

Our Animal Bodies

STEVEN HORST

As this essay brings together two subjects that are not often spoken about in the same circles, moral psychology and cognitive science, I shall begin with some preliminary remarks about what I mean by the crucial terms 'moral psychology', 'cognitive science', and 'the passions'.

1. PRELIMINARIES

1.1 What Is Moral Psychology?

Moral psychology, as I shall use the term, is the study of the person (or self, our soul)¹ with an eye towards its good, rather than its merely theoretical study. Thus conceived, moral psychology involves at least three components:

1. A theoretical study of the nature of the self, or at least of some of its particular faculties.
2. A theory of what the good of the self consists in.
3. A set of therapeutic methods for achieving or moving towards the good of the self, thus conceived, in light of this theoretical understanding of its nature.

Two of the most basic questions we ask ourselves are, *What sort of being am I?* and *How shall I live?* From these arise a host of more specific and practical questions, among which few have either received or deserved so much attention as the passions. Anger, lust, jealousy, hope, despair, greed, the desire for power—all these are a part of our nature, and it is a fortunate person indeed who has never acted on their promptings only to regret it later. For anyone engaged in the task of living a moral life, the passions present both a theoretical problem (*What are the passions, and how do they operate?*) and a practical one (*How shall I control them as best I may?*). They thus present problems that fall in the sphere of *moral psychology*, a term which I shall explain in the next section.

My intention in this essay is to provide the basis for a theoretical account of the passions that is also practically and therapeutically useful. It takes its guiding metaphor from Plato's myth of the beast in Book IX of the *Republic*, but fleshes the metaphor out in more exact terms taken from contemporary cognitive science. In a nutshell, the view is as follows: the passions are a collection of separate special-purpose mechanisms built into animals (including human animals). One can tell a plausible story about each of these animal mechanisms in terms drawn from cognitive science, particularly evolutionary biology, that shows how they could have been selected for some adaptive advantage conferred upon an individual or its phenotype. These stories often illuminate puzzling features of our motivations and behavior. More importantly, they allow us to see how our passions are good things in their own right, without thereby being a good guide to conduct, especially not a good guide to moral conduct. And this analysis also provides vocabulary for talking about various ways these animal mechanisms can fail to function properly, and can be out of line with our moral aims even when functioning properly.

Moral psychology has its beginnings in the recognition that facts about *what kinds of beings we are* have implications, both for what vision we conceive of the human good, and for what methods we should adopt to pursue it. Moral psychology differs from moral *theory* or ethics, both in its emphasis upon factors that depend upon the facts of human nature and in its practical therapeutic emphasis. It asks not merely "What is the good?" (much less "What do we mean by 'good'"), but "How can a being who is constituted as I am move towards achieving it?" Moral psychology differs from *theoretical* psychology in that its goals are not merely theoretical understanding (much less prediction and control of behavior), but achieving the good of the soul, thus invoking a normative component which academic psychology lacks. It also differs from the bulk of *clinical* psychology, which is often leery of any conception of the good of the person which goes beyond the cure of mental illness and the pursuit of such value-neutral goals as integration and integrity of the personality.

While a thoroughgoing moral psychology would require all three of the components listed above, one may make contributions to the project of moral psychology by way of insights in any of the three areas, provided that they can be used to inform the others towards the overall goal. Thus, for example, a realistic description of some psychological feature such as empathy or one of the passions may be a useful contribution insofar as it allows us better to understand the resources or impediments to moral improvement that lie within us, even though the description itself may lack any normative commitments.

1.2 What Are the Passions?

It is widely held that Plato, in his early period, held that virtue consists wholly in knowledge of the good, and that happiness and the good life consist wholly in

the possession of virtue. Proper interpretation of Plato on this sort of point is, of course, a notoriously thorny bramble into which I do not care to wander. At any rate, he surely intended such a theory to be *recognized* by the reader of early dialogues such as the *Gorgias*. However, as Plato himself clearly understood by the time he wrote the *Republic*, this is a very dubious theory. For we all have had the experience of *believing* that one course of action was right and yet *doing* something else. We sometimes say we were “overcome by” hunger, or lust, or jealousy, or rage—that we suffered “weakness of the will” or *akrasia*. The obvious interpretation of such experiences, vividly illustrated by Plato in the *Republic*, is that the processes that lead to our actions are not entirely a rational matter. Instead, reason (or, better, our considered beliefs about what is right) must *compete for the control of behavior* with various other forces within us, which Plato groups as *thumos* (the kind of spiritedness characteristic of warriors and guard dogs) and the appetites. In short, *being good* is not simply a matter of *knowing* what is good and doing it; rather, our psychology is much more complicated than that, and even the person who succeeds in the nontrivial task of figuring out what is the right thing to do may not succeed in doing it. St. Paul expresses this eloquently in his letter to the Christians at Rome when he says, “The good that I wish to do I do not do, but the very evil that I wish not to do I find myself doing! Wretched man that I am! Who will save me?” (Rom. 7).

These nonrational parts of ourselves are often called “the passions,” from the Greek word *paschô*, from which we also have the English word “passive.” The movements of passions such as jealousy, anger, lust, or hunger are not things that we bring about by action; rather they are things that *happen to us*, even though they happen entirely *within us*. We are passive with respect to them in two senses. First, they come upon us unbidden, and we cannot simply will their coming and going. Second, we are often unable to resist their control and act “out of anger,” “out of fear,” etc., in spite of our considered desire to behave otherwise. As a result, the usage of the word ‘passions’ has a broader and a narrower sense. In the broad sense, which I shall employ here, the word ‘passion’ is a generic term for nonrational motivations such as the emotions and appetites. The word is sometimes also used more narrowly, when restricted to the operation of these motivations *when they are “out of control” in us*—i.e., in those cases where we might rightly say that it is not *I* but my *anger* (greed, lust, etc.) that is controlling my actions. To repeat, I use the word ‘passions’ here in the broad sense.

Understandably, discussions of the passions comprise a large part of classical and medieval discussions of ethics,² as the problems of dealing with the passions make up a large part of the practical endeavor of striving towards a moral life. There are many possible attitudes we can take towards the passions. Among them:

- that they are unnecessary evils of which we need to be rid (e.g., Socrates’ suggestion in the *Phaedo* [be it serious or no] that the philosopher ought to welcome death because it frees him from the desires of the body);
- that neither they themselves nor actions based upon them are signifi-

- cant, as they belong to the body, and the soul will be freed from them and their effects after death (e.g., the extreme Gnostic dualism of Marcion);
- that they are temptations that need continually to be actively repressed and resisted;
- that they are temptations that need to be mortified through some sort of ascetic practice;
- that they are perversions of good mechanisms that can be redeemed or re-educated or retrained;
- that in fact there is no objective good, and hence the desires arising from the passions are themselves the only standard by which things are called “good.” (A position Plato puts compellingly in the mouth of Callicles in the *Gorgias*.)

The attitude we take towards the passions will have significant effects on how we go about the enterprise of moral improvement. Do we mortify the passions? Indulge them? Treat them like petulant children (Theresa of Avila) or domestic animals (Plato)? Wait for a miracle of grace to change us (St. Paul)? How to regard and treat the passions is surely one of the central practical issues in moral psychology, and its answer depends in great measure upon empirical facts about our nature: how we *ought* to treat our own passions depends on *what treatment of the passions could possibly be effective*, and this in turn depends on what kinds of things they are, which is an empirical fact about human psychology.

1.3 What Is Cognitive Science?

The resources to be applied to this problem here are taken from cognitive science. ‘Cognitive science’ is a kind of grab-bag term for work done in the various disciplines studying cognitive phenomena such as reasoning, learning, conditioning, development, language, and the control of behavior. It embraces work done in a wide variety of fields: psychology, artificial intelligence, neural modeling, ethology, neuroscience, linguistics, and philosophy. There are a number of different paradigms at work in cognitive science today. Among them are:

- 1) The computational paradigm employed in artificial intelligence since the 1940s, which applies computational notions such as symbolic representation and rule-governed programs to psychological models.
- 2) The Darwinian paradigm and study of animal behavior, which applies such concepts as selection, mutation, and adaptation.
- 3) The neural networks paradigm, which explains high-level behavioral and cognitive phenomena as emergent from properties of networks like those found in the brain.
- 4) The complex systems paradigm, including cybernetics and much work in neural modeling, which seeks explanations in terms of abstract properties that can be rendered precise in mathematical models.
- 5) Neuroscience and psychopharmacology, which draw direct connections

between cognitive and behavioral phenomena and structural and chemical phenomena in the organism's brain, nervous system, and body chemistry. Each of these may be viewed as at least a potential source of explanatory and therapeutic resources for moral psychology. The bulk of what will be said here, however, draws chiefly upon resources taken from Darwinian biology.

2. BASIC APPROACH

The basic approach I shall suggest to understanding the passions finds its grounding in cognitive science, but is clearly foreshadowed by Plato in Book IX of the *Republic*. There, he describes the soul as a kind of beast with many heads:

Mould, then, a single shape of a manifold and many-headed beast that has a ring of heads of tame and wild beasts and can change them and cause to spring forth from itself all such growths. . . . Then fashion one other form of a lion and one of a man and let the first be far the largest and the second second in size. . . . Join the three in one, then, so as in some sort to grow together. . . . Then mould about them outside the likeness of one, that of the man, so that to anyone who is unable to look within but who can see only the external sheath it appears to be one living creature, the man.

Let us, then say to the speaker who avers that it pays this man to be unjust, and that to do justice is not for his advantage, that he is affirming nothing else than that it profits him to feast and make strong the multifarious beast and the lion and all that pertains to the lion, but to starve the man and so enfeeble him that he can be pulled about whithersoever either of the others drag him, and not to familiarize or reconcile with one another the two creatures but suffer them to bite and fight and devour one another. And on the other hand he who says that justice is the more profitable affirms that all our actions and words should tend to give the man within us complete domination over the entire man and make him take charge of the many-headed beast—like a farmer who cherishes and trains the cultivated plants but checks the growth of the wild—and he will make an ally of the lion's nature, and caring for all the beasts alike will first make them friendly to one another and to himself, and so foster their growth. (*Republic*, 588c-589, Shorey translation, responses emended for quotation here)

There are several salient points to Plato's metaphor. (1) The soul is not unified with regards to motives and action: rather, there are a number of different sources of motivation that compete for control of action (including, in Plato's own taxonomy, reason, *thumos*, and the various appetites). In contemporary terms, Plato was an early proponent of psychological *modularity*. (2) These appetites are likened to animals, which can be domesticated or can go feral. (3) These appetites are each a *part* of the soul: there is no intimation that we can be rid of them. (There is no suggestion of having the animal heads surgically

removed, for example, as Heracles attempted to decapitate the hydra!) (4) If properly trained, each of them can be a good thing. (5) None of the appetites has the kind of global perspective necessary to make it a good guide for behavior: reason needs to be in the driver's seat. (6) The proper attitude of reason towards these "animals" is to care for or cherish them. (7) It must befriend them and domesticate them, like a gardener encouraging the domestic plants and weeding out the wild ones. (8) In order to do this, it is necessary to understand their natures.

Now this is an incredibly rich metaphor, and is in its own right suggestive of much in the way of therapeutic strategy: just as the trainer of animals must care for the animals and know a lot about their nature, the person who wishes to domesticate her own passions must understand them as well. She must know what they are for, what they feed on, what irritates them and soothes them, what useful work they can be made to do, how they can be made friendly to one another and to her overall interests, and so on. But I wish to push this insight in a slightly different direction, inspired by the study of behavior in humans and animals.

My guiding insight is this: animals are comprised of a large *grab-bag* of *special-purpose mechanisms* that have some sort of survival value. These mechanisms are generally relatively *stupid*—they respond to rather specific cues, have limited flexibility, and cannot themselves evaluate whether the actions they recommend are appropriate in the larger context in which the animal finds itself. A dog smells food and some inner mechanism goes off: it begins to salivate, it seeks the food, it is prepared to act in such a way as to eat it. This mechanism is useful, as nutrition is a necessary function. However, this mechanism does not itself evaluate other salient questions, such as whether there is also danger present. The dog's survival depends upon other mechanisms that detect various sorts of danger, and upon the competition between these mechanisms being such that it will, say, flee from tigers even when food is also present. It is, I think, deeply misleading to characterize the dog as *reasoning* about how to weigh these factors. Rather, there are special-purpose mechanisms that compete with one another, and the dogs that are made in such a way that the wrong mechanisms win out die young and do not pass on their genes.

Even if you believe that human beings are a result of a special act of creation by God, it is hard to deny that we are made up of many of the same mechanisms found in other animals. We, too, possess a grab-bag of animal mechanisms that have special narrow purposes that they clamor for with the mindless persistence of a baby bird or a Washington lobbyist. Many of these mechanisms are built-in. They go off without our bidding: you smell a steak or some other favorite food and a desire to eat it forms. A shapely body jogs by encapsulated in bright lycra and you have an inward response as surely as a kitten does when you drag a string in front of it or a lion does when a wildebeest wanders by on the Serengeti. For the most part, these mechanisms serve some purpose, whether for the individual (hunger) or for the species (sexual desire). They do not always *filter* for other relevant cues: your desires are as likely to be set off by the stimulus of your

neighbor's food (or spouse) as your own. They can be *well-adapted* or *maladapted* to their current environment. They can have bizarre and unhealthy *mutations*. (There are, to take the extreme cases, those who want to have sex with the wildebeest or eat their spouse.) In all this we are like the other animals, though of course the presence of language and general-purpose reasoning gives us special abilities in coordinating our desires—an important advance, to be sure, but simply one more feature among many that make up our psychology. Reason *stands alongside* the less intelligent mechanisms, and interacts with them with various degrees of control. It does not replace or eliminate them. We are animals that possess some measure of reasoning ability as well; reasoning does not make us cease to be animals.

2.1 *The Biological and the Moral*

There is also, however, a crucial difference of another sort which really does transform the story about animal mechanisms in us—though what it transforms is not their *nature* but their *significance*. With the other animals, the story we tell about the various mechanisms within them and their adaptive fitness with respect to their environment is *the whole story*. In them, these mechanisms have no moral overtones: both the other animals and the mechanisms within them are neither moral nor immoral. They are *nonmoral*. (If anyone thinks some of the higher animals also have moral properties, no matter, my point is simply that the mechanisms themselves are nonmoral.) Likewise, *our* animal mechanisms for hunger, sex, flight, battle, and what have you are *in themselves* nonmoral. They are merely biological—selected for particular purposes, and adapted or maladapted to their environments. However, unlike the other animals, we have this additional dimension to our lives: the moral dimension. And because moral predicates apply to us, these animal mechanisms take on new overtones in us that they do not have in the other animals.

I discovered, in the brief episode in my childhood when we tried to keep tropical fish, that Marble Mollies tend to eat their own young. This is, no doubt, the result of some mechanism within them—whether it is a special mechanism that has the gruesome function of bidding them devour their own offspring, or a general feeding mechanism that simply does not discriminate their own progeny from other small swimming objects, I do not know. But none of this has moral overtones for the Marble Molly. Fish are neither good nor bad. They are simply fish. We may use the expression “a good fish” to express some relation to our own interests (it is good to eat, it looks good in the aquarium), or even to evaluate it in relation to some biological end (it is healthy, or well-adapted to its environment, or flexible in adapting to a variety of environments). But, barring some surprising revelation about the inner lives of fish, it makes no sense to say of a fish that it is “good” or “bad” in the *moral* senses of those terms.

With humans it is quite different. I do not recall having heard of humans that have eaten their own children willingly (I believe Shakespeare has a story

in which some potentate's children are fed to him by an enemy); but there have been a number of human cultures that practiced cannibalism, and others that sacrificed their own children to some squalid backwater deity. Now if we were *merely* animals, this too would be a completely nonmoral fact about the behavior of the human animal, as it is a nonmoral fact about female black widow spiders that they devour the male after mating with it. However, while we are animals, we are not *merely* animals, but animals to which moral categories apply. And hence such acts, and the mechanisms that lead to them, take on moral overtones. Thus discussion of human psychology rapidly enters the domain of *moral* psychology as well. If we observe that in some animals, such as ducks and spiders, the sexual act is often associated with violence to the partner, this is a disturbing observation of animal behavior, but carries no moral overtones. When humans engage in the same behavior, it is a different matter.

This essay will thus attempt to flesh out the following ideas:

1. *The Modularity of Motivation*. We, like other animals, are not a single decision-theoretic unit; but a collection of special-purpose units that are often in competition with one another.
2. *Darwinian Teleology*. For the most part, these special-purpose units have some biological function, which in its normal operation would serve some biological interest of either the individual organism or its phenotype. (I do not mean this to represent any commitment to any particular theory of how these phenotypic traits arise in the first place: all this is equally compatible with gradualist, catastrophist, and creationist stories, though it does require some belief in *selection*.) An important distinction within this framework is the distinction between the biological interests of the *individual organism* and that of the *phenotype*: some traits are selected because they help the individual survive, while others are selected in spite of risk to the individual because they increase the overall probability of the phenotype surviving (e.g., the “self-sacrifice” of hive animals when the hive is threatened).
3. *Features of Implementation*. A psychological module is defined in Darwinian terms by what it *does*, but it must do it in some particular way. *Which* way it is designed will have implications—not only for how well it does its job, and under what circumstances, but for what side-effects may arise. (As, say, visual illusions are side-effects of the specific ways our visual systems are designed.) This takes place both with *global* features, like conditionability, and with more *specific* features, like the implementation of some emotional properties through hormone levels.
4. *Adaptation*. An animal mechanism within us may perform its task well in one environment but be maladapted to a different environment.
5. *Dysfunction*. An animal mechanism may be internally skewed so that it does not do its job well even in optimal circumstances, or so that it does something other than what it is supposed to do. This can take place either

for organic reasons (deficits, traumas to the brain or endocrine system), or due to conditioning (e.g., irreparable psychological and social damage done to monkeys deprived of contact during early critical periods).

6. *The Normativity of the Biological*. I draw a firm distinction between biological interests of either organism or species and the moral interests of a human being. Most species are not even subject to moral evaluation, and our own biological interests underdetermine (and even go against) our moral imperatives. Nonetheless, knowledge of our biological nature can be of assistance in helping us (a) to know what is really in the sphere of possibility for us (cf. Flanagan's [1991] "minimal realism") and (b) to go about the task of taming the passions.

This last point is quite important. We are *both* moral beings (in the weak sense of "beings subject to moral evaluation") and animals. The former is not, in my view, reducible to the latter. But the specific ways in which we are animals have a very real impact on the nuts and bolts of the task of being moral.

In the schema for moral psychology presented above—that it involves components of (1) theoretical psychology, (2) theoretical ethics, and (3) therapeutic methods—this essay is concerned principally with fleshing out the theoretical psychology in a way that might be useful in generating therapeutic methods for those already engaged with some theoretical view of the good life and morality. I shall say little or nothing here about the specifics of the good life or morality, and I think that little that I say here will depend upon one moral stance rather than another. What I present here is an approach to moral psychology that aims at giving those already engaged with an ethical position a theoretical stance towards the passions. It may well be that a study of our animal natures would, in the long run, lead us to reject some moral views as unrealistic. But that will not be the concern of this essay. The structure of the remainder of the essay is therefore as follows:

Modularity, Biological Explanation, and Ethics introduces two ideas from different branches of cognitive science which will be used in the rest of the essay, and assures the reader that they will not be used to *reduce* the ethical.

Tradeoffs in Design Constraints explains why a modular account of the passions makes good biological sense and presents an account of the nature of the passions as special-purpose mechanisms that are very good at spotting particular kinds of environmental cues quickly, but by no means perfectly or flexibly.

The sections *The Curse of the Sweet Tooth* and *The Dog Who Barked at Everything* explore two different ways that special-purpose mechanisms can go awry: maladaptedness and dysfunction.

The final two sections, *The Junkyard Dog of the Soul* and *The Job of Sex*, are brief applications of the resources developed here to two specific mechanisms within us: the boundary-protecting mechanism Plato calls *thumos* and sexual desire.

3. MODULARITY, BIOLOGICAL EXPLANATION, AND ETHICS

I wish in this section to exploit two important ideas. The first arises from recent work in cognitive science, particularly to the writings of Noam Chomsky (1980), Jerry Fodor (1983) and Zenon Pylyshyn (1984). It is the claim that the mind is not a single general-purpose reasoning tool, but has *modules* that serve particular functions. Some of these seem completely automatic and closed off from the information contained in the rest of the mind: e.g., your blink reflex goes into action when something rapidly approaches your face even if you know it will not hit you. Others share information very well, and still others are somewhere in between. Owen Flanagan (1991) has begun to explore the idea that modular design may be relevant to moral psychology—for example, that we have different faculties for rule-based and care-based reasoning. I wish to take this idea much further here, and claim that we have a nontrivial number of modules in us that produce motivations and representations relevant to moral psychology: e.g., the mechanisms responsible for sexual attraction, jealousy, anger, fear, hope, social dominance, etc. So the first idea is that *the mind is modular*.

The second idea is that these modular mechanisms can be looked at in biological (Darwinian) terms. In Darwinian explanation, a phenotypic trait is sometimes (though not always) explainable in terms of its survival value. Why does the sapsucker have its peculiar bill? Because, the story goes, some of its ancestors underwent a random mutation that made their bills longer and thinner than those of their cousins, and this gave them advantages in acquiring food. As a result, those with the long, thin bills as a group were more successful in living to produce viable offspring, and so their genes were perpetuated, and with the genes the phenotype that they control. There are two parts of this story: the mutational and the selective. Mutation is generally viewed as blind and random: the adaptedness of the sapsucker's bill did not cause it to emerge as a phenotype; it merely explains why, having appeared by random mutation, it made its possessors more successful at the biological task of spreading their genes. Of course, not all traits were selected for. Some are free riders. Some are merely epiphenomena of traits that were selected for. The point is not that biological explanation *always* works; it is, rather, that it provides a resource that goes beyond the resources of either mechanistic or Aristotelian-style teleological explanation.

There have been attempts, both in Darwin's lifetime and in the twentieth century, to try to *reduce* the ethical to the natural by way of biological explanation. Perhaps most famously in our own day, the trait of altruism, so difficult to explain in terms of egoistic theories of ethical reasoning, has been argued to be a trait that is explicable in terms of the survival advantages it provides for the species or population that possesses the trait, which is the unit of biological selection. I think that this sort of attempted reduction is misguided, for reasons that will emerge shortly. But it leads to two important distinctions which we must keep in mind if we are applying resources of biological reasoning in ethics.

First, we must distinguish between the biological "interests" of the *individ-*

ual organism and the biological "interests" of the *species*. I put the word 'interests' here in scare quotes, because there is some reason to doubt that lower animals, much less species, can have interests in the proper sense of the word. "Interests," for us, are intimately connected with what we do want and what we should want. But even with animals that may have nothing we could recognize as an inner life at all, there is something about being an animal of a particular kind (say, a dog or a fish) that carries with it some implications about what form of life would count as *thriving* for that particular animal, and what would count as the opposite. What makes for a good life for a dog would make a fish miserable or would even be impossible for it, and vice-versa. (Cf. Midgley, 1995) There is at least an extended sense of "interests" to be had here, and we may borrow some language from Spinoza, and say that the biological interest of the *individual* consists in *maintaining the form of life proper to it*. The biological interest of the individual consists in large measure in individual survival. It is not in the biological interests of the individual animal to die, and all animals have at least some mechanisms that help them maintain their own form of life.

However, what is good for the individual animal (biologically speaking) need not be good for the population and vice-versa. Among the social and the hive animals, for example, it is often the case that individuals will instinctively act in a manner that is harmful to their own well-being, such as bees and ants that will do things leading directly to their own deaths when there is a threat to the hive. I find it unlikely that bees and ants *think* about what they are doing at all, much less of the sacrifice they are making; and I doubt that they have any thoughts about individual survival or "the good of the colony." Their behavior is most readily interpreted as an expression of a phenotypic trait that was perpetuated because populations in which it was present were more viable as breeding populations than those that lacked it. The presence of this phenotypic trait was in the biological interests of the population (i.e., it provided a differential advantage in passing on of its genes), even though it was not in the biological interest of the individual animal. These two notions of individual and population biological interest are not incompatible, even though they often conflict. Rather, they are separate explanatory resources which are of use in explaining different aspects of animal behavior. It is, of course, the notion of the "interest" of the *phenotype* that arises specifically out of Darwinian explanation, though mechanisms that are aimed at individual biological interest are often a means to perpetuating the phenotype as well.

Second, neither individual nor species biological interest is the same as the moral interests of an individual. In the case of ants and bees, it is no more than a crude anthropomorphism to describe their behavior as "altruistic" or moral. While they may provide the basis for useful moral parables to tell our children, moral properties cannot seriously be imputed to them at all unless they are secretly possessed of much richer inner lives than we know. And with human beings, self-sacrifice is never in a person's individual biological interest, though it is sometimes morally good. And while the particular trait of altruism may often

be in our species' interest (though presumably not always—if the only breeding male dies protecting the females, there will be no next generation), there are surely things that are in the biological interest of the species that are immoral. Consider, for example, a scenario in which there is one woman and one man left alive, and she wants nothing to do with him. It is surely in the biological interest of the species that they reproduce together, regardless of how it is accomplished. But rape can never be moral.

My purpose is thus not to try to reduce the moral good to any form of biological interest, but to use biological interest as a tool for seeing why the various mechanisms in us *represent* certain courses of action as "good"—and indeed how they *are* "good" in *some* sense of the word, though not necessarily in the moral sense. Our animal mechanisms were designed to spot biological, not moral, salience.

4. TRADEOFFS IN DESIGN CONSTRAINTS—CREATION MYTH

Another key idea coming from cognitive science (and biology) is the distinction between *function* and *implementation*. Bird eyes and human eyes do much the same things, but are built differently. Natural hearts and artificial hearts do the same things, but are made of different materials, and as a consequence have different strengths and weaknesses (e.g., artificial hearts are sensitive to electromagnetic fields of the sorts found at airport checkpoints, while natural hearts are not). Whenever God or nature designed something in an animal, it was done in some particular way, and doing it in that way had consequences that often have little to do with the function of the mechanism.

It is no secret that we human beings are not omniscient, and that most of nature is even dumber than we are. Nonetheless, even profoundly stupid animals seem to do a pretty decent job of getting around, often on the basis of a few well-adapted instincts and "good tricks" that serve individual animals well enough and serve the population quite well indeed. Horseshoe crabs are so stupid and so badly sighted that they attempt to mate with anything shaped like a con-specific—rocks, upturned bowls—but there are enough conspecifics around that they will occasionally get lucky, and when they do, thousands of offspring result. It is enough to keep the species going. And presumably the rocks do not mind.

Much of what animals do—both what succeeds and what seems laughably peculiar, like the horseshoe crab's amorous bumbings—is a *result of basic design issues* that are a result of how viable animals are built out of relatively simple and often coarse-grained mechanisms. Some of these, like the trade-off between discriminatory power and speed, and the stability-plasticity dilemma, are general features of adaptive systems. Others are more particular to specific cases. In what follows, much will be based around the consideration of principles through real examples and thought experiments, and then applications of these insights to moral psychology.

4.1 *Fast vs. Fine-Grained Discriminations: Sometimes It Is Good to Be Stupid*

Creation myth: On the fifth day, around one in the afternoon, God appoints two angels to make some unguulates. (Animals with multiple stomachs that chew the cud.) Angel #1 makes deer. He makes them with a minimum of brain matter and hence they are stupid. They spend most of their time eating, sleeping, and testing the air for signs of predators, and when something rustles in the bushes they run like hell. They lead dull lives, and are often interrupted from eating and sleeping when a twig snaps, but they do not get eaten much by predators because they are quick to run, and boy are they fast! Angel #2 makes unicorns. He makes them smart and thoughtful. They think about high-minded things like deontology and set theory and, being far too cultivated to run from snapping twigs, carefully consider whether there is really a tiger present, what the best course of action would be, what might be done about this tiger problem once and for all. . . . Sometimes they come up with clever answers, but inevitably this happens just before they are eaten by tigers, so it comes to nothing. (And that's why you've never seen one to this very day.)

The point is that there are trade-offs between different ways of designing a cognitive system that can do useful things like detect predators. A quick and dirty system (the deer) is fast and sloppy: it picks up on cues that sometimes indicate predators, but produces many false positives. False positives are costly, as they cause stress to the animal and interrupt other necessary activities. False negatives, however, are even more costly when the task is predator detection, as they result in immediate death. So you want your system biased towards the less costly kind of mistake, though of course if it is biased too far the animal will always be running and die of hunger or exhaustion. A slow and careful system (the unicorn), by contrast, can screen out false alarms, but the cost for this is processing time, and the more carefully the animal processes things, the more time is lost. So for mechanisms that deal with fast-moving life-or-death situations in particular, a quick and dirty system is not only easier to come up with (i.e., it takes less sophisticated wetware, and less of it), it has better survival value as well. Hence things like sharks and cockroaches have remained in the same model-year for several eons, even though as individuals they are profoundly stupid. And it is good that even beings like ourselves, who have fairly formidable cognitive abilities, also have quick and dirty systems as well, ranging from the blinking reflex, to the surge of adrenaline when we are startled, to the rise of anger when we are attacked or our space is invaded. And this is an important point: quick and dirty mechanisms and slow and deliberate mechanisms are *different in kind, but they are not incompatible*, in the sense that they can be present in the same individuals. Indeed, human beings are a notable case in point.

4.2 *What Are Emotions?*

Many of the mechanisms that we call "emotions" can usefully be viewed as quick and dirty mechanisms that adapt the animal's behavior on a short time scale to

potentially relevant environmental cues,³ like the presence of food and water, physical danger, potential mates, and in the case of the social animals, other relevant social relations to conspecifics, like the arrival of a group member of higher, lower, or ambiguous status in the dominance hierarchy. *Cognition* can also serve the function of determining the presence of such factors, and can carefully sort out ambiguous situations. Emotion is fast and sloppy. It often gets the cues right in ways that baffle explicit cognition. (Think of how a socially astute person simply *knows* what is going on in a way that cannot be explained to a person without this astuteness who tries to reason it through.) However, it produces lots of false positives as well. Look, for example, at our reactions to things in movies and television: our emotions react as though the person we see on screen were really flirting with us or a *T. Rex* threatening us, experiencing real arousal or fear, and our bodies are primed for further moves in the internal scripts for sex or conflict, much like the bird that sings and preens for its own image in the mirror. But then cognition kicks in and says, "Hey, it's only a movie! Get a grip!"

How, then, ought we to regard such emotions? Thinking folk often have a tendency to undervalue these emotions, both for the somewhat reasonable reason of their high rate of error and for the less sensible reason that they seem "brutish." As a result, many intellectual folk (surely not all) are famously underdeveloped in their abilities to pick up on cues in their physical and social environment, as they neglect their rich endowment of special-purpose detectors. However, it is important to note two virtues of emotional awareness that accrue even to beings that also possess heightened cognitive abilities: (1) they are faster than reasoning, and sometimes you need to respond faster than you can think things through, and (2) because they are useful hacks rather than careful programs, they are sometimes brilliantly right in ways that are not penetrable to our reason: you just *know* that this person is up to something, that that one is the alpha member of the group, etc. And, biologically, this is not surprising: if you are going to build a social animal, and weed out less adapted traits over long periods of time, it is not surprising that the animals will have mechanisms that are specifically tuned to socially relevant stimuli, even though these may not be observable to, say, a Martian observer whose brain is not endowed with the same special-purpose hack to spot human (canine, dolphin, chimpanzee . . .) social cues.

Dealing with the emotions is not unlike keeping a watchdog. The dog will likely detect visitors before you will and start barking, and prepare for a hostile encounter. You are not nearly as good as the dog is at hearing (much less smelling) people skulking through the property, but once alerted, you generally have a much better idea of who is friend and foe and how to behave properly towards them than does the dog. (You *want* the dog to bark at strangers, but the mailman is *not* your enemy just because the dog barks at him.) You and the dog work as a team: the dog alerts you to a situation, and indeed if someone is rushing you the dog may deal with it before you have a chance to think, but once you are alerted (due, mind you, to the dog's advance warning) you may have time to reason out an appropriate response that is simply unavailable to the dog.

4.3 *Stability and Plasticity—The Joy of SpheX*

A closely related adaptive dilemma is the stability-plasticity dilemma. On the one hand, it is useful for an animal to be able to adapt to new situations—to be *plastic* in its behavior—since its environment may very well change during its lifetime.⁴ On the other hand, plasticity is *costly*, in that it is easier and more efficient to repeat a well-established pattern of behavior than to search for a new one. Plasticity is stressful, and has costs in terms of energy and time. (Imagine if you could never remember how to do something and had to always improvise, or if each morning the world had changed so much that none of your previous coping strategies could work.) Thus there is a drive towards *stability* as well—towards finding a lasting equilibrium with the environment.

The favorite example in recent years of animals who exhibit extremes of stability is the spheX wasp, which exhibits extreme behavioral stability in the form of a highly ritualized sequence of actions it performs in laying its eggs. The female spheX digs a burrow in which to lay her eggs and then paralyzes a cricket with her sting, but without killing it. She brings it to the mouth of the burrow and enters the burrow, presumably to check on its condition. She then places the cricket in the burrow and lays her eggs within the burrow as well. When the eggs hatch, the wasp larvae then feed upon the still-living cricket. What is interesting to biologists about the gruesome story is that, if the wasp is interrupted in the middle of this routine, she will retrace her steps unnecessarily rather than improvising. In particular, if the stunned cricket is moved while the wasp is examining the burrow—say, moved a few inches away from the aperture—the wasp will *not* simply drag it into the burrow directly, but will instead repeat the steps of dragging it to the mouth of the burrow, examining the burrow, and then emerging to look for the cricket. Woolridge (1963) reports that on one occasion this procedure was repeated forty times with the same result. (If you are the sort who finds it entertaining to play practical jokes upon frighteningly large insects, duplicating this observation might prove a satisfying afternoon's diversion.)

The spheX wasp is, in layman's terms, incredibly stupid, as stupidity is the opposite of intelligence, and intelligence is, in rough and ready definition, the ability of an individual organism to adapt to its environment. Its behavior, like most insect behavior, is completely inflexible and hard-wired. At a species level, this has proven to be a good evolutionary strategy, at least when combined with the ability to lay vast numbers of eggs. However, at the individual level, it is quite another story.

Apart from hard-wired, inflexible behavior (which of course may itself be the result of adaptation at a genetic level over the slow timescale of evolution), there are also at least two kinds of adaptive strategies: the *conditional* and the *improvisational*. Even very primitive animals are subject to some degree of flexibility through conditioning, and higher animals can adapt through conditioning to rapidly changing environments. Mammals and birds are often able to *improvise* as well—that is, to construct new strategies for dealing with a new situation

“on the fly.” (Raccoons and otters, for example, are famous for this.) In human beings, this improvisational ability is greatly enhanced by the flexibility of reasoning and language, and by our extraordinary capacities for imagination and will. The first three of these (reasoning, language, imagination) allow us to conceive of, model, and evaluate alternative possibilities that may never have been actually experienced, while the fourth (will) enables us to execute strategies even if they are contrary to our conditioning.

Stability and plasticity are not themselves moral categories. Both stable patterns and innovations can be good, bad, or indifferent. (The thumbscrew and biological weapons, after all, were innovations.) However, the stability-plasticity axis is of use in understanding particular moral situations and coping with them. At a very fundamental level, the very possibility of moral reflection and change are dependent upon high-level plasticity and the ability to model the effects of our own actions. And some of the most severe forms of mental illness, such as paranoia, monomania, and obsession, can be seen as maladaptive forms of extreme stability. (That is, deeply entrenched patterns that remain in the face of evidence that they are out of sync with the environment and with the organism's interests.) The spheX is neither good nor bad; a person whose life is similarly compulsive is, not to put too fine a point on it, in hell.

At another level, it is useful to use the notions of stability and plasticity in looking at our passions and the relation between reason and the passions. It is a good thing for us, in both biological and ethical terms, that we are capable of a rich *interplay* between stable patterns (such as Aristotle's moral virtues) and flexibility (such as the ability to reflect on whether our practices are good ones). We would *not* be well served by putting our decisions solely in the hands of a cognitive faculty that had to figure out each situation we encounter *de novo*. And practical reason is notoriously susceptible to undue influence by representations of short-term effects: when it comes to courage or charity, it is better to train yourself into a resolve (i.e., train yourself into a stable pattern of responding to particular sorts of situations) and then stick by your guns than to try to figure things out when the bullets, or the harsh words, are flying.

Likewise, in the project of domesticating the passions, it is important to ask of a particular passion in what ways it can be trained. There are doubtless a few things in human beings that are hard-wired, such as the activation of nonspecific arousal in the face of sudden loud noises. But what passion that nonspecific arousal grows into depends very much on training. The raw recruit may feel terror or confusion when a shell detonates; hopefully, by the end of boot camp, it incites a firm resolve instead. Or as Diotima suggested to Socrates, the erotic reaction to a beautiful body can be trained into a concern for the spiritual welfare of all people. It can likewise be trained in more destructive directions as well, as the Jeffrey Dahmers of this world continually remind us. How one trains the passions makes a profound moral difference; and those who would be morally profound must look to how they train their passions, and this requires us to know how to work with them.

5. THE CURSE OF THE SWEET TOOTH—MALADAPTEDNESS AND CHANGE

While most mechanisms have some function that provides some kind of benefit to the individual or the phenotype, there are a variety of ways that a mechanism can go awry. We may distinguish between *first-order* mechanisms, which do some special-purpose function, from *second-order* mechanisms, which make an organism more flexible and adaptive, allowing it to come up with new ways of providing first-order functions. First-order mechanisms can be adaptations on an evolutionary scale to some environmental feature (say, the phototropism in moths) or adaptations through learning in a single organism (say, learning to respond to an attack using specific karate moves). Second-order mechanisms are also selected in an evolutionary timeframe, but allow adaptations on a much faster scale to changes in environment within an organism's lifetime (i.e., the mechanisms that underlie the capacities for various types of learning).

A first-order mechanism is an adaptation to a specific environment. It may cease to be well adapted if the environment changes. The woolly mammoth's thick coat was a useful temperature regulator during the ice age. It became a disastrously maladapted temperature regulator when the Earth thawed out. The monogoose's instinct for fighting snakes serves it well in India, where it faces the comparatively slow cobra, but was disastrous when it was exported to Brazil to combat the much quicker bushmaster. In short, adaptation is always adaptation to a particular environment.

The same is true of us. A mechanism in us can be perfectly good in its own right, but maladapted to our current environment. And, what is slightly different, it can function well in one environment, but not in another. This would seem to have variations at the levels both of instinctive and learned behavior. First, consider our appetites for food. Many of us suffer an inordinate attraction to sweets or fats. Arguably, there is a straightforward evolutionary explanation for this: we were designed for an environment in which these things were relatively scarce. Green plants are fairly common in a variety of biomes, but fruits, nuts, and ready animal protein are harder to come by, and generally require more work to acquire. If you are designing an animal that needs things that are scarce as well as things that are plentiful, it makes sense to design stronger mechanisms for seeking and consuming the scarce things. Thus we prefer filets and desserts to salads and broccoli. However, if you were designing an animal for an environment rich in these rich foods, you would want to give it a completely different weighting to different food sources, de-emphasizing those that, while nutritious in small quantities, cause problems when consumed regularly in large quantities, like saturated fats and simple carbohydrates. When you take an animal designed for scarcity and put it into an environment of abundance of things that should not be eaten in large doses over extended periods, it will tend to harm itself with "good" things.

Of course, maladaptation is not the only way the appetites and other passions can come out of order, but it is neither insignificant nor useless to know

that our bodies may systematically mislead us about what they need because they are operating in a different environment than that which they were designed for. An altimeter is set to give the right atmospheric pressure for a given assumption of ambient barometric pressure. Fly into the heart of a storm and what it tells you is systematically wrong. Your body's demand for sugars and fats may have nothing to do with any real nutritional deficit, but may be an artifact of how our hunter-gatherer ancestors were optimized to survive by hunting and gathering fruits. Let them loose in a Denny's for a year, and you would have a pack of fat, constipated hominids with heart trouble and indigestion.

Adaptation of the appetites is not a moral matter in and of itself. But the appetites surely affect our ability to be morally good at times. And if we construe the good as "the good life" and not simply as the moral good—that is, if we adopt a *eudaimonistic* view of the good—it surely matters that we give the body what it needs to support a good life. Perhaps even more basically, though, this paradigm I am exploring provides a way of seeing the body as good without it following from that that we ought to take its demands as things we ought necessarily to give in to. Even setting aside moral questions for the moment, our appetites need not even lead us to what is in our individual biological interest: our dietary appetites may lead us to a diet that causes us discomfort, lack of mental acuity, and early death, and places great indirect costs upon our loved ones and neighbors. But this does not mean that these appetites are bad in themselves. It is better to say that they are decently crafted tools that are not perfect and that work better in one environment than in another. Consider by analogy the incredible difficulty of suppressing the suffocation reflex that bids us take a breath, even when we know the air we are breathing is poisonous. It is a good thing that we have this reflex, lest we all suffocate; but in the wrong environment it can kill us.

If this is true of biological needs, it is all the more true of ethical considerations. It is a biological fact that I have a mechanism that urges me to eat steak and candy. (This instinct seems to be the overriding factor in Belgian and English cuisines, so far as I can tell.) But what amount of this is conducive to a good and healthy existence? Moreover, how was this meat acquired? Were the cattle kept in cruel conditions? Is this delicious meat perhaps dolphin or chimp meat? Am I eating more than my share of the world's protein supply while others are starving? Was this candy produced by slave labor? The passions, when not entirely dysfunctional, are a very limited sort of "good-detector"—they are suited to picking out things that are, in some amount, good for us, and to pick them out in particular conditions. As stated in previous sections, they are not perfect detectors even in optimal conditions—some things that seem good may be poisonous, for example. But the fidelity of the detectors decreases dramatically as the conditions diverge from the optimal. And, moreover, the detectors only indicate *prima facie* goods. You cannot expect God or nature to design animals with appetites that distinguish their own food from their neighbors'. For that you need a separate module that is sensitive to property rights.

This gives us a useful perspective on the passions: they are useful advisors, and often are experts in their own sphere. However, the wisdom they have is less

like the wisdom of Solomon and more like the wisdom of a pet dog. One hopes that the family dog is a legitimate expert on when there is a person (a dog, a skunk, a cat) in the yard. It is generally not an expert on how to interact with the mailman. We need to keep the passions in proper working order, but also to keep a perspective on how far their advice may be relied on.

Very similar considerations apply to many of the acquired personality traits that we call virtues. It is a familiar claim, at least since Hume, that what counts as a virtue is quite variable, and depends upon circumstances. What traits make a person useful or admirable in the heroic age of the *Iliad* are quite different from those that make him useful or admirable in a highly structured urban society. Achilles, the "best man" of the Achaeans, would not be a great man in modern New York or Tokyo. He would be a dangerous lunatic—prone to violence, willing to engage in kidnapping and murder over a friend's wife's infidelity, acting largely on the basis of voices that he hears. He would be the more dangerous precisely because of the "manliness" (*andreia*) that makes him a great warrior—his pride, his placing personal loyalties above laws, his desire for ascendancy, his appetite for physical conflict. Now if you are living in a small tribe that is always being attacked by other small tribes, it is in your communal interest to have some people like Achilles around. They are the kind of people you want on your side when the next tribe comes to steal your children or your cattle or Menelaus' wife. But these same traits are extremely maladapted to other situations, like those of a law-governed society living in close quarters in which physical threats are comparatively rare. A person who is a "hero" in the Greek sense may well be the worst kind of threat to civil society. His character traits are maladapted to his environment, whether as a matter of innate temperament or of conditioning.

6. THE DOG WHO BARKED AT EVERYTHING— DYSFUNCTION AND CONDITIONING

In graduate school, I lived next door to a man who owned a schnauzer. I soon came to the conclusion that schnauzers only bark under two conditions: when they see something move—and when they do not. Having met more schnauzers since then, I now think that, while there is a definite tendency in the breed to be high-strung, this particular dog was (whether by conditioning or unspeakable crimes against natural selection on the part of previous generations of dog owners) a particularly sorry specimen.

What had gone wrong? Dogs are by nature both social and territorial. They bark when something approaches their territory, and this both acts as a threat to the intruder and communicates a warning to the rest of the pack. Different breeds of dogs have pronounced differences in their inclinations to bark, ranging to the extreme of dobermans, which hunt silently. Schatzie (that was the schnauzer's name) was at the opposite extreme. He barked at everything—and at nothing. The mechanism within all normal dogs that causes them to bark in response to boundary violations had gone sadly (and annoyingly) awry. It was

set off, not only by humans or other dogs that approached his territory, but by squirrels and birds, by people he spotted half a mile down the road, by blowing leaves, and, in the end at least, by silence itself. The territorial instinct in him was hopelessly fubared. He was a doggie neurotic.

This was an example of a *dysfunction* of a standard-issue mechanism. In Schatzie, the territorial instinct had developed a hair trigger. Its input conditions had been hopelessly broadened until it was of no use at all, and probably caused the dog (and certainly caused its neighbors) untold stress. This kind of dysfunction happens in us as well: to take the closest example, there are plenty of human beings who falsely perceive all events in their environment as threatening, and many more who register false positives only some of the time, but far more than is good for them. Likewise for people who think that everyone wants to have sex with them, or hates them, or that inanimate objects have beliefs and desires, etc. There are other kinds of dysfunction as well. A mechanism can go wrong on the output side: most forms of neurotic behavior are some kind of transformation of the output side of a primitive animal mechanism. The anger we feel in response to an encroachment is displaced onto a less threatening target, or transformed into another, less threatening emotion, or internalized as stress; sexual desire can get fixated in fetishistic ways; and so on. Some such transformations are not harmful and are even adaptive; but others are quite dysfunctional. Likewise, normal kinds of outputs can be maladaptive and dysfunctional in degree: one becomes enraged at small slights, or falls completely in love with someone on the basis of a kind word or a smile.

7. THE JUNKYARD DOG OF THE SOUL

The preceding sections have all dealt with general and systemic features of organisms. I said earlier that it can also be fruitful to look at specific modules that may be specific to a species or shared across species. I shall conclude with a discussion of two of these. The first follows quite naturally from the preceding discussion of the barking dog. It is what Plato called *thumos*, a Greek word that has no adequate English equivalent, but is variously rendered as "anger," "incensive power," "spiritedness," or "warrior spirit." *Thumos* is, in short, what you see wide receivers displaying in the endzone after they score a spectacular touchdown, and is, as Plato quite naturally notes, the aspect of the soul that needs to be developed in order to be a good warrior. But Plato also makes some additional observations about *thumos* that are particularly astute, even for Plato. First, he likens it, and the class of guardians of whom it is the characteristic trait, to guard dogs. He even elaborates on this to suggest that, to fulfill its proper function, it must be *trained* so as to bark at enemies but be gentle with members of the household. The function of *thumos*, in other words, is to *protect our boundaries*—to respond with hostility to trespasses against our personal boundaries. (It is not accidental that boundary metaphors like 'trespass' form part of the basic vocabulary for sin in many languages.) Thus we feel *thumos* rise within us when there

are violations of our bodies, our personal space, our property, or anything that we are capable of thinking of as "our own." *Thumos* is, if you will, *the junkyard dog of the soul*.

Plato's second important observation is that *thumos*, like a dog (or a warrior ruler) needs to be properly trained so it shows hostility to the right things and not to the wrong ones. No matter how good the dog may be at deterring burglars, it is not a good watchdog if it attacks the baby too! And of course something of this is built into dogs, and arguably into *thumos*, as well: the family is fairly central to what is regarded as "one's own," and hence is protected rather than attacked. That is, it is protected if everything is functioning properly. Some dogs bark at everything. Some people get angry at the wrong things, or get angry in an inappropriate degree even at the right things (Aristotle, *Ethics* 1109a25).

Plato's third great observation about *thumos* is that, in us, it goes beyond the function of protecting outer boundaries and serves an additional, *regulative* role *within* the soul. In Plato's tripartite model, the soul can be properly balanced only if reason makes an ally of *thumos* (the lion head in Book IX) and enlists its help in domesticating the appetites. (In this capacity *thumos* is perhaps less like a guard dog and more like a sheep dog, keeping the much stupider bleating passions in line.) This line of thought has been most notably explored by the Greek Fathers of the Christian church—some of whom, like Evagrius (cf. 1970 trans.), were strongly influenced by Plato by way of Origen and the Cappadocians—and the line of Christian spirituality that extends from them into Eastern Orthodoxy. (See, for example, Theophan the Recluse's *The Spiritual Life*.) It has perhaps also been rediscovered by certain currents in modern therapeutic psychology that point out that we can often "say no" to the "negative tapes" playing in our heads, in a fashion that is remarkably reminiscent of the older Christian notion of the casting out of demons.

There are important therapeutic insights here. First, there is at least a partial answer to the question of how we ought to regard *thumos* in ourselves and others. We are too easily swayed to the extremes of (a) thinking anger unconditionally a bad thing, or else (b) letting it have control over us. But *thumos* is neither a bad thing nor is it itself sufficiently intelligent to be a guide to conduct. On the one hand, it is, in biological terms, a good and even a necessary faculty. An organism of our complexity that cannot detect and respond to boundary violations is perilously flawed. I at least am inclined to think that a person who has too great a deficit or an inhibition about responding to boundary encroachments is in some sense *morally* flawed as well, as (a) personal integrity is an important kind of virtue, and (b) our repulsion of boundary encroachments by others is an important social mechanism used to teach others respect for boundaries. On the other hand, as Aristotle liked to point out, feeling *thumos* is easy, but feeling it in the right amount in response to the right stimuli (and finding the right way to express it) is difficult. In biological terms, *thumos* is a faculty that goes fairly far down the phylogenetic spectrum. It is a rough-and-ready special-purpose mechanism. It is not—and arguably could not be—pre-programmed with appropriate responses to *all* of the many and varied situations an organism could find itself

in. And, once again, it can become dysfunctional due to organic or environmental factors. In short, it needs something else to guide it and train it—and at least part of this "something" is, as Plato suggested, a more flexible, intelligent, and general-purpose faculty of reasoning. The moral is that *thumos* has its proper functions in both the biological and the moral sphere; but in order to fulfill those functions, it must be kept in proper working order and guided by more discriminating faculties.

A second sort of therapeutic insight is provided by the suggestion that *thumos* can have a regulative role in the inner economy of the soul. Indeed, we all have the experience of being angry at some of our own thoughts and feelings. But not all such episodes are therapeutic. Indeed, the *misuse* of anger against our inner selves is arguably one of the chief impediments to inner thriving. We go wrong when the passions run amok and go unchecked, but we go equally wrong when we hate some fundamental part of ourselves. However, we are now in a position to see the beginnings of a path through this swamp:

1) As Plato points out, *thumos* needs a rational understanding of what is good for the soul in order to be properly deployed in checking the passions.

2) As Plato's model also makes clear, the role of *thumos* is *not* to *destroy* the appetites, nor to starve them, but to domesticate them, and keep them in line.

3) One obvious case in which the application of *thumos* is called for is when one or more of the appetites is usurping the function of pilot for the whole ship of the soul. This is an ancient insight—that appetites that threaten to become tyrannical require some inner coercion. Unfortunately, they generally require healing or redirection as well, and those who rely *solely* on *thumos* often find themselves doing what they hate.

4) But our analysis here lets us see more: the appetites and emotions are special-purpose mechanisms. They are biologically good, and also to be cherished because they are a part of us, and they cannot be hated without hating ourselves. However, even when they function properly in their own right, they are *stupid* (non-intelligent), and the goods they are programmed to aim at are often not our moral good or even our biological good. Contrary to Callicles and others who think that "the good" is equivalent to "what I want," biological mechanisms aim at goods that they may misrepresent. There are animals such as ants, and even mammals like Wolverines and perhaps some kinds of ferrier, that will attack anything, regardless of its size or the danger it represents. *Thumos*, or its equivalent, has run amok in them, and misrepresented what is good for the individual. Sometimes it is better to run than fight. (Or to say no to food or drink or sex. Pick your own passion.) Thus Aristotle points out the distinction between *cour-age* and *rashness*.

One therapeutic application is that, while reason may not be able to completely *turn off* the mechanism that makes battle or food or sex seem desirable, it can *put them in context*. The passions each know only their own goods, and loudly and persuasively proclaim them when they are set off. They are often very effective lobbyists and admen. (It is perhaps not accidental that rhetoric and the passions are so closely linked by Plato in the *Gorgias*!) One thing we can do,

though, is to see their claims, which arrive couched in the language of an absolute good, in terms more relative to their function. I smell rich food and a mechanism that is optimized for fat and protein goes off. What it says is, in some sense, true: I *am* an animal that needs fat and protein. But this is all that mechanism knows, and it reminds me of it every time it is exposed to the smell of a steak. It is like a dog that barks for food whenever it hears the can opener. What can I do about this?

First, I must recognize what is going on, and make an assessment of my overall situation: *do I really need to eat now? Is this the right time for fighting (drinking, mating)?* Second, if it is a time to show restraint, I treat my appetite as I would a pet or a petulant child: I say something on the order of, "Yes, I know you're hungry, and we'll get something to eat eventually, but not just yet." That is, I acknowledge the goodness of my appetite, make concessions to the fact that it cannot really understand very much, and try to gently put it off. Third, if this is ineffective, I can use *thumos* to provide some negative reinforcement in attempt to dissipate the urge. (This is done in the tone of "No, Rover, sit!" rather than "Baaaaa dooog!") Fourth, over time I can employ longer-term strategies to try to *shape* the nature of my appetites. Perhaps through conditioning and associations I can lessen my response to fattier meats and encourage my response to leaner sources of protein. I can desensitize myself to spiders and learn to be more wary about running up my VISA bill, etc.

5) If appetites and emotions are stupid and monomaniacal even when functioning properly, they can also be dysfunctional. My appetites can function all the time, or not often enough, or in response to the wrong things. It is possible to examine and sometimes to treat dysfunctions by various means, including psychotherapy and self-examination. Things like Freudian mechanisms of displacement and transference are useful diagnostic and therapeutic concepts here, and can be supplemented by our more biological model.

8. THE JOB OF SEX

Our final topic is sex. With the possible exception of altruism, no aspect of animal behavior has received more attention from the Darwinian standpoint than sex and the behaviors surrounding it. This is not due merely to the prurience of ethologists: the drive to reproduce is central to the whole Darwinian schema of biological explanation. This has often been badly distorted into the claim that passing on one's genes is our central *duty*. This, of course, is nonsense: the reproductive "imperative" is not a matter of duty, as (1) it is present in many animals incapable of duty, (2) reproduction often does not coincide even with individual biological interests, (3) it is itself *nonmoral*, and (4) in particular instances even conflicts with the moral. (Any *prima facie* duty to reproduce would be defeasible, for example, in light of a duty to not commit rape.)

All the same, the reproductive urge—that is, the output of mechanisms within us that lead us to pursue sex—often *presents itself* as an imperative, even as a categorical (i.e., unconditional) imperative! And indeed in many cultures

the drive to perpetuate the family genes is transformed into at least a *social* duty. When we are caught in the grips of it, it is among the most difficult to get ourselves out of by rational means.

The richness of human associative mechanisms and the lack of a specific breeding season have resulted in the fact that for our species (perhaps alone among terrestrial species) sex is also thoroughly interlaced with much of the rest of our lives. Much has been written around the subject of the perplexing, comical, and frightening aspects of human behavior: its relation to love, nurturance, self-esteem, security, status, jealousy, conflict, dominance, violence, and many other aspects of our lives; the notorious mismatch of the sexual and personal drives of women and men; the difficulties men and women have in understanding one another; and so on. I believe that from a purely theoretical ethological perspective, evolutionary explanation can shed much light upon our peculiarities. It makes sense that selection would favor such otherwise peculiar-sounding traits as:

- sexual arousal mechanisms that are sensitive to any breedable partner, and indeed mechanisms that respond to much more specific cues such as primary and secondary sexual features, behaviors characteristic of mating displays, pheromones, etc. (In humans, these are subject to cultural specification as well, so that, for instance, some cultures find obesity arousing while others find it an inhibition to arousal—of course there are "deviants" within any culture as well.);
- mechanisms in males that maximize the distribution of their genes by spreading them around (philandering, playing the field);
- mechanisms that lead us to favor our own children;
- mechanisms in females that maximize the prospects of their genes by trying to keep the father of their children around. (Not just *any* male will do as well, as there are mechanisms that lead us to favor our own children. Unfortunately, this may be a special-purpose mechanism that is not easily disengaged by the known abusive behavior of the father.);
- mechanisms that allow us to extend nurturance to others' children, especially those who are linked by either tribal bonds or some perceived similarity to oneself;
- preferential breeding with alpha members of the group, particularly common in social mammals, and hence intermix issues of sex and the various factors involved in dominance among intelligent mammals (strength, intelligence, capacities for bluff and deception, aggressiveness, charisma);
- mechanisms that aim at genetic diversity through sexual novelty (in DNA studies of chimp tribes, it turns out that the majority of young are in fact sired, apparently on the sly, by peripheral males);
- mechanisms that guard one's genetic prospects by restricting competitors' access to the breeding population (jealousy, territoriality).

These traits all seem plausible in biological terms in the simple sense that one can see how, if they emerged in the first place through mutation, they could easily

propagate at a statistical level through a population, as those who possessed them would have a statistical edge in passing on their genes. One may note that some of these traits seem to be in opposition to one another. But this need not be either surprising or problematic if we assume we are talking about a cluster of special-purpose traits rather than a general-purpose reasoning apparatus. It is not self-contradictory, and surely not impossible, to have special-purpose mechanisms that compete with one another, though it may be comical or tragic. (Indeed, here may lie the origins of both comedy and tragedy.)

In the case of sex, I think it is in the comic aspects of our situation that the therapeutic touchstone is to be found, for if we do not view our sexual nature in comic terms, it will likely lead us to tragic actions. As with all of the passions, we must navigate between the unacceptable extremes of hating our own natures and taking them as a good guide to conduct. Both of these extremes are particularly difficult to avoid in the case of sex. Witness on the one hand the perverse view, found in many ages, at least in the West, that just about any action is excusable if it is done out of romantic love. (Courtly love, in which a young man courted an older married woman, was a particularly depraved form of this.) Witness on the other hand the equally perverse views that sexual desire is necessarily evil, the widespread misconception that the original sin of Adam and Eve had anything to do with sex, and the vilification of women as the cause of men's sexual temptations. Sex is so powerful a drive that it is difficult to deal with. It is difficult to keep from acting upon the drive, and even if we refrain from acting on it, it is difficult to keep from *experiencing* it in unwanted ways. It drives us to deeds, or at least thoughts, that in the clear light of day are hateful to us. And hence we can come to hate the part of ourselves that is responsible for our grief. It is clear that for many, like Augustine, these desires (especially once indulged) have caused considerable pain, and one might reasonably wish to be rid of them. Yet if one thing is clear from centuries of ascetic literature, it is that these desires are things we cannot be rid of—except perhaps for some rare individuals in whom they are missing due to some organic deficit—in spite of the tendencies of moralists over the ages to blame us for our desires. It is simply not among our options to stop being animals, and animals have a drive towards reproduction.

But what insights can we glean here?

- 1) Reproduction is a biological drive whose goal is the biological good of a population. This need not coincide with the moral or even the biological good of the individual.
- 2) Sexual urges are the manifestations (the outputs) of any number of special-purpose mechanisms that have been selected for what they contribute towards reproduction.
- 3) Thus, sexual urges are far too narrow in their scope to be a good guide to behavior.
- 4) These special-purpose mechanisms, even when they are functioning properly at a biological level, are not errorproof. A person can be as aroused by a picture or a memory of a conspecific as by their actual presence. Sexu-

ally immature conspecifics may manifest behaviors associated with mating without themselves being ready to mate. The cause of the arousal (a pheromone left in the air, a previous encounter, nonspecific arousal caused by exercise, surprise, or fear) may not be associated with the person who occupies one's attention, and hence the association of arousal and potential partner on your part does not necessarily indicate that the person you have designs on is in fact a willing partner. Nor does the fact that a person wears clothes you find arousing mean that she or he wants to sleep with you.

- 5) These mechanisms can go askew in various ways: They can be too narrowly fixated (redheads), or too broadly (mammals), or both at once (the fetishist fixates only on women with garters, and even finds the garters themselves his primary object of desire). They can become inflexibly and maladaptively associated with other emotions (fear, anger, aggression, dominance, jealousy) or with formative developmental experiences (Oedipus and Electra complexes). They can be compulsive or too easily aroused (nymphomania, satyriasm) or inhibited. And so on. These sorts of things can often be diagnosed and treated through traditional modes of therapy.
- 6) The intensity and urgency that nature has programmed into these mechanisms (as well as their apparent ubiquity in human situations) require that we treat them with especial suspicion, much as we would the boy who is known to cry "wolf!" on the least provocation (or none at all). This is the more true given the very real moral and social consequences of rape, adultery, untimely pregnancy, and even sex without love. *It often requires that we develop special psychological and social measures to regulate conduct and even thought.*
- 7) But since there is a level beyond which sexual response is beyond our control, it seems *inappropriate to direct moral censure towards its mere appearance as a reflex*. For most post-adolescent and pre-geriatric humans, exposure to primary or secondary sexual traits of a conspecific of appropriate sex will produce arousal. It is not clear that we can any more prevent this than one can short-circuit the snake reflex that goes off in the kitten when a piece of string is dragged in front of it, though of course we can expect some control to be exercised both in the behavior that follows this reaction and in preventing the situation from arising at inappropriate times. Making people feel ashamed for being aroused is tilting at windmills, and moreover it directs attention away from the things we can do something about and adds the new problems engendered by shame. Hating or despising ourselves as passionate beings does no therapeutic good. (Though it may cause one to seek divine salvation instead of or in addition to therapy.)
- 8) We can, however, act therapeutically on what happens after specific arousal has taken place. One way to begin is to see our situation as comical: to recognize that what is going on is that our body produces in us (for good biological reasons which may not be good for us individually) reactions and desires that are often ridiculous, inappropriate, and even contradictory. It is not unlike always being in the middle of one of those "Funny Animal

Video" shows on television. This is the attitude of St. Francis, who referred to his body as "Brother Ass" because of its ridiculous behavior. The comic stance is, I think, more potentially productive than the shame stance. At best, shame can inhibit behavior and even later stages of thought, but usually at the costs of repression. Comedy can do these things as well, without the cost of repression (laughter releases pressure rather than causing it!), and opens the door to further therapy. It also allows one to have a positive attitude towards one's own nature. One may be embarrassed by a silly animal, but one is not ashamed for it, because animals are not proper subjects of shame. Our mechanisms of attraction are not themselves more than animal. They may be causes for embarrassment, but not for shame.

9) This being said, it is necessary to have ways to carve out space for wiser faculties to decide what kind of conduct is actually in order. How this is best done (the carving out of space, not the moral deliberation) will depend a great deal on the pathology of the particular passion. If a passion tends to be overwhelming and tyrannical, like lust or jealousy, it tends to inhibit other modules that might represent other goods. (Indeed, in lust and jealousy, even some of the senses can be dimmed.) In these cases, it is important to have a plan beforehand as to how one will act if one should ever get into such an overwhelming situation, just as in the martial arts one learns some moves that one will reflexively perform in particular situations in which deliberation will not be possible. One programs oneself to think of one's spouse and children, or of some image incompatible with desire, or to use *thumos* to rebuke the desire, or to take a cold shower or engage in some other distracting behavior (some ascetics took to whipping themselves or rolling in nettles, though of course this may not work for the fetishist!), or one learns to simply "put the clutch in" so that thoughts do not turn into actions.

9. CONCLUSION

The purpose of this essay has been to present a new framework for viewing the passions. Despite its length, this piece has done little more than to sketch some of the territory. The passions provide vast amounts of material for personal reflection—far more than I could cover even in a small book. My purpose here is to plant the seeds of an idea that may bear moral fruit in the reader, and to give a few examples of how to cultivate the idea in the case of a few selected examples. My hope is that this will provide a model for further reflection that can also lead to productive moral change.

NOTES

1. I had originally used the term 'soul' here, but found that it conjured too much theological and metaphysical baggage for several early readers of the manuscript. Unfortunately, the alternative terms, such as 'self' and 'mind' are also not without particular associations.

2. It is an interesting question why moral psychology seems to receive so much less attention from early modern philosophers, though with the notable exceptions of Hume and Kant (in his writings on moral anthropology, not those on the metaphysics of morals). I suspect it has to do with the return to popularity of the decision-theoretic model of practical reasoning and action.

3. I am not sure who deserves credit for first putting forth this idea. I first heard it from Professor Stephen Grossberg of Boston University, one of the pioneers in neural networks and complex systems, while on sabbatical from Wesleyan in 1993 at BU's Center for Adaptive Systems. Thanks to Grossberg, the Center, and Wesleyan for making that venture possible.

4. This notion of the "stability-plasticity dilemma" is one I derived from studying with Stephen Grossberg.

REFERENCES

- Chomsky, Noam. 1980. "On Cognitive Structures and Their Development: A Reply to Piaget." In *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, ed. Massimo Piattelli-Palmarini. Cambridge, Mass.: Harvard University Press.
- Evagrius the Solitary. 1970. *Praktikos and Chapters on Prayer*. Translated from the fourth-century Greek text by John Eudes Bamberger. Spencer, Mass.: Cistercian Studies Series, Number 4.
- Flanagan, Owen. 1991. *Varieties of Moral Personality*. Cambridge, Mass.: Harvard University Press.
- Fodor, Jerry. 1983. *The Modularity of Mind*. Cambridge, Mass.: MIT Press/Bradford Books.
- Lorenz, Konrad. 1952. *King Solomon's Ring*. Trans. M. K. Wilson. New York: Crowell; London: Methuen.
- Lorenz, Konrad. 1965. *Evolution and the Modification of Behavior*. Chicago: University of Chicago Press.
- Midgley, Mary. 1995. *Beast and Man: The Roots of Human Nature*. Rev. Ed. New York: Routledge.
- Plato, *Republic*. In *Plato in Twelve Volumes*, Vols. 5 & 6, trans. Paul Shorey. Cambridge, Mass.: Harvard University Press; London: William Heinemann Ltd., 1969.
- Pylyshyn, Zenon. 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, Mass.: MIT Press.
- Theophan the Recluse. 1995. *The Spiritual Life and How to Be Attuned to It*. Translated from the nineteenth-century Russian text by Alexandra Dockham. St. Herman of Alaska Brotherhood.
- Tinbergen, Niko. 1951. *The Study of Instinct*. Oxford: Oxford University Press.
- Tinbergen, Niko. 1953a. *The Herring Gull's World*. London: Collins. Reprinted New York: Basic Books, 1961.
- Tinbergen, Niko. 1953b. *Social Behavior in Animals*. London: Methuen.
- Tinbergen, Niko. 1968. *Curious Naturalists*. New York: Anchor Books.
- Woolridge, D. 1963. *The Machinery of the Brain*. New York: McGraw-Hill.