

CHAPTER 1

~ HUMAN KIND ~

We human beings are animals, governed by the laws of biology. Our life and death are biological processes, of a kind that we witness in other animals too. We have biological needs and are influenced and constrained by genes with their own reproductive imperative. And this genetic imperative manifests itself in our emotional life, in ways that remind us of the body and its power over us.

For centuries poets and philosophers have told stories about erotic love—Plato leading the way. These stories have endowed the object of love with a value, a mystery, and a metaphysical distinction that seem to place it outside the natural order. And in these stories biology seems hardly to figure, even though they are stories that would make little sense were it not for our condition as reproductive animals, who have established their niche by sexual selection.

We are territorial creatures, just like chimpanzees, wolves, and tigers. We claim our territory and fight for it, and our genes, which require just such an exclusive

claim over habitat if their replication is to be guaranteed, depend upon our success. Yet when we fight it is, as a rule, in the name of some high ideal: justice, liberation, national sovereignty, even God Himself. Once again, it seems that we are in the habit of telling ourselves stories that make no reference to the biological realities in which they are rooted.

The most noble of human attributes also have their biological underpinning—or so it seems at least. The self-sacrifice that causes a woman to lay everything aside for her children, the courage that enables human beings to endure the greatest hardships and dangers for the sake of something that they value, even those virtues such as temperance and justice that seem to require us to vanquish our own desires—all these things have seemed to many people to have their counterparts among the lower animals and to demand a single explanation, generalizable across species. Personal affection has been brought within the fold of biology, first by Freud’s highly metaphorical and now largely discredited theory of the libido and more recently by the attachment theory of John Bowlby, for whom love, loss, and mourning are to be explained, at least in part, as phylogenetic products of our need for a “secure base.”¹ Bowlby was a psychiatrist, acutely aware that human beings do not merely inherit their emotional capacities but also adapt and refine them.

¹John Bowlby, *Attachment and Loss*, vols. 1–3 (New York: Basic Books, 1969–1980); John Bowlby, *A Secure Base* (New York: Routledge, 1988).

Nevertheless, he described love, grief, and mourning as biological processes and argued that “the child’s tie to his mother is the human version of behaviour seen commonly in many other species of animal.”²

By putting that behavior in its ethological context Bowlby was able to give a far more plausible account of our primary attachments than those given by Freud and his immediate successors. Our personal affections, he argued, are to be explained in terms of the function that they perform in our “environment of evolutionary adaptedness,” and the explanation will not be couched in terms that make any radical ontological division between us and other mammals. The discovery of the hormone oxytocin, and its effect in predisposing animals of many different species toward affectionate relations with their own kind, has further encouraged the view that attachment can be understood and explained without reference to the stories with which we humans embellish it.

When Darwin and Wallace first hit on the idea of natural selection, the question arose whether our many “higher” characteristics, such as morality, self-consciousness, symbolism, art, and the interpersonal emotions, create such a gap between us and the “lower” animals as to demand explanation of another kind. Wallace at first thought that they did not but later changed his mind, coming to the conclusion that there is a qualitative leap in the order of things,

²Bowlby, *Attachment and Loss*, vol. 1, p. 183.

setting the higher faculties of human kind in a different category from those features that we share with our evolutionary neighbors. As he put it: “We are endowed with intellectual and moral powers superfluous to evolutionary requirements,”³ and the existence of these powers could therefore not be explained by natural selection for fitness.

Darwin, however, remained attached to the view that *natura non facit saltus* and in writing *The Descent of Man* tried to show that the differences between humans and other animals, great though they are, can nevertheless be reconciled with the theory of stepwise development.⁴ For Darwin the moral sense is continuous with the social instincts of other species.⁵ Through the theory of sexual selection, he gave an enhanced account of the resources on which natural selection can draw and made the suggestion, taken up in our own time by Steven Pinker and Geoffrey Miller, that many of the “higher” faculties of man, such as art and music, which seem, on the face of it, to be devoid of any evolutionary function, should be seen as resulting from selection at the sexual level.⁶ Darwin went

³A. R. Wallace, *Natural Selection and Tropical Nature: Essays on Descriptive and Theoretical Biology* (London: Macmillan, 1891). See also A. R. Wallace, *Darwinism: An Exposition of the Theory of Natural Selection with Some of Its Applications* (London: Macmillan, 1889), chapter 15.

⁴Charles Darwin, *The Descent of Man*, vol. 1 (New York: Appleton and Co., 1871).

⁵*Ibid.*, pp. 71–72.

⁶Steven Pinker, *How the Mind Works* (London: Allen Lane, 1997), pp. 522–524; Geoffrey Miller, *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature* (New York: Doubleday, 2000).

on to give an account of human emotions in which their expression in the face and gestures is compared with the expression of emotion in other animals: and his purpose in all this was to suggest that the perceived gap between us and our evolutionary cousins is no proof of a separate origin.⁷

GENETICS AND GAMES

This controversy has taken on an entirely different character since R. A. Fisher's pioneering work in population genetics.⁸ Problems with which Darwin wrestled throughout his life—the sexual selection of dysfunctional features (the problem of the peacock's tail), for example, or the “altruism” of insects (the problem of the anthill)—are radically transformed when the locus of evolution is seen as the self-replicating gene, rather than the sexually reproducing animal.⁹ And as John Maynard Smith and G. R. Price showed in an elegant essay,¹⁰ the new way of looking at natural selection, as governed by the replicating “strategies” of genes, permits the application of game

⁷ Charles Darwin, *The Expression of the Emotions in Man and Animals* (New York: Appleton and Co., 1898).

⁸ R. A. Fisher, *The Genetical Theory of Natural Selection* (1930), revised ed. (New York: Dover, 1958).

⁹ See the lively account in Helen Cronin, *The Ant and the Peacock: Altruism and Sexual Selection, from Darwin to Today* (Cambridge: Cambridge University Press, 1991).

¹⁰ J. Maynard Smith and G. R. Price, “The Logic of Animal Conflict,” *Nature* 246 (1973): pp. 15–18.

theory to genetic competition, which in turn delivers a neat solution to another famous problem—that of aggression, noticed by Darwin and spelled out in detail by Lorenz.¹¹ The rut among stags can be derived as an “evolutionally stable strategy”: one that enables the genes of rutting stags to reproduce themselves while providing the genes of hinds with the best return for their reproductive investment. This approach, generalized by R. Axelrod,¹² has had profound consequences, for example, in showing that there might be an evolutionary advantage in reciprocally altruistic cooperation, even when not part of kin selection (as when female bats allegedly share their booty of blood with other unsuccessful females in a colony). It has also suggested a general theory of “altruism,” held by its supporters to explain not only the inflexible self-sacrifice of the soldier ant but also the fear-filled and heroic self-sacrifice of the human soldier.¹³ In short, we seem to have been brought a step nearer the proof of Darwin’s contention that the moral sense is continuous with the social instincts of other species.

¹¹Konrad Lorenz, *On Aggression*, trans. Marjorie Kerr Wilson (New York: Harcourt Brace, 1966).

¹²R. Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

¹³See, for example, Matt Ridley, *The Origins of Virtue: Human Instincts and the Evolution of Cooperation* (London: Viking, 1991). It is important to recognize that the game-theoretic approach to altruism is distinct from the theory of “inclusive fitness,” defended in W. D. Hamilton, “The Genetical Evolution of Social Behaviour,” *Journal of Theoretical Biology* 7 (1964): pp. 1–16, according to which altruism extends to kin and in proportion to the degree of kinship.

The genetic approach has not been without its critics. Advocates of “group selection” have argued that selection must occur at higher levels than that of the gene if we are to account for such socially complex behavior as the self-limitation of populations and the dispersal patterns of herds.¹⁴ Others have been skeptical of the assumption that there can be small-scale transitions that lead by a chain of changes from the social behavior of animals to the social behavior of people. In particular, Chomsky has argued that the acquisition of language is an all-or-nothing affair, which involves acquiring a rule-guided and creative capacity that cannot be built up from singular connections between words and things.¹⁵ A Chomskian would be dismissive of those attempts to inflict language on animals—on chimpanzees and dolphins, for example—that were once greeted with such enthusiasm, as the proof that they are like us or we are like them.¹⁶ Whatever the

¹⁴V. C. Wynne-Edwards, *Animal Dispersion in Relation to Social Behaviour* (Edinburgh: Oliver and Boyd, 1962). The original inspiration here is Lorenz, *On Aggression*. Wynne-Edwards is somewhat cantankerously criticized by Richard Dawkins in *The Selfish Gene* (1976), revised ed. (New York: Oxford University Press, 1989), pp. 7–10.

¹⁵See especially Noam Chomsky, *Language and Mind* (1968), 3rd ed. (Cambridge: Cambridge University Press, 2006), in which language is described as “an example of true emergence—the appearance of a qualitatively different phenomenon at a specific stage of complexity of organization” (p. 62).

¹⁶For the attempts, see Eugene Linden, *Apes, Men and Language* (New York: Saturday Review Press, 1974); for the enthusiasm, see Mary Midgley, *Beast and Man: The Roots of Human Nature* (London: Routledge, 1978), pp. 215–251.

interest of the word-thing/word-experience connections that animals can make, these are connections of a radically different kind than those embedded in a transformational grammar. They are piecemeal associations that, detached from generative rules and semantic organization, remain no more vehicles of thought, dialogue, and interrogation than the warning cries of birds and bonobos or the wagging tails of dogs. Again, the objection is not widely regarded as conclusive, and geneticists have advanced theories of “protolanguage” that attempt to show both that there could be piecemeal advances toward linguistic competence and that these advances would be selected at the genetic level.¹⁷

GENES AND MEMES

We know that the human species has adapted to its environment; but we also know that it has adapted its environment to itself. It has passed adaptations to its offspring not only genetically but also culturally. It has shaped its world through information, language, and rational exchange. And while all those features can be acknowledged by biology and given a place in evolutionary theory,¹⁸ that theory will not,

¹⁷ See, for example, John Maynard Smith and Eörs Szathmáry, *The Major Transitions in Evolution* (Oxford: W. H. Freeman, 1995), pp. 303–308.

¹⁸ As exemplified, for instance, by Kim Sterelny, in his theory of cumulative niche construction. See his *Thought in a Hostile World: The Evolution of Human Cognition* (Oxford: Blackwell, 2003).

in the first instance, concern the replication of genes but, rather, the reproduction of societies. Moreover, human societies are not just groups of cooperating primates: they are communities of persons, who live in mutual judgment, organizing their world in terms of moral concepts that arguably have no place in the thoughts of chimpanzees. It is possible that cognitive science will one day incorporate these moral concepts into a theory of the brain and its functions and that theory will be a biological theory. But its truth will be tested against the distinctively human capacities that, according to Wallace, seem “superfluous to evolutionary requirements,” and not against the features of our biological makeup that we share with other animals.

Now, philosophers who argue in that way find themselves confronting a powerful current of opinion that has flowed through all the channels of intellectual life since the publication of Richard Dawkins’s *The Selfish Gene*. Natural selection can account for all the difficult facts presented by human culture, Dawkins suggests, once we see culture as developing according to the same principles as the individual organism. Just as the human organism is “a survival machine” developed by self-replicating genes, so is a culture a machine developed by self-replicating “memes”—mental entities that use the energies of human brains to multiply, in the way that viruses use the energies of cells. Like genes, memes need *Lebensraum*, and their success depends upon finding the ecological niche that

enables them to generate more examples of their kind. That niche is the human brain.¹⁹

A meme is a self-replicating cultural entity that, lodging in the brain of a human being, uses that brain to reproduce itself—in the way that a catchy tune reproduces itself in hums and whistles, so spreading like an epidemic through a human community, as did “La donna è mobile” the morning after the first performance of *Rigoletto*. Dawkins argues that ideas, beliefs, and attitudes are the conscious forms taken by self-replicating entities, which propagate themselves as diseases propagate themselves, by using the energies of their hosts: “Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation.”²⁰ Dennett adds that this process is not necessarily harmful:²¹ there are, among parasitic organisms, both symbionts, which coexist harmlessly with their hosts, and mutualists, which positively amplify the host’s ability to survive and flourish in its environment.

¹⁹For various attempts to give a memetic theory of culture, see Robert Aunger, ed., *Darwinizing Culture: The Status of Memetics as a Science* (Cambridge: Cambridge University Press, 2000). The theory of the meme is dismissively criticized by David Stove in “Genetic Calvinism, or Demons and Dawkins,” in *Darwinian Fairytales: Selfish Genes, Errors of Heredity, and Other Fables of Evolution* (New York: Encounter Books, 2006), pp. 172–197.

²⁰Dawkins, *Selfish Gene*.

²¹Daniel C. Dennett, *Breaking the Spell* (London: Allen Lane, 2006).

To make the theory remotely plausible we must distinguish memes that belong to science from memes that are merely “cultural.” Scientific memes are subject to effective policing by the brain that harbors them, which accepts ideas and theories only as part of its own truth-directed method. Merely cultural memes are outside the purview of scientific inference and can run riot, causing all kinds of cognitive and emotional disorders. They are subject to no external discipline, such as that contained in the concept of truth, but follow their own reproductive path, indifferent to the aims of the organism that they have invaded.

That idea is appealing at the level of metaphor, but what does it amount to in fact? From the point of view of memetics, absurd ideas have the same start in life as true theories, and assent is a retrospective honor bestowed on reproductive success. The only significant distinction to be made when accounting for this success is between memes that enhance the life of their hosts and memes that either destroy that life or coexist symbiotically with it. It is one of the distinguishing characteristics of human beings, however, that they can distinguish an idea from the reality represented in it, can entertain propositions from which they withhold their assent, and can move judge-like in the realm of ideas, calling each before the bar of rational argument, accepting them and rejecting them regardless of the reproductive cost.

It is not only in science that this attitude of critical reflection is maintained. Matthew Arnold famously

described culture as “a pursuit of our total perfection by means of getting to know, on all matters which most concern us, the best which has been thought and said in the world, and, through this knowledge, turning a stream of fresh and free thought upon our stock notions and habits.”²² Like so many people wedded to the nineteenth-century view of science, Dawkins overlooks the nineteenth-century reaction—which said, “Wait a minute: science is not the only way to pursue knowledge. There is moral knowledge too, which is the province of practical reason; there is emotional knowledge, which is the province of art, literature, and music. And just possibly there is transcendental knowledge, which is the province of religion. Why privilege science, just because it sets out to *explain* the world? Why not give weight to the disciplines that *interpret* the world and so help us to be at home in it?”

That reaction has lost none of its appeal. And it points to a fundamental weakness in “memetics.” Even if there are units of memetic information, propagated from brain to brain by some replicating process, it is not they that come before the mind in conscious thinking. Memes stand to ideas as genes stand to organisms: if they exist at all (and no evidence has been given by Dawkins or anyone else for thinking that they do), then their sempiternal and purposeless reproduction is no concern of ours. Ideas, by contrast, form part

²²Matthew Arnold, *Culture and Anarchy: An Essay in Political and Social Criticism* (London, 1869).

of the conscious network of critical thinking. We assess them for their truth, their validity, their moral propriety, their elegance, completeness, and charm. We take them up and discard them, sometimes in the course of our search for truth and explanation, sometimes in our search for meaning and value. And both activities are essential to us. Although culture isn't science, it is nevertheless a conscious activity of the critical mind. Culture—both the high culture of art and music and the wider culture embodied in a moral and religious tradition—sorts ideas by their intrinsic qualities, helps us to feel at home in the world and to resonate to its personal significance. The theory of the meme neither denies that truth nor undermines the nineteenth-century view that culture, understood in that way, is as much an activity of the rational mind as science.

SCIENCE AND SUBVERSION

The concept of the meme belongs with other subversive concepts—Marx's "ideology," Freud's unconscious, Foucault's "discourse"—in being aimed at discrediting common prejudice. It seeks to expose illusions and to explain away our dreams. But it is itself a dream: a piece of ideology, accepted not for its truth but for the illusory power that it confers on the one who conjures with it. It has produced some striking arguments—not least those given by Daniel Dennett in *Breaking the Spell*. But it possesses the very fault for

which it purports to be a remedy: it is a spell, with which the scientific mind seeks to conjure away the things that pose a threat to it.

Reflecting on this, it seems clear to me that Wallace had a point in the emphasis that he put on the features that seem to place humanity in a world apart, though he was surely wrong to think of those features as “surplus to evolutionary requirements,” for if any of our attributes is adaptive, rationality surely is. But then, rationality is, in one sense of that difficult expression, “of our essence.” Wallace was therefore pointing to the fact that we human beings, even if we are animals, belong to a kind that does not occupy a place in the scheme of things comparable to that of the other animals. And the *philosophical* controversy here—a controversy adjacent to that among biologists and evolutionary psychologists concerning the significance of culture—is precisely a controversy about human nature: To what kind do we belong?

Dawkins sets out to explain goals and rational choices in terms of genetic materials that make no choices. He describes these materials as “selfish” entities, motivated by a reproductive “goal,” but (at least in his less rhetorical moments) he recognizes that genes are not, and cannot be, selfish, since selfishness is a feature of people, to be characterized in terms of their dispositions and their rational projects.²³ In a cogent

²³Though David Stove takes Dawkins to task for his constant reference to “selfishness” and his failure to say what it could possibly mean in this context: see Stove, “Genetic Calvinism.”

biological theory all such teleological idioms must be replaced with functional explanations.²⁴ And that is what the recourse to game theory and similar devices is supposed to authorize. A player wants to win and therefore adopts a winning strategy: that is a teleological explanation of this behavior. Natural selection tells us that winning strategies will be selected, even when they describe the behavior of genes that want nothing at all. That is a functional explanation, which says nothing about intentions, choices, or goals.

Functional explanations have a central place in biology.²⁵ The fact that birds have wings is explained by the function of wings, in enabling birds to fly. The process of random mutation at a certain point produces a winged creature: and in the competition for scarce resources, this creature has the decisive advantage over its rivals. Note, however, that this reference to function only amounts to a causal explanation because it is supplemented by the theory of random mutation—a theory that tells us *how* the existence of a trait is caused by its function. This point bears heavily on the “explanations” of altruism and morality

²⁴How teleological thinking can be replaced by functional explanation is one theme of Richard Dawkins’s subsequent book, *The Blind Watchmaker* (Oxford: Oxford University Press, 1986). For an illuminating discussion of functional explanations and their application outside biology, see G. A. Cohen, *Karl Marx’s Theory of History: A Defence* (Princeton: Princeton University Press, 1978).

²⁵See Ron Amundson and George V. Lauder, “Function without Purpose: The Use of Causal Role Function in Evolutionary Biology,” in D. Hull and M. Ruse, eds., *The Philosophy of Biology*, Oxford Readings in Philosophy (Oxford: Oxford University Press, 1998), pp. 227–257.

advanced by Axelrod and Maynard Smith. A population genetically averse to cooperation, to parental affection, to self-sacrifice on behalf of children, and to sexual restraint and the control of violence is a population endowed with traits that are dysfunctional relative to reproduction. Hence it will disappear. From this trivial truth, however, we can deduce nothing about the causes of moral conduct or moral thought and nothing about their grounds. It does not follow that morality is the result of natural selection rather than group selection within the species; nor does it follow that morality originates in our biological makeup rather than in the workings of rational thought. In fact nothing follows that would serve either to bypass or to undermine the work of philosophy in exploring the foundations of moral judgment and its place in the life of a rational being.

It is a trivial truth that dysfunctional attributes disappear; it is a substantial theoretical claim that functional attributes exist *because of* their function.²⁶ And until the theory is produced, the claim is without intellectual weight. You may think that genetics provides the needed theory: for it implies that altruism is the “evolutionally stable” solution to genetic competition within our species. But that explanation only gives a *sufficient* condition for “altruism,” and only by

²⁶A similar objection can be mounted, it seems to me, against the defense of Marx’s theory of history presented by G. A. Cohen (*Karl Marx’s Theory of History*). That dysfunctional institutions disappear is no ground for thinking that the existence of an institution is caused by its function.

re-describing altruism in terms that bypass the higher realms of moral thought. If Kant is right about the categorical imperative, then there is an independent sufficient condition, namely, rationality, that tells us to act on that maxim that we can will as a universal law.

Moreover, practical reason explains not only altruism, in the minimalist description favored by geneticists, but also the superstructure of moral thought and emotion. It also suggests a theory of *the kind to which we belong*, and it is a theory at odds with that suggested by the game-theoretic account of genetic self-sacrifice. According to Kant, the kind to which we belong is that of *person*, and persons are by nature free, self-conscious, rational agents, obedient to reason and bound by the moral law. According to the theory of the selfish gene, the kind to which we belong is that of *human animal*, and humans are by nature complicated by-products of their DNA. Kant saw his theory as raising the human being “infinitely above all the other beings on earth.”²⁷ But it is also true that his theory allows that nonhuman beings may nevertheless belong to the same kind as us: angels, for instance, and maybe dolphins too. The selfish gene theory would dismiss the suggestion as nonsense.

In the hands of their popularizers, the biological sciences are used to reduce the human condition

²⁷Immanuel Kant, *Anthropology from a Pragmatic Point of View*, ed. Robert Loudon and Manfred Kuehn (Cambridge: Cambridge University Press, 2006), p. 1.

to some simpler archetype, on the assumption that what we are is what once we were and that the truth about mankind is contained in our genealogy. The previous wave of pop genetics, which called itself “sociobiology,” came up with deliberately disturbing conclusions, such as this one: “Morality has no other demonstrable ultimate purpose than to keep human genetic material intact.”²⁸ Such conclusions depend upon using the language of common sense while at the same time canceling the presuppositions on which commonsense terms depend for their meaning. This trick can be played in almost any area of human thinking and is never more effective than when it is used to pour scorn on our moral and religious ideas. Ordinary people are in the unfortunate position of believing things that are true but which they cannot defend by any rational argument that will withstand the force of scientific reasoning, however flawed that reasoning may be. Hence, by targeting ordinary beliefs—beliefs that, if backed up at all, are backed up by religious faith and not by scientific argument—scientists score easy points and conceal the weakness of their case.²⁹

²⁸E. O. Wilson, *On Human Nature* (Cambridge, Mass.: Harvard University Press, 1978), p. 168.

²⁹This accusation was strongly made against Dawkins, in the context of the original TV series of *The Selfish Gene*, by Mary Midgley (*Beast and Man*, pp. 102–103). Whether Midgley’s objections are fair is a moot point; but she deserves credit for recognizing that the challenge presented by Dawkins goes to the heart of philosophical anthropology. Her criticisms of sociobiological writers are more pertinent and have been amplified in her *Evolution as a Religion*, revised ed. (London: Routledge, 2002).

UNDERSTANDING LAUGHTER

I do not deny that we are animals; nor do I dissent from the theological doctrine that our biological functions are an integral part of our nature as human persons and also the objects of fundamental moral choices.³⁰ But I want to take seriously the suggestion that we must be understood through another order of explanation than that offered by genetics and that we belong to a kind that is not defined by the biological organization of its members. The “selfish gene” theory may be a good account of the origin of the human being; but what a thing is and how it came to be are two different questions, and the answer to the second may not be an answer to the first. It may be as impossible to understand the human person by exploring the evolution of the human animal as it is to discover the significance of a Beethoven symphony by tracing the process of its composition.

Consider one of those features of people that set them apart from other species: laughter. No other animal laughs. What we call the laughter of the hyena is a species sound that happens to resemble human laughter. To be real laughter it would have to be an expression of amusement—laughter *at* something, founded in a complex pattern of thought. True, there is also “laughter at what ceases to amuse,” as Eliot puts it. But

³⁰This view is eloquently defended by Pope John Paul II in the encyclical *Veritatis Splendor*, August 6, 1993, sections 47 et seq.

we understand this “hollow” laughter as a deviation from the central case, which is the case of amusement. But what is amusement? No philosopher, it seems to me, has ever quite put a finger on it. Hobbes’s description of laughter as “sudden glory” has a certain magical quality; but “glory” suggests that all laughter is a form of triumph, which is surely far from the truth. Schopenhauer, Bergson, and Freud have attempted to identify the peculiar thought that lies at the heart of laughter: none, I think, with more than partial success.³¹ Helmuth Plessner has seen laughing and crying as keys to the human condition, features that typify our distinctiveness.³² But his phenomenological language is opaque and leads to no clear analysis of either laughter or tears.

One contention, however, might reasonably be advanced, which is that laughter expresses an ability to accept our all-too-human inadequacies: by laughing we may attract the community of sentiment that inoculates us against despair. This fact about laughter—that it points to a community of sentiment—has been well brought out by Frank Buckley.³³ From that suggestion, however, another follows. Only a being who makes judgments can laugh. Typically we laugh at

³¹See R. Scruton, “Laughter,” in *The Aesthetic Understanding* (London: Methuen, 1982), pp. 180–194.

³²Helmuth Plessner, *Laughing and Crying: A Study in the Limits of Human Behavior*, trans. J. Spencer Churchill and Marjorie Grene (Evanston: Northwestern University Press, 1970).

³³F. H. Buckley, *The Morality of Laughter* (Ann Arbor: University of Michigan Press, 2003).

things that *fall short* or at witticisms that place our actions side by side with the aspirations that they ridicule. If the laughter of children seems not to conform to that suggestion, it is largely because the judgments of children, like the laughter that springs from them, are embryonic—stages on the way to that full readiness of social assessment that is the basis of adult life. Insofar as children are amused by things, it is because, in their own way, they are comparing those things with the norms that they challenge. Putative cases of amusement in chimpanzees should, it seems to me, be understood in a similar way.³⁴ Creatures coaxed by their human masters to the verge of judgment are on the verge of amusement too. And by getting to the verge they reveal how wide for them is the chasm that human children will cross with a single stride.

To explain laughter, therefore, we should have to explain the peculiar thought processes involved in our judgments of others; we should have to explain the pleasure that we feel when ideal and reality conflict and also the peculiar social intentionality of this pleasure. Of course, we can make a stab at this kind of explanation, postulating cognitive programs in the human brain and the biological “wetware” in which they are imprinted. But as yet the explanation will be a pure speculation, with little or no input from genetics.

³⁴For an example, see the case of Roger and Lucy—two chimpanzees with some competence in the “Ameslan” sign language—described in Linden, *Apes, Men and Language*, p. 97.

I envisage evolutionary psychologists offering the following account of laughter. By laughing together at our faults, they might say, we come to accept them, and this makes cooperation with our imperfect neighbors easier, since it neutralizes anger at our shared inadequacies. Hence a community of laughing people has a competitive advantage over a community of the humorless. A moment's reflection will reveal the emptiness of that explanation. For it assumes what it needs to explain, namely, that laughter promotes cooperation. Admittedly my way of describing laughter suggests that this is so. But it suggests it by quite another route than that presented by biology or the theory of genetics.

I was describing a thought process, involving concepts such as those of fault and ideal that can have no clear place in evolutionary biology, as we now know it. I was assuming that laughter is an expression of understanding and that this understanding may be shared. And at no point did I assume that the sharing of laughter benefits anybody's genes in any of the ways that feature in the theory of genetics. Indeed, so far as my account was concerned, laughter might be an entirely redundant by-product of human life. It seems otherwise only because of my account, which is not a scientific account at all but an exercise in what Dilthey called *Verstehen*—the understanding of human action in terms of its social meaning rather than its biological cause.³⁵

³⁵See Rudolf Makkreel, *Dilthey: Philosopher of the Human Studies* (Princeton: Princeton University Press, 1993). Makkreel is currently

Suppose a group of zoologists were to come across a species that sat around in groups, pointing and emitting laughter-like sounds. How would they set about explaining this behavior? They would *first* have to know whether what they observed was real laughter. In other words they would have to know whether these creatures were laughing *at* something and pointing *at* something. And this word *at* does not yield easily to scientific analysis. It is a marker of intentionality, the “mental direction upon an object,” as Brentano described it,³⁶ and can be deciphered only if we are able to interpret the thought processes from which the behavior in question flows. All the work of explanation, therefore, depends upon a prior work of interpretation, the point of which is to settle the question whether these creatures are like us in being amused by things or whether, on the contrary, they are not like us at all, and their laughter-like behavior is something to be explained in another way. If we come to this second conclusion, the apparatus of ethology can indeed be imported into the case: we can begin to ask what function this laughter-like behavior might perform in securing an ecological niche for the genes of those who engage in it. If we come to the first conclusion, then we need to understand these creatures as we understand one another—in terms of the way they

editing an accurate and scholarly English edition of Dilthey's works, which is in the course of publication by Princeton University Press.

³⁶Franz Brentano, *Psychology from an Empirical Standpoint*, trans. L. McAlister (London: Routledge, 1974), p. 77.

conceptualize the world and the values that motivate their response to it.

I used the phrase “like us,” implying that amusement is one of our characteristics. And the question before us is how we should unpack that phrase. What do we mean when we refer to “creatures like us”? Do we mean to include only humans? Or do we have some wider, or perhaps narrower, category in mind? Homer tells us of the “laughter of the gods,” and Milton of laughter among the angels. Here is the beginning of a profound metaphysical problem. We belong to a natural kind, the kind *Homo sapiens sapiens*, which is a biological species. But when we talk of creatures like us, it seems that we do not necessarily refer to our species membership.

One last point about laughter. As I described it laughter seems to have a beneficial effect on human communities: those who laugh together also grow together and win through their laughter a mutual toleration of their all-too-human defects. But not everything that confers a benefit has a function. Entirely redundant behavior—jumping for joy, listening to music, bird-watching, prayer—may yet confer enormous benefits. By calling it redundant I mean that those benefits are the effect of the behavior, not its cause. That is how it is with laughter. There are communities of the humorless in which laughter is perceived as a threat and severely punished. But the humorless community is not for that reason dysfunctional; in itself it is as well equipped for survival as a community of comedians.

It is arguable indeed that the humorless Puritanism of the Massachusetts colonists was an important stimulus to their survival strategies during the early years. But the thing that they lacked would nevertheless have been a benefit to them, since laughter is something that rational beings enjoy.

THE GENEALOGY OF BLAME

I turn now to another feature of the human condition that divides us from our simian relatives: the feature of responsibility. We hold each other accountable for what we do, and as a result we understand the world in ways that have no parallel in the lives of other species. Our world, unlike the environment of an animal, contains rights, deserts, and duties; it is a world of self-conscious subjects, in which events are divided into the free and the unfree, those that have reasons and those that are merely caused, those that stem from a rational subject and those that erupt into the stream of objects with no conscious design. Thinking of the world in this way, we respond to it with emotions that lie beyond the repertoire of other animals: indignation, resentment, and envy; admiration, commitment, and praise—all of which involve the thought of others as accountable subjects, with rights and duties and a self-conscious vision of their future and their past. Only responsible beings can feel these emotions, and in feeling them, they situate themselves in some

way outside the natural order, standing back from it in judgment. From Plato to Sartre, philosophers have differed radically in their attempts to account for these peculiar features of the human condition: but almost all have agreed in searching for a philosophical rather than a scientific account.

There is one interesting historical exception to that claim, however, and that is Nietzsche, who, in *The Genealogy of Morals*, tries to explain the origins of responsibility in a way that anticipates the more recent attempts of geneticists to account for the moral life in terms of survival strategies that benefit our genes. Nietzsche envisages a primeval human society, reduced to near-universal slavery by the “beasts of prey,” as he calls them—namely, the strong, self-affirming, healthy egoists who impose their desires on others by the force of their nature. The master race maintains its position by punishing all deviation on the part of the slaves—just as we punish a disobedient horse. The slaves, too timid and demoralized to rebel, receive this punishment as a retribution. Because they cannot exact revenge, the slaves expend their resentment on themselves, coming to think of their condition as in some way deserved, a just recompense for their inner transgressions. Thus is born the sense of guilt and the idea of sin. From the *ressentiment*, as he calls it, of the slave, Nietzsche goes on to derive an explanation of the entire theological and moral vision of Christianity.

According to Nietzsche’s genealogy, the master race benefits from the subjection of the slaves—and

you can see this as the premise of a protobiological, even protogenetic, explanation of its social strategy. The master race secures its position by a regime of punishment, and in due course the punishment is internalized by the slave to engender ideas of guilt, blame, desert, and justice. But why should the slave understand punishment in these elaborate and moralized terms? Why should the internalization of punishment lead to *guilt* rather than *fear*? A horse certainly fears the whip: But when has it felt guilty for provoking it? Why is the original exercise of force seen as a *punishment* rather than a mere need on the part of the one who inflicts it?

What, after all, is the distinction between suffering inflicted as a means to securing one's ends and suffering inflicted as a punishment? Surely the difference lies in the mind of the agent. The trainer thinks that the suffering he inflicts is *needed*; the one who punishes thinks that it is *due*. That is due which is deserved, and that is deserved which may be rightly and justly inflicted. In short, punishment is a moral idea, to be unpacked in terms of those concepts of justice, desert, and responsibility that Nietzsche was supposed to be explaining. His genealogy of morals works only because he has read back into the cause all the unexplained features of the effect. In other words, it is not a genealogy at all but a recognition that the human condition, in whatever primitive form you imagine it, is the condition of "creatures like us," who laugh and cry, praise and blame, reward and punish—that is,

who live as responsible beings, accountable for their actions.³⁷

There are other momentous truths about the human condition that, while often overlooked or downplayed by biologically minded thinkers, occupy a central place in the outlook of ordinary people: for example, there is the fact that we are persons, who regulate our communities through laws ascribing duties and rights. Some philosophers—Aquinas notably but also Locke and Kant—argue that it is “person,” not “human being,” that is the true name of our kind. And this prompts a metaphysical question brought to the fore by Locke and still disputed, which is that of personal identity. What is the relation between “same person” and “same human being” when both are said of Jill? Which description engages with the fundamental kind under which Jill is individuated and reidentified? I mention that question not so as to suggest an answer to it but in order to highlight the difficulties confronting the view that Jill is in some way reducible to the biological processes that explain her.³⁸ Under what conditions do those processes reproduce the person who Jill *is*?

³⁷Nietzsche’s attempted derivation of the moral sense has been undertaken from the standpoint of evolutionary biology by Philip Kitcher, in *The Ethical Project* (Cambridge, Mass.: Harvard University Press, 2011). It is for Kitcher’s readers to judge whether he succeeds in explaining the emergence of the moral sense without assuming it.

³⁸Moves toward an answer are given in David Wiggins, *Sameness and Substance Renewed* (Cambridge: Cambridge University Press, 2001), chapter 7.

There is also the division that separates merely conscious creatures from *self*-conscious creatures like us. Only the second have a genuine “first-person” perspective, from which to distinguish how things seem to *me* from how they seem to *you*. Creatures with “I” thoughts have an ability to relate to others of their kind that sets them apart from the rest of nature, and many thinkers (Kant, Fichte, and Hegel preeminently) believe that it is this fact, not the fact of consciousness per se, that creates or reveals the central mysteries of the human condition. Although dogs are conscious, they do not reflect on their own consciousness as we do: they live, as Schopenhauer put it, in “a world of perception,” their thoughts and desires turned outward to the perceivable world.

I have tried to illustrate the way in which, in order to construct vivid biological explanations of our mental life, we are tempted to read back into the biology all the things that it ought to be trying to explain. To aim for a plausible theory of human nature we must first of all resist that temptation. And we must be prepared to admit that such laws of species-being as we have established—the laws of genetics and the functional account of inherited characteristics—are not yet adequate either to describe or to explain our normal behavior. They fall short of the target, for the very reason that what we are is not the thing that they assume us to be. We are animals certainly; but we are also incarnate persons, with cognitive capacities that are not shared by other animals and which endow us

with an entirely distinctive emotional life—one dependent on the self-conscious thought processes that are unique to our kind.

THE EMBODIED PERSON

This returns us to the problem of the relation between the human animal and the person. This problem, as I see it, is not biological but philosophical. I can make only a tentative suggestion in response to it—a suggestion that has something in common with what Aristotle meant when he described the soul as the form of the body and with what Aquinas meant when he argued that, while we are individuated through our bodies, *what* is individuated thereby is not the body but the person.³⁹ I would suggest that we understand the person as an emergent entity, rooted in the human being but belonging to another order of explanation than that explored by biology.

An analogy might help. When painters apply paint to canvas they create physical objects by purely physical means. Any such object is composed of areas and lines of paint, arranged on a surface that we can regard, for the sake of argument, as two-dimensional. When we look at the surface of the painting, we see those areas and lines of paint and also the surface that contains them. But that is not all we see. We also

³⁹ Aristotle, *De anima*; Thomas Aquinas, *Summa Theologiae*, 1, 19, 4.

see—for example—a face that looks out at us with smiling eyes. In one sense the face is a property of the canvas, over and above the blobs of paint; for you can observe the blobs and not see the face, and vice versa. And the face is really there: someone who does not see it is not seeing correctly. On the other hand, there is a sense in which the face is not an additional property of the canvas, over and above the lines and blobs. For as soon as the lines and blobs are there, so is the face. Nothing more needs to be added in order to generate the face—and if nothing more needs to be added, the face is surely nothing more. Moreover, every process that produces just these blobs of paint, arranged in just this way, will produce just this face—even if the artist is unaware of the face. (Imagine how you would design a machine for producing Mona Lisas.)

Maybe personhood is an “emergent” feature of the organism in that way: not something over and above the life and behavior in which we observe it but not reducible to them either. Personhood emerges when it is possible to relate to an organism in a new way—the way of personal relations. (In like manner we can relate to a figurative picture in ways that we cannot relate to something that we see merely as a distribution of pigments.) With this new order of relation comes a new order of explanation, in which reasons and meanings, rather than causes, are sought in answer to the question “Why?” With persons we are in dialogue: we call upon them to justify their conduct in our eyes, as we must justify our conduct in theirs. Central to

this dialogue are concepts of freedom, choice, and accountability, and these concepts have no place in the description of animal behavior, just as the concept of a human being has no place in the description of the physical makeup of a picture, even though it is a picture in which a human being can be seen.

There is another thought that is helpful in describing the relation between persons and their bodies, a thought first given prominence by Kant and thereafter emphasized by Fichte, Hegel, Schopenhauer, and a whole stream of thinkers down to Heidegger, Sartre, and Thomas Nagel. As a self-conscious subject I have a point of view on the world. The world *seems* a certain way to me, and this “seeming” defines my unique perspective. Every self-conscious being has such a perspective, since that is what it means to be a subject of consciousness. When I give a scientific account of the world, however, I am describing objects only. I am describing the way things are and the causal laws that govern them. This description is given from no particular perspective. It does not contain words such as *here*, *now*, and *I*; and while it is meant to explain the way things seem, it does so by giving a theory of how they are. In short, the subject is in principle unobservable to science, not because it exists in another realm but because it is not part of the empirical world. It lies on the edge of things, like a horizon, and could never be grasped “from the other side,” the side of subjectivity itself. Is it a real part of the real world? The question is surely wrongly phrased, since it misconstrues

the deep grammar of self-reference and of the reflexive pronoun. When I refer to myself I am not referring to another *object* that is, as it were, hidden in the lining of the observable Roger Scruton. Self-reference is not reference to a Cartesian self but reference to this thing, the thing that I am, namely, an object with a subjective view.

We are not entitled to reify the “self” as a distinct object of reference. Nor can we accept—given the force of Wittgenstein’s antiprivate language argument—that our mental states exhibit publicly inaccessible features that somehow define what they really and essentially are.⁴⁰ Nevertheless, it is still the case that self-reference radically affects the way in which people relate to one another. Once in place, self-attribution and self-reference become the primary avenues to what we think, intend, and are. They

⁴⁰Though we should note the tenacity of the view that the felt “quale” of a mental state is a *fact* about it, inwardly but not outwardly observable, and bound up with its essential nature. It seems to me that the notion of qualia is an empty hypothesis, a wheel that turns nothing in the mechanism, as Wittgenstein would put it. In an interesting essay, however, Ned Block—one of the most sophisticated defenders of qualia in the current literature—argues that Wittgenstein inadvertently commits himself to the existence of qualia, in a form that goes against the tenor of his philosophy. Ned Block, “Wittgenstein and Qualia,” *Philosophical Perspectives* 21, no. 1 (2007): pp. 73–115. The debate here goes so far beyond the scope of these lectures that I can only refer the reader to the brilliant summary by Michael Tye (an equally sophisticated defender of qualia) in the Stanford Encyclopedia of Philosophy: <http://plato.stanford.edu/entries/qualia>. The position I take can be gleaned from my “The Unobservable Mind,” *MIT Technology Review*, February 1, 2005, <https://www.technologyreview.com/s/403673/the-unobservable-mind/>.

permit us to relate to each other as subjects and not as objects only; and that is what lies at the heart of those ideas for which Nietzsche gave his pseudo-scientific genealogy: ideas of responsibility, accountability, guilt, praise, and blame. By relating to Jill in this way, I come face-to-face with her: her essential being as a person “emerges” from her bodily reality, in the way that the face emerges from the colored blobs on the canvas.

INTENTIONALITY

In a series of books and essays Daniel Dennett has argued for the view that human beings are “intentional systems”—organisms that exhibit intentional states that are systematically connected.⁴¹ The behavior of intentional systems can be explained or predicted by attributing “propositional attitudes”: by describing them as both representing the world and seeking to change it. Not all intentional systems are human: Some animals exhibit intentional states; maybe computers, when sophisticated in the way that Turing foretold, can exhibit them too. Dennett himself takes an easygoing attitude, allowing anything to be an intentional system if our treating it as such gives us some ability to predict its behavior—so that even

⁴¹D. C. Dennett, “Intentional Systems,” *Journal of Philosophy* 68 (1971), reprinted in *Brainstorms* (Cambridge, Mass.: MIT Press, 1978).

a thermostat is an intentional system in Dennett's view.⁴² His motive in taking this line is to make way for a "genealogy" of intentionality, building toward "aboutness" from simple feedback mechanisms that operate unmythically in the ordinary physical world. But it is not necessary to follow Dennett in this. Whatever the genealogy of the intentional, we must recognize the very real difference that exists between behavior that is caused by and expressive of an intentional state and behavior that is not.

Brentano's original insight has been taken by subsequent philosophy to imply that an intentional state is founded on a reference that may fail or a thought that may be false.⁴³ We can attribute such a state only where there is the possibility of referential failure. Animals exhibit intentionality through their beliefs and desires; they may even exhibit the kind of nonpropositional intentionality in which an object is "before the mind" and mentally targeted—as when a dog barks at an intruder, whether or not an intruder is there. It is certainly true that we are intentional systems and that this is a feature of our biological organization. Our brains are not merely devices for mediating between

⁴²See, for example, D. C. Dennett, *Kinds of Minds* (London: Weidenfeld, 1996), p. 34.

⁴³See the classic essay R. M. Chisholm, "Sentences about Believing," *Proceedings of the Aristotelian Society* 56, no. 1 (1955–1956): pp. 125–148. It is doubtful that this interpretation represents what Brentano really meant, however. See Barry Smith, *Austrian Philosophy* (LaSalle, Ill.: Open Court, 1994).

stimulus and response but instruments that enable us to think about and perceive the world and which lead us at times to think about it and perceive it wrongly.

In referring to the emergence of personality and self-consciousness, however, I am not referring only to this familiar feature of the human condition. I am referring, as Dennett has pointed out,⁴⁴ to a higher level of intentionality, one that is only doubtfully exhibited by other animals and which has certainly not been simulated by a computer.

A dog sees its owner as a living thing, capable of eye contact; but there is no place in its mental repertoire for the thought of its owner as a “subject of consciousness,” capable also of I-contact. By contrast, we humans respond to each other and to other animals *as* intentional systems, recognizing a distinction between how things are in the world and how they seem to other observers and adopting the “intentional stance” that Dennett again has emphasized in a series of books and essays.⁴⁵ But once we admit the existence of the intentional stance—the stance that interprets the behavior of other creatures in terms of the propositional attitudes expressed in it—we must recognize a higher (because more conceptually complex) level of intentionality. Our attitude to a dog is toward a creature with beliefs and desires; our attitude toward

⁴⁴D. C. Dennett, *Consciousness Explained* (London: Allen Lane, 1991).

⁴⁵D. C. Dennett, *The Intentional Stance* (Cambridge, Mass.: MIT Press, 1987).

a normal human being is toward a creature that *attributes* beliefs and desires to itself and to others and therefore to us.

Recognizing that others take this perspective on us, we become accountable for what we think and do, and we try to understand and relate to one another as responsible subjects of consciousness, each of whom has a unique perspective that informs his or her thoughts and actions. By describing this personal perspective as an “emergent” feature of the organism I am offering no theory of its nature—any more than I am offering a theory of pictures when I say that they emerge from the physical marks in which we see them. Rather, I am saying that at a certain level of complexity, a way of seeing others and ourselves becomes available to us and through this way of seeing we are confronted with another world than that described by evolutionary biology. This other world is the world in which we live—the *Lebenswelt*, to use Husserl’s term—the world of interpersonal attitudes.⁴⁶

⁴⁶The view I am arguing for has some connection with that defended by P. F. Strawson in “Freedom and Resentment,” in his *Freedom and Resentment and Other Essays* (London: Methuen, 1974), pp. 1–28. Unlike Strawson, however, I believe that the human being is truly represented in our interpersonal attitudes and falsely represented in those attitudes Strawson calls “objective.” The higher-order intentionality to which I refer—which is the ability to form mental representations of mental representations (one’s own and other people’s)—has been described, in important psychological studies by Alan Leslie and others, as “metarepresentation.” See, e.g., A. Leslie and D. Roth, “What Autism Teaches Us about Metarepresentation,” in S. Baron-Cohen, H. Tager Flusberg, and

Cont.