cf. e.g. Turchin, *Ultrasociety*. Such accounts do not include the decisive element of critical reflection and convictions based on reasons. Cf. for some critique Buchanan and Powell, *The Evolution of Moral Progress*, 396 ff. On the thesis of a “culture–gene coevolutionary process” as the key to understanding human cooperation, Henrich, *The Secret of Our Success*, 210, 319 f.

<sup>105</sup> Tomasello, *A Natural History of Human Morality*, 106.

process. Therefore, the newly acquired capabilities are passed on to the next generations. These next generations then move up the ladder, acquiring new, enhanced cognitive skills due to more mutations and passing them on to their offspring, until the level of proximate cognitive mechanisms enabling the formation of an “objective” morality of humans is reached. There is an important qualification: Some of the proximate cognitive mechanisms are understood as “spandrels,” cognitive structures that evolved because of certain nonmoral functions they serve but recruited to new tasks with potentially no (immediate) adaptive value.

All of this means that the baseline of the account is the assumption of consecutive accidental mutations that lead, step by step, to the set of cognitive capacities that ultimately formed the psychological mechanisms that enable human beings to judge and act morally.<sup>106</sup> Some crucial aspects of these capacities are understood as “spandrels,” the primary function of which was not moral. Joint intentionality, for instance, opened up a whole range of new possibilities of human life.<sup>107</sup>

As we have seen, the described development does not make unique evolutionary sense on the level of the evolution of the species, so the problem of lacking evidence for this particular trajectory cannot be circumvented.

Moreover, the assumption of a *gradual, incremental* evolution in small steps in hominids and *Homo sapiens* is not the only possibility. On the contrary, as we will see in greater detail when reviewing other evolutionary accounts, it is rather implausible given the evidence of the development of complex traits in various organisms and the paleoanthropological evidence about human development. It is in fact widely accepted that after very long periods of comparatively little cognitive development, as evidenced by the lack of major change in tools used, a qualitatively different kind of intelligence emerged with the cognitively modern *Homo sapiens*. Rather than incremental steps, this indicates a major evolutionary reorganization of the cognitive capacities of modern human beings. How and why this evidence fits into current evolutionary theory is something the following section will explore.

We can already learn something important from these interesting and differentiated arguments about the incremental evolutionary growth of human morality and their critique, however: The incremental succession of mutations, including the creation of spandrels, is not more probable than other accounts allowing for more far-reaching change, let alone the only evolutionary account that makes sense. This is even more so if one stays mindful of the fact that – as the theory underlines itself – there is no evidence that the described stages were those through which hominids in fact passed. It is possible that they are pure imagination.

The cognitive mechanisms enabling human morality, created by mutations that reorganized human cognitive abilities, were the *preconditions* (not the *consequences*)

<sup>106</sup> Tomasello, *Becoming Human*, 312 f.

<sup>107</sup> Tomasello, *Becoming Human*, 340, underlining that “the fact that a psychological adaptation is ‘aimed at’ a specific ecological challenge does not constrain its subsequent application.”

of qualitatively new forms of collaboration (not of collaboration as such) and other forms of human thought and action – for example, based on joint or collective intentionality, ideas of self–other equivalence or respect and the morality of altruism and justice that have structured human ways of living ever since. Given the “sparse evidence” of what happened in the history of the human species, *the question of the evolution of these cognitive capacities underlying morality thus continues to be an open one*. Acknowledging these limits of our current understanding is of fundamental importance for any evidence-based scientific endeavor.

These sobering findings have an important consequence: *So far, there is no reason whatsoever to assert that human beings, given what we know about the evolution of human moral cognition, can only have a small-group morality irreconcilable with the ethical principles embodied in human rights*. The simple reason is: We know so little about this process that there are no sufficient grounds to come to any such far-reaching conclusions.

The summarized theory therefore does not provide a fully convincing account of the evolution of morality. But what it does do, which is of substantial interest, is help to determine important elements of the set of psychological mechanisms that allow (among other things) for the specific forms of human cooperation, in particular on moral grounds. It helps to distinguish the cognitive apparatus of the nearest relatives of human beings from humans’ mental capacities. It has substantial things to say about the content of morality. There is no explicit reference to the problem of human rights, but it seems clear that rights play a role in the context of the discussion of notions such as “deservingness” or fairness. Morality’s contents of sympathy, fairness and – importantly – equal respect for everybody are in line with the human rights idea. One should take this seriously, at least as showing that contemporary evolutionary theory does not necessarily mean human rights skepticism. These findings are further underlined by other approaches of contemporary evolutionary theory.<sup>108</sup>

The theory of the incremental growth of human morality through interdependence makes another important point that shows us which direction to pursue further, given the preliminary results of our reflection: The reference to “spandrels” opens the door to evolutionary theories that take mechanisms of evolution operating in addition to natural selection seriously, among which nonadaptive structures such as spandrels play an important role, albeit not the only one. Such approaches help to paint a much more differentiated picture of the mechanisms of evolution than those which have been reviewed by us so far. These approaches form some of the most influential thought on evolution today and therefore merit closer investigation.

<sup>108</sup> Cf. e.g. the biocultural theory of inclusivist morality by Buchanan and Powell, *The Evolution of Moral Progress*.

## 7.7 EVOLUTIONARY PLURALISM

7.7.1 *The Contested Scope of Evolutionary Theory*

The review of some influential and paradigmatic approaches to the evolution of morality (which have some direct consequences for the understanding of the law) renders it crucial to answer the following questions: How much scope do evolutionary possibilities offer for the development of human beings' cognitive capacities? Can any well-defined constraints on the mental mechanisms possible (whatever they may be) be derived from evolutionary theory?

Thus far, our discussion of attempts at an evolutionary explanation of the idea of human rights, including the normative principles that underpin these rights, has provided no reason to assume that this idea is formed contrary to humans' cognitive nature – because the cognitive faculties underlying human morality are geared to produce a narrow tribal code of proper behavior, for instance. Nor are there reasons to assume that the idea of human rights is a misleading cognitive illusion produced by a hardwired cognitive gizmo. How do these results fit into the framework of the current understanding of evolution?

Evolutionary biology has developed since Darwin, and the transformations of evolutionary thinking in the *Modern Synthesis* have become a controversial field.<sup>109</sup> Nevertheless, current evolutionary biology appears to provide strong support for our analysis so far. To begin with, it is worth noting that in a considerable number of accounts of the evolution of morality, a set of assumptions play a guiding role: An organism is understood as an ensemble of singular traits. These traits are designed by natural selection for a specific fitness-enhancing function. Given the obvious fact of the interaction of traits, trade-offs are allowed, which are, however, understood as the best possible solutions in the face of the different competing functional demands expressed by natural selection. These are the tenets of what Stephen J. Gould and Richard Lewontin famously called the adaptationist program or the Panglossian paradigm.<sup>110</sup>

However, some very good arguments suggest that this picture of evolution is too simple. Evolution involves more than just gradual, simple lineage modification through natural selection. First of all, it is not so easy to say what a “trait” actually is, because organisms are “integrated entities, not collections of discrete objects” that allow for an atomization of properties.<sup>111</sup> In addition, the function that a trait possesses is not necessarily obvious. Take a well-known example: “Bones serve the function of providing rigidity to the body and attachments for muscles. But they also

<sup>109</sup> Cf. Robert Berwick and Noam Chomsky, *Why Only Us: Language and Evolution* (Cambridge, MA: MIT Press, 2016), 15.

<sup>110</sup> Stephen Jay Gould and Richard Lewontin, “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme,” *Proceedings of the Royal Society of London, Series B* 205, no. 1161 (1979): 585.

<sup>111</sup> Gould and Lewontin, “The Spandrels of San Marco and the Panglossian Paradigm,” 585.

are the sites for the storage of calcium, and the bone marrow is the tissue within which new red blood cells are produced. Depending on the causal pathway of interest, ‘bones’ are either macroscopic structural elements or collections of cells that secrete calcium or embryonic tissue of the circulatory system.”<sup>112</sup> Often, there is no evidence available that would allow us to choose between the possible options.<sup>113</sup>

Traits that have developed themselves in turn define the conditions of evolution: Without hands that are anatomically able to make tools, tool-making remains beyond an organism’s reach. In order to cultivate fire, an organism needs sufficient body mass to collect enough wood. This is also important for cognitive traits: “[T]he evolutionary questions about cognition are questions both of the evolution of cognition and the effects of cognition on evolution.”<sup>114</sup>

Furthermore, for theory-building it is important that – as Darwin himself underlined<sup>115</sup> – natural selection is certainly a central, but not the only causal factor of evolution. Many other factors need to be considered, as the example of spandrels already indicates. Evolution is, for instance, a stochastic, not a deterministic process. Very different factors that are unrelated to natural selection can influence the evolutionary pathway. Evolution plays out, for example, in finite, not infinite populations. Therefore, sampling effects in small populations – genetic drift – can shape the traits of a given population. Even without the influence of natural selection, because of stochastic inheritance patterns, some traits (including adaptive ones) may become extinct while others are passed on.<sup>116</sup>

<sup>112</sup> Richard Lewontin, *The Triple Helix: Gene, Organism, and Environment* (Cambridge, MA: Harvard University Press, 2000), 79.

<sup>113</sup> Richard Lewontin, “The Evolution of Cognition,” in *Methods, Models, and Conceptual Issues: An Invitation to Cognitive Science*, Vol. 4, eds. Don Scarborough, Saul Sternberg and Daniel Osherson (Cambridge, MA: MIT Press, 1998), 119: “The problem of origin is the problem of reconstructing the functions of traits in long-extinct environments together with their long extinct forms. While on purely mechanical grounds, we may exclude some explanations, we cannot choose among many allowable ones. Did the dinosaur stegosaurus use the large leaflike plates along its back for physical defense, for appearing deceptively large to potential predators, for sexual attraction, for thermoregulation, for all four, for some at one time and others at another, or none of the above? We will never know.”

<sup>114</sup> Lewontin, “Evolution of Cognition,” 113: “Eighteen-inch monkeys may remind us of humans and seem clever when we watch them in the zoo, but they are too short and too weak to raise a weight high enough and bring it down with enough force to break rocks, or to gather and process large chunks of fuel wood needed to maintain and control fire, and so they could never mine ore and smelt iron.” On ecological niches, Buchanan and Powell, *The Evolution of Moral Progress*, 367.

<sup>115</sup> Cf. Darwin’s well-known comment, Charles Darwin, *The Origin of Species* (London: John Murray, 1872), 421: “As my conclusions have lately been much misrepresented, and it has been stated that I attribute the modifications of species exclusively to natural selection, I may be permitted to remark that in the first edition of this work, and subsequently, I placed in a most conspicuous position – namely at the close of the Introduction – the following words: ‘I am convinced that natural selection has been the main, but not the exclusive means of modification.’ This has been of no avail. Great is the power of steady misrepresentation.”

<sup>116</sup> Gould and Lewontin, “The Spandrels of San Marco and the Panglossian Paradigm,” 585 ff.; Berwick and Chomsky, *Why Only Us*, 16 ff.

Furthermore, the problem of the stochastic “gravity well” has to be overcome.<sup>117</sup> Newly evolved adaptive traits may disappear again because there are only a small number of carriers of the trait that all become extinct before becoming sufficiently numerous to reproduce successfully – for example, because they all accidentally fall prey to some predator:

It is, for example, of little value to be the smartest member of your species, if, in an environment crawling with predators, you are also the slowest – or even just the most unfortunate. What’s more, in an indifferent world your reproductive success may not in the end have much to do with how magnificently you are adapted to any one thing. Whether or not that predator gets you, or whether or not you get the girl, may simply be a function of blind luck and circumstance.<sup>118</sup>

There are stochastic migration patterns that can lead to hybridization with existing organisms, with evolutionary effects.<sup>119</sup>

In addition, genes interact in complex ways – as epitomized by Darwin’s observation that blue-eyed cats are deaf. Genetic change may therefore have effects unrelated to natural selection. Natural selection operates on the whole organism – mutations leading to the genotype with the alleles XX may have different effects on fitness if combined with the alleles YY as compared to a genotype XX, ZZ.<sup>120</sup> Looking at the fitness of single traits therefore possibly misses important factors of development. Other conditions of development have to be considered, too: Fitness levels, for instance, can fluctuate if the frequency of the trait increases, as in the case of overpopulation.<sup>121</sup>

Nonadaptive mutations can persevere in the evolutionary process because of such influences. There are also nonadaptive side effects of adaptive properties that are established by natural selection. Architectural constraints are an important limit on the acquisition of new traits – not all developmental paths are even theoretically possible: “[T]here are no animals with wheels presumably because there is no way to make an appendage that rotates on an axle and still can be supplied with blood and nerves.”<sup>122</sup> Moreover, only a limited amount of what is theoretically possible has been explored by evolution.<sup>123</sup> Consequently, the traits of an organism can be determined by path-dependent phylogenetic trajectories that circumscribe the possible future development independently of natural selection and create the oft-observed “inertia” of evolution: “There appear, then, to be basic body plans that

<sup>117</sup> Berwick and Chomsky, *Why Only Us*, 22.

<sup>118</sup> Ian Tattersall, *Masters of the Planet: The Search for Our Human Origins* (New York: St. Martin’s Press, 2013), XIX f.

<sup>119</sup> Berwick and Chomsky, *Why Only Us*, 29.

<sup>120</sup> Cf. Lewontin, *The Triple Helix*, 82 ff., with the example of chromosomal polymorphisms in the grasshopper *Moraba scurra*; Berwick and Chomsky, *Why Only Us*, 23.

<sup>121</sup> Berwick and Chomsky, *Why Only Us*, 23.

<sup>122</sup> Lewontin, “Evolution of Cognition,” 114.

<sup>123</sup> Berwick and Chomsky, *Why Only Us*, 19 n. 9, 169 n. 9, 176.

are maintained through immensely long evolutionary periods despite dramatic changes in the life activity patterns of organisms and the functions of their parts.<sup>124</sup> There are, for instance, no vertebrates with six limbs or insects with eight, not six.<sup>125</sup> Or, to take another example: “[T]he contingent fact that we have five fingers and five toes may be better explained by an appeal to how toes and fingers develop than that five is optimal for their function.”<sup>126</sup> There is substantial debate about the reasons for the uniformity of certain common structures of organisms.<sup>127</sup> In addition, evidently, any development can only unfold within the framework of natural laws (physics, chemistry, etc.) and the physicochemical constraints that they impose.

Some of an organism’s adaptive, nonadaptive or neutral traits can be coopted for new functions – “exaptation” is another factor to be aware of in evolutionary theory.<sup>128</sup> This term clarifies an ambiguity of the term “adaptation,” which refers either to the genesis of a trait and its reason (features developed by natural selection for their present role) or to current features with fitness-enhancing utility irrespective of how they arose.<sup>129</sup> Darwin himself noted the existence of traits developed because of “laws of growth” but coopted for other (adaptive) purposes.<sup>130</sup> Mixing both dimensions might lead to “a common flaw in much evolutionary reasoning – the inference of historical genesis from current utility.”<sup>131</sup> One example of an exaptation that opened up striking new functional possibilities is the evolution of feathers. Finds such as gigantic feathered dinosaurs support the theory that feathers initially served thermoregulation and were only later coopted for flight. Other examples include the acquisition of limbs in a marine environment that later were used for movement on land, or the wings of insects.<sup>132</sup> After a trait is recruited for a new function, it may evolve further, becoming more adapted to this new function.<sup>133</sup> The cooptation of existing traits is of great importance for the evolutionary process:

<sup>124</sup> Lewontin, “Evolution of Cognition,” 117.

<sup>125</sup> Lewontin, “Evolution of Cognition,” 117.

<sup>126</sup> Berwick and Chomsky, *Why Only Us*, 60.

<sup>127</sup> Berwick and Chomsky, *Why Only Us*, 58 ff.

<sup>128</sup> Stephen Jay Gould and Elisabeth Vrba, “Exaptation – A Missing Term in the Science of Form,” *Paleobiology* 8, no. 1 (1982): 6.

<sup>129</sup> Gould and Vrba, “Exaptation,” 1 ff.

<sup>130</sup> Darwin, *Origin of Species*, 158: “The sutures in the skulls of young mammals have been advanced as a beautiful adaptation for aiding parturition, and no doubt they facilitate, or may be indispensable for this act; but as sutures occur in the skulls of young birds and reptiles, which have only to escape from a broken egg, we may infer that this structure has arisen from the laws of growth, and has been taken advantage of in the parturition of higher animals,” as quoted by Gould and Vrba, “Exaptation,” 5.

<sup>131</sup> Gould and Vrba, “Exaptation,” 14.

<sup>132</sup> On feathered dinosaurs, cf. Xing Xu et al., “A Gigantic Feathered Dinosaur from the Lower Cretaceous of China,” *Nature* 484 (2012): 94, noting the alternative possibility of feathers as a display structure, too. On the evolution of feathers and limbs, Gould and Vrba, “Exaptation,” 7 ff.; Tattersall, *Masters of the Planet*, 68, 210 f. On wings of insects, Lewontin, “Evolution of Cognition,” 119.

<sup>133</sup> Gould and Vrba, “Exaptation,” 11.

“[T]he enormous pool of nonadaptations must be the wellspring and reservoir of most evolutionary flexibility. We need to recognize the central role of ‘cooptability for fitness’ as the primary evolutionary significance of ubiquitous nonadaptation in organisms. In this sense, and at its level of the phenotype, this nona[da]ptive pool is an analog of mutation – a source of raw material for further selection.”<sup>134</sup>

We already encountered the idea of spandrels, which are unavoidable byproducts of architectural choices – for example, to construct a dome on rounded arches.<sup>135</sup> Given that they exist, they are put to good use for ornamental purposes and often are so accomplished that it seems as if they are built precisely for this decorative reason. Analogously, traits of an organism can serve a function even though they may have evolved not because of this function, but because of such architectural constraints: “One must not confuse the fact that a structure is used in some way . . . with the primary evolutionary reason for its existence and conformation.”<sup>136</sup>

It is even less admissible to go a step further and conclude from the assumption that selection could produce only one kind of trait in an organism that this organism in fact has this and no other trait. As explained above, one also encounters this functional fallacy in accounts about the evolution of morality that maintain that human beings have a naturally tribal, small-group morality – not because there is strong evidence that human beings indeed have this kind of morality, but because only such a morality is held to make evolutionary sense.

As indicated earlier, in Tomasello’s account of the evolution of morality we encountered a reference to spandrels at a crucial point of the argument: Proxy psychological mechanisms producing cognitive judgments of self–other equivalence are assumed to have been recruited for moral purposes – which only underlines the importance of these broader evolutionary perspectives.

The evolutionary processes captured by the idea of exaptation are of particular importance when considering the evolution of human beings’ higher mental faculties: As Gould and Vrba note on the debate between A. R. Wallace and Darwin on the evolutionary explanation of the development of the brain, Darwin recognized

that the brain, though undoubtedly built by selection for some complex set of functions, can, as a result of its intricate structure, work in an unlimited number of ways quite unrelated to the selective pressure that constructed it. Many of these ways might become important, if not indispensable, for future survival in later social contexts . . . . But current utility carries no automatic implication about historical origin. Most of what the brain now does to enhance our survival lies in the domain of exaptation – and does not allow us to make hypotheses about the selective paths of human history.<sup>137</sup>

<sup>134</sup> Gould and Vrba, “Exaptation,” 12.

<sup>135</sup> Gould and Lewontin, “The Spandrels of San Marco and the Panglossian Paradigm,” 581.

<sup>136</sup> Gould and Lewontin, “The Spandrels of San Marco and the Panglossian Paradigm,” 587.

<sup>137</sup> Gould and Vrba, “Exaptation,” 13. They conclude: “Thus, the two evolutionary phenomena that may have been most crucial to the development of complexity with consciousness on our

7.7.2 *Nature Does Not Make Leaps, Does It?*

A classic topic that already occupied early evolutionary theory is whether evolution is exclusively based on micromutations, as Darwin for instance assumed,<sup>138</sup> echoing the traditional metaphysical idea that nature does not make leaps (*natura non facit saltum*). There is conclusive evidence to show that this is an overly narrow conception of what is evolutionarily possible. There are examples demonstrating that rapid change is possible, leading to a very fast divergence of closely related forms, arguably based on small genetic changes – in particular in regulatory genes – that have far-reaching effects, including the classic case of the stickleback, a small fish with spines, or the evolution of the basic elements of eyes (light-sensitive cells, pigment cells), among others.<sup>139</sup> Thus, the possible evolutionary trajectory of a species needs to include more than just evolutionary accounts of incremental change.

Furthermore, the evolutionary theory of human cognition faces particular challenges. One such challenge we have already encountered: the lack of evidence for the existence of particular cognitive capabilities of early human beings. If certain kinds of artifacts are found, say a tool or a flute, we can be sure that the cognitive abilities necessary for making this kind of tool or this musical instrument must have been in place. So, too, must other capacities – for example, those making it possible to identify the need for tools, or the capacity for purposive functional inventions, for using tools or for producing and enjoying music. But how are we to know whether a group of modern humans sang a song while making the tools or the flute? What were their social relations like? How did they communicate? Did the members of this species tell tales while sitting at campfires? Did they use arithmetic to keep track of the quantity of their provisions? Did they possess a sense of justice that guided them when dividing up prey? Such cognitive activities do not leave any traces behind. So how are we to know?

planet (if readers will pardon some dripping anthropocentrism for the moment) – the process of creating genetic redundancy in the first place, and the myriad and inescapable consequences of building any computing device as complex as the human brain – may both represent exaptations that began as nonaptations, the concept previously missing in our evolutionary terminology.”

<sup>138</sup> Darwin, *Origin of Species*, 156: “For natural selection can act only by taking advantage of slight successive variations, she can never take a leap, but must advance by the shortest and slowest steps.”

<sup>139</sup> Cf. Tattersall, *Masters of the Planet*, 94 ff., including the example of the development of *Homo ergaster*, as shown by the Turkana Boy – another example that genetic modification involving a radical change of morphology does not need to lead to “hopeful monsters,” derided in evolutionary theory: “Perhaps the Turkana Boy’s radically new bodily confirmation can be attributed to a genetic event of similar kind. A minor mutation had occurred in the Boy’s lineage that, through altering gene timing and expression, had radically changed its possessor’s morphology – and had, entirely accidentally, opened new adaptive avenues to them. . . . Something routine and unremarkable on the genomic level had occurred among the Boy’s precursors; and it just happened to change the course of hominid history,” *ibid.* 98; Berwick and Chomsky, *Why Only Us*, 2, 5, 26 ff., 31 ff., 67 ff.

Giving an evolutionary description of the developmental process of human cognition is equally difficult. It is far from clear who the predecessors of modern humans are, as the paleoanthropological record contains insufficient evidence to determine the lines of relationship between species and distinguish ancestral lines from those forms that are not ancestors.<sup>140</sup> It is equally unclear what cognitive capabilities any of the species that may be part of the lineage of modern human beings possessed, as here, too, the problem of evidence for such capabilities rears its head. The split between Neanderthals and modern humans happened 400,000–600,000 years ago. Tool-making was typical of Neanderthals.<sup>141</sup> Recent research on Neanderthal artwork indicates symbolic thought on the part of this certainly cognitively gifted species, but the evidence is controversial and very sparse. And, whatever the results of future research may be on Neanderthal symbolic behavior, there is no evidence of any systematic use of symbolic thought – in contrast to modern humans.<sup>142</sup> We have already observed that the closest living relatives of modern humans are primates, chimpanzees and bonobos, about 14 million years of evolutionary time away from modern humans – a substantial amount of time in which to develop differently. As animals adapted in crucial aspects to life in forests, and in some cases adapted behaviorally to life in the savannah,<sup>143</sup> there already is a difference from early hominids, who – enabled by their new body form – moved out of forests and into other environments and modes of life.<sup>144</sup> The number of existing “close” relatives is sparse, which makes it difficult to determine a trait’s successive changes: “The evolutionary space is too sparsely populated to be able to connect the points sensibly.”<sup>145</sup> In addition, the paleoanthropological evidence suggests that early hominids developed cognitive skills that are not available to primates living today: “The evidence of tool use, and yet more of tool-making, tells us that the bipedal apes had graduated – perhaps as much as 3.4 million years ago, and at least before 2.6 million years ago – to a cognitive state that lay beyond anything we can infer for the apes as we know them today.”<sup>146</sup> Tattersall argues that early hominids’ “entirely new and radical way of interacting with the world around them” had its roots in an early “leap of nature” and later exaptation of cognitive

<sup>140</sup> This is a problem that applies to the emergence of the genus *Homo*, too: “To put the situation in a nutshell, there is not one fossil among all those known in the period before two million years ago that presents itself as a compelling candidate for the position of direct progenitor of the new hominids to come,” Tattersall, *Masters of the Planet*, 85. Lewontin, “Evolution of Cognition,” 115, 118.

<sup>141</sup> Tattersall, *Masters of the Planet*, 175 ff.

<sup>142</sup> Ian Tattersall, “The Minimalist Program and the Origin of Language: A View from Paleoanthropology,” *Frontiers in Psychology* 10, no. 677 (2019); Tattersall, *Masters of the Planet*, 180.

<sup>143</sup> Tattersall, *Masters of the Planet*, 70.

<sup>144</sup> Tattersall, *Masters of the Planet*, 57, 70.

<sup>145</sup> Lewontin, “Evolution of Cognition,” 122.

<sup>146</sup> Tattersall, *Masters of the Planet*, 49 f. Cf. e.g. throwing with precision, understanding properties of materials or tool-making, *ibid.* 53 f. On the latter cf. n. 160. On vocal skills, *ibid.* 61.

skills. The best explanation is that “the cognitive potential to make stone tools was born in the large genetic alteration that must have been involved in the acquisition of the new and radically different bipedal body form; and that this potential lay dormant for some time before being expressed in the invention of stone tool making.”<sup>147</sup> The fact that evolution is not a continuous, linear, incremental process creates specific problems for the understanding of the evolution of cognition, because it is possible for closely related forms to diverge very rapidly.<sup>148</sup> The similarity of traits can vary from trait to trait – an organism may be similar in some respects but quite different in others.<sup>149</sup> And there is another crucially important possibility, namely that the trait “simply does not exist in some or all related lines, that it is a novelty, and so has no observable evolutionary history.”<sup>150</sup>

Furthermore, there is the functional change of traits – less closely related organisms may share more properties with certain organisms than with others that are more closely related, as illustrated by an evolutionarily convergent system such as vocal learning in songbirds and humans, which are separated by millions of years of evolutionary time.<sup>151</sup> There is a need to differentiate between homologous and analogous structures and to identify homologies, especially because homologous structures may change functionally and because of the abovementioned possibility of novelties without homologies in other species.<sup>152</sup> In any case, as in other organisms, it is not singular traits that determined the survival of modern humans but the human organism as a whole, with its many specific adaptive and not-so-adaptive features and their complex interactions,<sup>153</sup> which renders an account of the evolutionary history of any one element of this integrated organism even more difficult.

All of this arguably leads to the conclusion that there is simply not enough evidence to ever decisively settle the question of how exactly modern humans’ cognitive capacities came into being over the course of evolution.<sup>154</sup>

It goes without saying that any evolutionary theory of the development of human cognition has to take into account the existing evidence of human cognitive development. There is widespread consent, given the current stage of knowledge, that around 200,000 years ago anatomically modern humans were living in Africa.<sup>155</sup> About 80,000 years before the present day, there is uncontested evidence of symbolic behavior such as beads, shell ornaments and geometric engravings, though there are some earlier artifacts.<sup>156</sup> This indicates that at least from this time onwards, modern

<sup>147</sup> Tattersall, *Masters of the Planet*, 67 f.

<sup>148</sup> Lewontin, “Evolution of Cognition,” 116 with some examples of such rapid change.

<sup>149</sup> Lewontin, “Evolution of Cognition,” 117.

<sup>150</sup> Lewontin, “Evolution of Cognition,” 118.

<sup>151</sup> Berwick and Chomsky, *Why Only Us*, 12 ff., 140 ff.

<sup>152</sup> Lewontin, “Evolution of Cognition,” 124 ff.

<sup>153</sup> For some examples cf. Tattersall, *Masters of the Planet*, 212.

<sup>154</sup> This is the famous conclusion in Lewontin, “Evolution of Cognition,” 108.

<sup>155</sup> Tattersall, *Masters of the Planet*, 184 ff.

<sup>156</sup> Tattersall, *Masters of the Planet*, 142, on earlier objects of a contested nature, including the “Venus” of Berekhat Ram, concluding that there is “nothing in the material record to suggest

humans had the cognitive capacity for symbolic thought and used it systematically. Other early indicators for new cognitive abilities are bladelets (stone flakes sunk into handles), making tools from artificially hardened materials, pressure flaking, barbed harpoons and, on the social level, the functional division of living spaces.<sup>157</sup>

The capacity for symbolic thought is plausibly taken as an indicator that cognitively modern humans had arrived on the scene: With modern humans, an unprecedented and rapid development set in, with cultural and technological innovations that transformed the world – and by now are threatening the continued existence of this simultaneously creative and destructive species. This development is dramatic proof that these beings possessed some revolutionary new creative abilities in both thought and action. Before that, over the millions of years of development there had been many striking innovations like tool-making, the use of fire, spears, composite tools, “prepared core” implements and blades.<sup>158</sup> However, there was a disconnect between technological innovation and the development of a new species of the *Homo* genus.<sup>159</sup> Technological innovations persisted unchanged for hundreds of thousands of years and were of very limited scope in comparison to what happened later, although they were very remarkable innovations in themselves, beyond the cognitive capacities and other abilities of great apes:<sup>160</sup> Mode 1/Oldowan tools (sharp stone flakes) were invented 2.5 million years before the present and mode 2/Acheulean hand axes 1.5 million years before the present, although they were widely applied only hundreds of thousands of years later. These tools continued to be used even though new species like *Homo ergaster* developed during this time.<sup>161</sup> There is consequently no necessary connection between the appearance of a new kind of hominid and technological innovation, nor is there evidence of any gradualism in the appearance of technologies for long stretches of time.

The cognitive abilities evident in Acheulean hand axes are already remarkable – foresight and the planning of complex sequences of actions, the knowledge of materials, anticipating need by keeping stocks of suitable stone, among others. Hand axes even indicate economic and social specialization.<sup>162</sup> Later developments furnish further examples of the advanced cognitive skills of the various branches of hominids. But with *Homo sapiens*, change starts to take place at breathtaking speed. In just 80,000 years the species traveled from hand axes that had already been used

that the symbolic manipulation of information was in any way a regular part of the cognitive repertoire of *Homo heidelbergensis*. Had it been, we would surely expect to see more material evidence of it.”

<sup>157</sup> Tattersall, *Masters of the Planet*, 200 ff.

<sup>158</sup> Tattersall, *Masters of the Planet*, 112, 138 ff.

<sup>159</sup> Tattersall, *Masters of the Planet*, 62.

<sup>160</sup> Cf. e.g. on chimpanzees’ inability to produce such hand axes, arguably not only because of their cognitive abilities, but also because of the form of their hands, which are ill formed for this purpose, although chimps are able to learn to use existing hand axes and to produce other tools, Tattersall, *Masters of the Planet*, 43 ff., 49 ff. (on the spear-making of chimpanzees).

<sup>161</sup> Tattersall, *Masters of the Planet*, 42 ff., 103, 116 f., 124 ff., 138, 187.

<sup>162</sup> Tattersall, *Masters of the Planet*, 126 ff.

in different forms for over 2 million years to the technical civilization of today. This change cannot but be based on cognitive abilities of modern human beings that are qualitatively new in comparison to those of their predecessors, abilities that have increasingly been put to use.<sup>163</sup> Symbolic thinking and action is one such ability, in itself of far-reaching importance for culture-building and science. Technical skills also point clearly to certain abilities – for example, the analysis of complex causal chains as a precondition for developing new technologies in the making of tools and weapons. The existence of art is an indicator not only of symbolic thought, but also of a very rich inner life of early humans, including the search for meaning and beauty – and this tells us a lot about the species' cognitive capacities. In the case of other cognitive abilities, there is only little evidence until much later in history. Theory-building therefore turns to the second-best option: interpreting the emergence of symbolic thought and action as a proxy for the existence of other cognitive abilities that are crucial to being human but leave no material evidence behind.

Language is at the forefront of many debates, as an instrument of complex thought and communication,<sup>164</sup> but any other aspect of the human mental world qualifies as well, of course. One example is human moral understanding and judgment. There are some indications of forms of care for others in hominid history: One example is a 1.8-million-year-old skull of an elderly male who already lost his teeth during his lifetime but survived nevertheless. This seems to indicate help from members of the group to which he belonged, although other possible explanations also exist. If help was indeed provided, the skull “furnishes us with the first putative instance of social concern in the hominid record.”<sup>165</sup> Other traces are (different forms of) cannibalism or burial practices that indicate concern (or lack thereof) for species members.<sup>166</sup>

Beyond this, as for other possible mental abilities, there is no direct evidence of the exact kind of cognitive capacities existing in hominid history that constituted the foundation for the moral world in which modern humans lived, answering the question of which concepts and principles they applied, what moral feelings they entertained and how these were spelled out in social practices and institutions. Did modern humans have a sense of fairness? Did they experience shame? Did they feel bound by obligations?

If one accepts symbolic thought as a proxy for a fully developed modern human cognition because of the reasons outlined, one has to conclude that early modern humans possessed the same cognitive abilities to form moral ideas as contemporary

<sup>163</sup> Tattersall, *Masters of the Planet*, 205.

<sup>164</sup> Cf. Tattersall, *Masters of the Planet*, 183 (Neanderthals: no language), 214 ff.; Berwick and Chomsky, *Why Only Us*, 1 ff.

<sup>165</sup> Tattersall, *Masters of the Planet*, 124. He asks whether such attitudes may have their roots in forms of empathic behavior in chimpanzees towards wounded or oppressed group mates, despite the chimpanzees lacking the technical abilities to provide help. For the example of a disabled aged male Neanderthal, *ibid.* 171.

<sup>166</sup> Cf. Tattersall, *Masters of the Planet*, 152 f. (*Homo antecessor*), 172 ff. (Neanderthal “survival cannibalism” and burial).

humans (whatever these abilities may turn out to be and to whatever use they may have been put).

When considering these problems, it is useful to distinguish between the evolution of cognitive capacities (human beings have the cognitive capacity to produce symbolic artifacts, primates do not) and the development of skills and insights within the framework of these evolved cognitive abilities. This is particularly important when considering the human species. Evidently, the great historical change that gave rise to the transformative creative innovations highlighted above is a central element of the history of modern humans. There is, however, no evidence of any evolution of modern humans' cognitive capacities themselves that underpins this striking history of change and innovation. Therefore, it comes as no surprise that one could transfer a child born in a contemporary tribal society whose way of life is similar to a Stone Age culture to Zürich, New York or Beijing and know that it will develop in the same way as any other kid in the world, becoming unhealthily attached to its smartphone at far too early an age.

It is consequently reasonable to conclude that modern humans' basic abilities in the cognitive domain have stayed unchanged since their appearance at least 80,000 years ago. If this is indeed the case, we can project backwards whatever we know about the structure of the mind of humans living today: If human beings today are distinguished by a general learning ability, then it is plausible to assume that this formed the basis of their early development as well. If they possess the ability for shared intentionality, then plausibly early modern humans had this capacity, too. If there is evidence for other additional or alternative cognitive abilities – say, a language faculty – then most probably language was among the cognitive tools with which modern humans started their journey. Accordingly, the current cognitive capacities of modern humans can be taken as a clue to the mental parameters of life in the remote past.

One should note that making this assumption is justified because it is the most plausible interpretation of what we know at present. Nevertheless, within this general framework, much is unclear. Regarding language, for instance, one prominent account in fact speculates about an evolutionary pathway along these lines: First the semantic–conceptual system of language evolved, then the systems of the externalization of language through speech and other means such as signs.<sup>167</sup> This is relevant for the assessment of the cognitive abilities of all hominids, including Neanderthals, for instance: “It is merely assertion that complex stone working, fire control, clothing, ochre, and the like require language. We may have them all, but that does not mean Neanderthals had to have all the features that co-occur in us just because they had some of them.”<sup>168</sup> This cautionary observation underlines that confident assertions about the precise path of the evolution of human cognition are simply not possible.

<sup>167</sup> Berwick and Chomsky, *Why Only Us*, 80, 87. For comments e.g. Tattersall, “The Minimalist Program and the Origin of Language.”

<sup>168</sup> Berwick and Chomsky, *Why Only Us*, 153.

There is another problem: For many such cognitive traits, we do not know what kind of heritable variations existed, and it is difficult to determine their reproductive function for the first individuals possessing the trait within a species, which is decisive for these traits to become established.<sup>169</sup> Nor do we know whether and how these cognitive traits led to the replacement of other species. Take language, a central topic of evolutionary research. Despite the large number of contributions to the theory of the evolution of language, the function of language is less clear than it may appear.<sup>170</sup> A cofounder of evolutionary theory, Wallace, famously wondered what function human language may serve, because he saw no biological function that could not be met by other, simpler means than human language<sup>171</sup> – a problem still relevant today (although Wallace’s proposed solution, divine intervention, may now be less accepted).<sup>172</sup> The proposals entertained today range widely, including language as a tool of successful cheating, of sexual selection, of communication for the purpose of cooperation or of structured thought, in the latter case a tool that evidently is still very useful.<sup>173</sup> Moreover, it is impossible to determine whether the function of such cognitive traits had any substantial effect on reproductive success: “[W]e cannot measure the survival advantage, if any, in our remote ancestors of the ability to do arithmetic.”<sup>174</sup> Thus, one cannot determine whether any cognitive trait that might have had an effect on reproductive rates *did in fact* have an effect on reproductive rates.<sup>175</sup>

Such problems need to be considered for morality, too. Experiencing altruistic obligations in a group with members devoid of such inclinations may simply make the respective individual easy prey. The functional value of morality for collaboration is not obvious either, as, for instance, it precludes effective forms of

<sup>169</sup> Lewontin, “Evolution of Cognition,” 112: “Evolution by natural selection occurs when individuals within a species possess a trait that gives them a reproductive or survival advantage over others within the species that lack the trait. It is an explanation of how a new trait spreads *within* a species, not how the species may replace other species once the trait has been incorporated. . . . Thus a species that possesses linguistic competence may indeed take over the earth as a consequence of the technological and managerial capabilities that are the result of language, but in a species lacking linguistic competence, the rudimentary ability to form linguistic elements by a few individuals may be taken as a sign of difference that causes them to be expelled or even killed” (emphasis in original).

<sup>170</sup> Berwick and Chomsky, *Why Only Us*, 65 ff.

<sup>171</sup> Berwick and Chomsky, *Why Only Us*, 3.

<sup>172</sup> Berwick and Chomsky, *Why Only Us*, 80 ff. with a review of the debate including comments from Nobel Laureate Salvador Luria and François Jacob.

<sup>173</sup> Berwick and Chomsky, *Why Only Us*, 80 ff.; Tecumseh Fitch, *The Evolution of Language* (Cambridge: Cambridge University Press, 2010), 297 ff.; Szabolcs Számadó and Eörs Szathmáry, “Selective Scenarios for the Emergence of Natural Language,” *Trends in Ecology & Evolution* 21, no. 10 (2006): 555 ff.

<sup>174</sup> Lewontin, “Evolution of Cognition,” 111.

<sup>175</sup> Lewontin, “Evolution of Cognition,” 120. To ascertain this, one needs to find: first, contrasting groups, one that possesses the trait and one that does not; second, the differences in reproductive rates need to be substantial enough to be measured; and third, there must be evidence that a trait is genetically inherited, *ibid.*

collaboration such as anthill- or beehive-like human societies. Collaboration for survival is not only possible on moral terms.

In this context, too, we should stay mindful of the observation that just because a trait is used in certain ways does not mean that it evolved because of this function. The fact that morality makes possible certain sophisticated forms of collaboration based on solidarity and respect does not mean that it evolved because of this function. The mental abilities creating the possibility of moral thought might have any of the evolutionary origins mentioned and form, for instance, a side effect of an (overall) adaptive cognitive reorganization or an example of exaptation, a spandrel perhaps, which turned out to be a major influence on the modern human life form.

There is no reason to assume that any account of human moral cognition has to satisfy the constraints of the adaptationist paradigm, namely to be favored by natural selection as an isolated trait within the framework of functional trade-offs between different traits. It is entirely possible from the point of view of evolutionary theory that various elements of human cognition are structures that evolved without any clear adaptive advantage but still define the kind of creatures we are.

On all accounts, the paleoanthropological timeline quite clearly indicates a case of rapid evolutionary development that led to a fundamental cognitive reorganization of human beings – albeit perhaps caused by minor genetic change, probably of regulatory genes, which had far-reaching consequences. In comparison to the 1 million years that passed between the invention of Oldowan stone tools and the designing of Acheulean hand axes as a major technological innovation, the development of our current civilization in a mere 80,000 years has been fast indeed. As this rapid change *did* happen, there is a pretty clear answer to the question of whether it *could* have happened.<sup>176</sup> With modern human beings, nature simply *did* make a leap.<sup>177</sup>

<sup>176</sup> Tattersall, *Masters of the Planet*, 63: “The extraordinary human cognitive style is the product of a long biological history. From a non-symbolic, non-linguistic ancestor (itself the outcome of an enormously extended and eventful evolutionary process), there emerged our own unprecedented symbolic and linguistic species, an entity possessing a fully-fledged and entirely individuated consciousness of itself. This emergence was a singular event, one that involved bridging a profound cognitive discontinuity. For there is a *qualitative* difference here; and, based on any reasonable prediction from what preceded us, the only reason for believing that this gulf *could* be ever have been bridged, is that it was” (emphasis in original). The Pleistocene offered arguably favorable conditions for “the local fixation of genetic novelties and for speciation. Both of these are processes that in creatures such as hominids depend on physical isolation, and small population sizes,” both features of the living conditions of hominids of that era, *ibid.* 149. On genetic evidence about a small founding African population, *ibid.* 194. Cf. Berwick and Chomsky, *Why Only Us*, 37. The current situation may have created reverse conditions: “Modern human populations have simply become too large and dense to witness the fixation of any significant genetic novelties that might in theory make us smarter and more protective of our own long-term interests,” Tattersall, *Masters of the Planet*, 231.

<sup>177</sup> Tattersall, *Masters of the Planet*, 207: “[W]e evidently came by our unusual anatomical structure and capacities very recently: there is certainly no evidence to support the notion that we gradually became who we inherently are over an extended period, in either the physical or

The conclusion to be drawn regarding the evolution of whatever psychological capacities enable human morality is, then: As things stand, there are good reasons to believe that these cognitive capacities were part of the set of faculties with which modern humans embarked on their journey at least 80,000 years ago – faculties that define modern humans and have made their journey since then *human* history.

### 7.8 THE EVOLUTIONARY POSSIBILITY OF HUMAN GOODNESS

This leads to the crucial lesson for the cognitive interests of this inquiry: Evolutionary theory provides no discernible argument as to why the cognitive faculties yielding both the idea of human rights and the identified justificatory principles of this idea could *not* be the products of evolutionary processes. It may have a “wonderful reductionist appeal” to think that the cognitive abilities of modern humans are determined by the living conditions of early modern humans<sup>178</sup> – for example, yielding a “small-group morality.” But this reductionism does not do justice to an evolutionary theory aware of the complexity of the human mind. Striking things evolved, such as the faculty to create a phenomenon as breathtaking as art,<sup>179</sup> and there is no reason to think that the ability for moral reasoning, motivation and emotion may not be a further example of an equally striking quality. There are no compelling grounds to restrict what is evolutionarily possible to such a “small-group morality” or any other form of psychological mechanisms that make the idea of human rights alien to human cognition.<sup>180</sup> Nothing in what we know about the history of the evolution of cognition, nor anything in evolutionary theory speaks against the possibility that human beings possess a sense of justice, evaluate genuine care and respect for others as morally good, regard just and morally good acts as obligatory and embed moral judgments in a rich, sometimes painful, sometimes ravishing world of moral emotions.

the intellectual sense. . . . [T]his suggests that the physical origin of our species lay in a short-term event of major developmental reorganization, even if that event was likely driven by a rather minor structural innovation on the DNA level.”

<sup>178</sup> Tattersall, *Masters of the Planet*, 228.

<sup>179</sup> “Decorating the dank and dangerous depth of caves with fabulous animal images and a whole vocabulary of geometrical symbols is, to put it mildly, a rather unusual pursuit,” Tattersall, *Masters of the Planet*, 205.

<sup>180</sup> Buchanan and Powell’s thesis, Buchanan and Powell, *The Evolution of Moral Progress*, that under out-group pressure an evolutionarily fixed, exclusivist, in-group morality thrives is thus not convincing. The supplementary thesis that (only) under favorable conditions without out-group threats and with an abundance of resources can the “luxury good” of inclusivist morality develop is not convincing either. There is a clear disconnect in social history between inclusivist morality and such socioeconomic circumstance. The human rights idea has convinced people under very dire conditions and motivated them, among others, to transform societies accordingly. One should not mistake favorable conditions for the social institutionalization of a normative idea for conditions necessary for individual humans to develop this idea.

The answer to our initial question on the origin of the cognitive capacities enabling the moral world of human beings is thus: There are many riddles hidden under the human skin, as Ahab despairingly realizes, and the attempt to hunt down the elusive essence of human existence can even lead to a self-destructive chase. But this much is clear: The scope of evolution is certainly wide enough to include cognitive mechanisms (whatever they turn out to be) that enable human beings to form notions of justice, concern and respect for others and, after a long development, ultimately articulate the remarkable idea of human rights.