EUJAP | VOL. 2 | No. 1 | 2006
ORIGINAL SCIENTIFIC PAPER
UDK: 1:17
17.021.2

RESTORING EMOTION'S BAD REP: THE MORAL RANDOMNESS OF NORMS

RONALD de SOUSA

University of Toronto

ABSTRACT

Despite the fact that common sense taxes emotions with irrationality, philosophers have, by and large, celebrated their functionality. They are credited with motivating, steadying, shaping or harmonizing our dispositions to act, and with policing norms of social behaviour. It's time to restore emotion's bad rep. To this end, I shall argue that we should expect that some of the "norms" enforced by emotions will be unevenly distributed among the members of our species, and may be dysfunctional at the individual, social, moral, or even species levels. I"ll adduce three considerations in support of that pessimistic view: The fallacy of adaptive fixation, the moral randomness of group selection, and the lack of fit between "natural norms" set up by evolution and those moral and social norms we would like philosophy to justify.

Key words: emotion, rationality, evolution, function, norm

Despite the fact that common sense taxes emotions with irrationality, philosophers have, by and large, celebrated their functionality. They are credited with motivating, steadying, shaping or harmonizing our dispositions to act, and with policing norms of social behaviour. It's time to restore emotion's bad rep. To this end, I shall argue that we should expect that some of the "norms" enforced by emotions will be unevenly distributed among the members of our species, and may be dysfunctional at the individual, social, moral, or even species levels. I"ll adduce three considerations in support of that pessimistic view: the fallacy of adaptive fixation, the moral randomness of group selection, and the lack of fit between "natural norms" set up by evolution and those moral and social norms we would like philosophy to justify.

First, talk of "fitness-enhancing" traits commonly assumes that if something is adaptive it will spread to fixation in a population. But fitness is often frequency dependent, implying that a stable outcome of selection may be far from opti-

mal from any normative point of view. The ratio of psychopaths in a population, or, for that matter, the ratio of males, may be cases in point.

Second, group selection is often credited with providing an explanation of the evolution of altruism. But altruism, in the relevant sense, characterizes fanatical suicide bombers or slavish devotion to tyranny. More generally, evolution does nothing for the sake of individuals, except accidentally.

That is part of a third, more general problem: the potential failure of natural "norms", enforced by emotions and/or by mechanisms of social sanction to fit in with the values to which rational deliberation might lead. While natural selection undoubtedly works on several levels or "units", individuals like you and me are not among its beneficiaries.

Here's my plan. I'll first briefly discuss what sorts of things norms might be, and what sort of enforcement of them one might have in mind entrusting to emotions. Specifically, I will contrast a narrow sense of prescriptive "norms" with a broader sense of the "normative" that embraces values, and sketch some ways in which emotions have or might be conceived to relate to both values and norms. Second, I"ll recapitulate the reasons for expecting fairly strong discrepancies between the objective biological functions a psychological or behavioral trait might serve, and the considered values that human beings might hold dear at the individual, social, and cultural levels. Third, I'll briefly confront some methodological scruples about what has been pejoratively called the adaptationist-innatist point of view. Fourth, I'll recap some reasons for thinking Nature doesn't always work to optimize anything. Fifth, I'll sketch three illustrative examples spanning a very wide spectrum of "normative" standards for which an equally wide range of emotions serve to guide and motivate conduct. The examples will range from emotions linked to epistemic strategies to the place of psychopaths in our overall conception of the normal and the normative. In between those extremes, I will look at some more arcane discoveries about the relation between the subjective emotional character and the underlying biological functionality of certain forms of sexual behaviour.

1. Norms and normativity: what are we talking about?

There is much talk in cognitive science of "normativity". Normativity is everywhere, frequently paired with the word "irreducible", with the implication that where there is normativity, naturalism is doomed. Normativity belongs to the category of value rather than fact, of *ought* rather than *is*. A naturalist theory of anything is one that is supposed to rest entirely on facts, not values, that is to say on how things actually are as opposed to the way they ought to be, or are supposed to be. As I'll argue below, a biological sense of normativity can be fully naturalised, but this doesn't tell us anything much about values.

One might question whether the terms "norm" and "normativity" don"t cover too diverse a range of things to be useful. In a forthcoming work, Ruwen Ogien and Christine Tappolet point out a number of differences between deontic norms and evaluative terms, and suggest that we should restrict the word "norm" to the former. Emotions relate to both; but perhaps they relate to them in different ways. Here are the more important contrasts drawn (Ogien and Tappolet forthcoming):

- A value is typically described in an indicative subject-predicate sentence; a norm, by an imperative or else one or a small set of a deontic expressions: "one must", "one ought" "it is forbidden", etc.

- Evaluative terms can be more or less "thick", where that term implies some empirical content: "harmonious", "depraved", "hideous", and "generous" are thick; "good" is thin. Deontic terms are intrinsically thin: to know that something must or must not be done is to know nothing about the properties of the "thing" in question.

- Only value terms admit of degrees. If I ought to do something, it's not the case that I ought to do it *to this or that degree*.

- Only value terms tend to have internal meaning relations to names of emotions ("disgusting"; "frightening"; "admirable"; "pitiful"). One can have a "feeling of obligation", which can admit of degrees; but the feeling as such is not specific to any particular kind of obligation.

- Norms apply to a more restricted domain of things within an agent's power: as the slogan has it, "ought implies can". Values extend to things that don't depend on some agent to whom the value applies. Thus a man might value beauty in himself or others but not have any means of procuring it.

- The direction of fit of norms is world-to-mind, that of values, mind-to-world. This holds even for those inclined to be projectivists about value: whereas a norm posits how the world should be, the attribution of a value to an object in the world amounts to a claim that the object really has that value.

- Values can be brought in to justify norms, but not vice versa (Mulligan 1998). Thus it makes sense to say: "Justice requires that the defendant be acquitted", but not vice versa.

These differences mark off the domain of *deontology* from *axiology*. The domain of Deontology is concerned only with acting, or refraining. Insofar as emotions are required only to motivate action, then, we can think of them as just pushing and pulling: swayed by the totality of your emotions and desires at any particular time, either you act or you don't. The pronouncements of deontology are essentially limited to what is forbidden, what is permitted, and what is required. In the sequel, when I refer to norms, I will mean to designate only strict deontological norms, as distinct from values.

Now if emotions are held to enforce norms, are we speaking of "norms" in the strict sense, or of the broader sense of the normative that includes values? Should we expect a distinction between emotions relevant to deontology and those that relate to values? It would seem that the appropriate emotions to enforce deontic norms will be those that pertain to what one ought to do, or should have done, or else ought or should not have done. That list is short. It's pretty much limited to guilt, regret, and maybe some emotion that we might refer to as the "sense of duty". Call this restricted set the "deontic emotions". Shame is sometimes included in the list, but it doesn't fit properly because shame is not limited in its objects to acts that lie in an agent's power. The broad spectrum of possible emotions, by contrast, confronts us with the complexities of the axiological domain, which varies only in terms of degrees, but also in terms of the dimensions in which those degrees might be measured. This contrast gives rise to the problem of "mixed emotions", and especially those mixed emotions that are responses to moral dilemmas. For deontological ethics, moral dilemmas are paradoxical: for it would seem that *I ought to do x* and *I ought to refrain from x* are contraries. From multi-dimensional axiological perspective, by contrast, mixed emotions aroused by a moral dilemma merely indicate the presence of incommensurable values (Greenspan 1978; Greenspan 1995).

Emotions admit of degrees, and their frequent link to evaluative terms suggests that they are closer kin to evaluations than to norms. But that doesn't show that they don't serve to enforce deontic norms. For though the norm may not admit of degrees, the influence of emotions on compliance (or rejection) of that norm could do so, in two ways. First, two norms could differ in importance. (Ogien gives the example of "don't talk with your mouth full" vs. "don't erase the hard disk in your mother's computer".) Pertinent emotional responses to the infraction of one or the other of these two prescriptions might well differ in intensity. Furthermore, prescriptive norms might be ordered, so that in case of practical conflict—when only lying could save a life, for example—the prescription *Act in such a way as to save a life* would trump *Don't lie*. And the emotional reluctance to lie may well be less intense than the emotional reluctance to allow someone to be killed. Secondly, that very ordering may itself be justified by the different *values* placed on truth telling and on life. If, as Mulligan (1998) noted, values *justify* norms, the degree of importance attributed to the value may seem to contaminate the prescriptive rule itself. But that, Ogien and Tappolet argue, may be an illusion. They adduce this example: suppose I owe $10000 in taxes and you owe $9000. If we both fail to pay, the punishment exacted by a criminal court for fraud could be proportioned to the value of the unpaid tax. So the gravity of the sanctions could be a matter of degree. A further question again (cf. Miller 2005) is whether the emotion of guilt you and I experience observes, or should observe, the same proportionality. But although you owe less than I do, you are not *less obligated* to pay.

In sum, then, there are three factors that could explain the temptation to find differences of degree among deontic norms: (1) gradations in the values that underlie and

justify the norms; (2) a ranking in importance of different norms; and (3) degrees of gravity in the sanctions prescribed for infraction of the norms. None of these amounts to a reason to attribute degrees to prescriptive norms in themselves.

Even if this is right, however, there are other variable aspects of norms that don't settle neatly into this dichotomy. At a very early age, children make a clear distinction between conventional and harm-based norms. From the first, but not the second, a change in the dictates of authority suffices to relieve us (Turiel 1983); (Nichols 2004). One might expect the emotional discomfort involved in infringing the latter to be more powerful than is involved in the infraction of merely conventional norms. But that's an empirical question. And actually it seems the emotional barriers against transgressing purely social norms are extremely powerful. The evidence is in Stanley Milgram's famous "subway experiment", in which he got graduate students to go up to people in an uncrowded subway and ask them to yield their seat.[1] The most striking observation that emerged from this was not so much the fact that nearly 70% of those asked willingly gave up their seat, but how extraordinarily difficult it was to ask. Yet it's worth remarking that whatever norm the experimenters were violating in the subway was one that is never taught, discussed, justified, or even made explicit. I'll return, towards the end of this paper, to cases of norms—if they can still be called that—that are completely hidden from normal awareness.

Another type of case that's hard to classify is the norm of *fairness*. Is it a value or a norm? A situation can be more or less fair. As a value, then, fairness seems to admit of degrees. In the *ultimatum game* in which one of two subjects gets to make a take-it-or leave-it offer to the other, the subject receiving the offer has an economic interest in accepting any sum greater than zero. It has been shown, however, that subjects are willing to give up surprisingly large sums of free money for the sake of punishing a radically unfair offer (Oosterbeek, Sloof and van de Kuilen 2004). But the actual response in the ultimatum game has to be Yes or No, so while there can be a gradual increase in the force of indignation, there is just a threshold where it snaps, and there the agent may simply say: "This offer *must* be rejected."

What then exactly is the relation of the emotion to the norm, and what is its relation to the value? If emotions *justify* judgments of values and values, in turn, can justify deontic norms, then emotions are twice removed from norms in terms of justification. So on the standard assumption that emotions can directly motivate behaviour, it is interesting to note—though I don't quite know what to make of it—that they can *justify* that same behaviour only indirectly.

Not every emotion justifies a value, nor does every value dictate some deontic norm. Many philosophers, however, have identified some emotions as specifically "moral"

[1] Milgram and Sabini 1978. The experiment has recently been informally replicated by reporters for the *NY Times*. See http://www.nytimes.com/2004/09/14/nyregion/14subway.html?ei=5090&en=cb9818cd9c7e70d2&ex=1252900800&partner=rssuserland&pagewanted=all.

ones. The sentiments of benevolence and compassion, dear to Hutcheson and Hume, march pleasantly in step with social virtues. One could cite similar thoughts from such widely different sages as Mencius and Aristotle. For the moment, however, I will simply ignore the delicate question of what emotions justify what values and when. I shall proceed on the assumption that there is no general and systematic formula for relating the justification of values to the functionality of emotions. This is not a merely evasive move, so much as a principled one. In the next section, I shall attempt to explain why.

2. "True" values and "objective" teleology

When Aristotle asked what this or that characteristic or disposition was for, he felt confident that an answer could be forthcoming in terms of the intrinsic good that would come of that characteristic for the kind of thing in question. And he was confident that one could discover how things were supposed to be in nature by observing what happened "always or for the most part" (Aristotle, *Met.* 6-vi). In the practice of biological sciences, this method is still of practical use. Studying the details of what actually takes place is essential to the discovery of functions. We must assume that things do in fact happen "normally" much of the time. In addition, we can sometimes observe the causal role played by some specific mechanism or component of a biological system in the accomplishment of the system's manifest role (Cummins 1975; Amundson and Lauder 1994). But the philosophical ground on which this method rests has crumbled with the discovery of evolution.[2] If species have arisen from a long dynamic process that has led from unicellular organisms to the elaborate forms of life familiar to us now, we are forced to infer that what usually happens is not a reliable guide to what is *supposed* to happen. Human beings are the outcome of a huge series of experiments, most of which were subjected to merciless destruction. At every stage of natural selection which brought our ancestors closer to *Homo Sapiens*, those of our ancestors who bore the innovative genome were what Aristotle would have regarded as freaks of nature.

Nevertheless, modern philosophy of biology has elaborated an essentially satisfactory substitute for the notion of inherent teleology. This is the dominant, if not quite undisputed, account of teleology provided by the *aetiological theory of objective proper function*. The basic idea is that the attribution of a function to certain effects rather than others can be justified without reference to any external goal or design. At most, one needs only the entirely general and quasi-tautological premise that the "goal" of life is to perpetuate itself. The aetiological conception of teleology makes use of the common-sense idea that functions, and teleological properties more generally, *explain*

[2] Admittedly this fact continues to be stubbornly denied in certain circles. The Aristotelian conception is still current in the Vatican, and with all those who oppose genetically modified products or artificial reproductive technology on the ground that these are "against nature". All these modern Aristotelians are one in spirit with the airline passenger in an old Gardner Rea cartoon who remonstrates with the flight attendant: "No thank you, I don't think Nature intended us to drink while flying".

the presence of the organ or phenomenon to which they are ascribed. They explain it not in terms of their efficient cause, but in terms of their effects. That conception also transposes into a different mode Aristotle's idea that certain non-existent effects (those he called *potentialities*) are privileged over others in that they only are *supposed* to take place, whether they end up doing so or not.

The most sophisticated elaboration of the aetiological theory of function is due to Ruth Millikan (1984; 1993). The following formula is a somewhat simplified version:

(P) An element X has the proper function F if and only if:

1. X results from the reproduction of an antecedent element Y;
2. Y effected F in the past, in virtue of properties (i....n) reproduced in X
3. X exists because Y effected F.

This formula accounts for the main intuitions behind the common notion of function, as used in biology: in particular, it allows us to understand the presence of an organ in terms of what it does, and to make the distinction between functioning and malfunctioning, as well as that between functions and accidental effects. It defines a concept of *objective teleology*. Whether or not something is a function may be very difficult or even impossible to discern in a particular case. The formula does not claim to supply an invariably useful criterion, but to specify the meaning of the claim of functionality.

What the aetiological conception of function does not do, on the other hand, is give us any reason whatever to think, as Aristotle did, that an effect is good because it's a natural actualization of a potentiality. If my "actualization" serves the purposes of alien replicators, it may or may not be conducive to values I hold dear. In the currency of replication, they alone are the ultimate beneficiaries of evolution. Each organism breaks the mould. My phenotype has never existed before. Genes, by contrast, like single-celled organisms, are effectively immortal. To be sure, their identity across time is not strictly maintained, since the very possibility of evolution rests on copying errors; but neither could evolution have taken place unless these errors were extremely rare. What is more, individual DNA molecules pass away like all material things. What subsists—as in the promise of Shakespeare's Sonnet LXV, *that in black ink my love may still shine bright*—is the *information* transmitted in the lineage of sexual cells from generation to generation: "The gene is not the DNA molecule; it is the transmissible information coded in the DNA" (Williams 1992, 11). So we could say that bodies fulfill their proper function in fostering the survival of genes. To put it more provocatively, *organism are organs of their genes.* Since individuals belonging to sexually reproducing species are never reproduced, the converse does not hold. Admittedly, the process we call "reproduction" suffices to maintain certain characteristics without which organisms would lack the fitness that enables them to transmit their genes. But organisms like ourselves are not capable of the sort of faithful reproduction that allows their copies to be reidentifiable across the ages.

3. Methodological scruples

I'm aware that thinking in terms of an evolutionary framework of this sort seems crazy to some people. It raises two red flags: adaptationism, and nativism. In an attack on Steven Pinker, Jerry Fodor (1999) urges that it's all right to be "nativists", but that we must reject adaptationism. His argument is a rhetorical question: "Does wanting to have a beautiful woman—or, for that matter, a good read—really require a further motive to explain it?" We don't *need* any explanation for, say, our innate propensity to love our children.

But we do. Titania, when she finds herself loving Bottom in the form of an ass, does not ask herself why. But we do; and actually her love at first sight is very much like every mother's. A spell had been cast:

> *What thou seest when thou dost wake,*
> *Do it for thy true-love take,*
> *Love and languish for his sake:*
> *Be it ounce, or cat, or bear,*
> *Pard, or boar with bristled hair,*
> *In thy eye that shall appear*
> *When thou wakest, it is thy dear.*

This fits the case of love of children rather well. Like Titania, a new mother's behaviour is controlled by being placed under the tutelage of an emotion. The origin of that emotion is chemical and mechanical, but the "love" it generates is perceived as wholly unlike anything chemical or mechanical. By its means, the will of another agent—in one case Oberon, in the case of mammalian mothers, the genes—passes into, or is expressed by, the "will" of the individual. Finally, while it does not necessarily have bad consequences, it *allows* them, as Oberon's wish manifests in the very next line:

> *Wake when some vile thing is near.*

Rejecting the demand for an explanation of the emotions we feel, and of the behaviour these trigger, on the basis of our failure to experience the *motivation* to propagate our genes, is no more sensible than claiming that since I'm not conscious of the tensing of my tendons when I *just* flex my finger, my capacity to do so requires no physiological explanation.

Although our emotional repertoire arises from the complex particulars of our experiences in infancy and beyond, those experiences would be meaningless and would form no emotional dispositions at all if what happens in our childhood did not arouse certain innate responses. In practical deliberation, the most obvious stopping points in a chain of justification refer to desires, typically linked to emotional responses. In turn, those desires and emotions have their roots in capacities for unconditioned response that can plausibly be supposed to be heritable, and subject to genetic variance. But what

is striking about those sorts of ultimate argument stoppers in emotional justification, is that while—at least in some cases—they make perfectly good sense as adaptations from the biological point of view, they are in no way acknowledged as such. Their biological function is completely divorced from their phenomenology. Genes (in the broad sense) have influenced our basic emotional dispositions—by which I do not mean stereotyped responses to specific stimuli, but the biasing of learning in certain ways that result in an increased propensity to go in for certain sorts of behaviour. But the "goals" of those genes don't coincide with the conscious aims or desires of the individual organism that pursues those goals. It is as if our genes, the more securely to secure our devotion to the tasks for which they have "designed" us, are better off concealing from their agents themselves the true purposes of the missions for which they were recruited. And that, of course, means that there is likely to be a screening off of the teleology of the behaviour from the conscious motivation. So if that is Adaptationism, it is significantly modified by the fact that it makes no inference about the desirability of the traits that result from the supposed "adaptation".

I will persist, then, in pursuing the hypothesis that our emotions sometimes enforce "norms" that do not necessarily correspond to what we consciously espouse. Furthermore, in the way that emotions evolved, we can see several good reasons for being ambivalent about their "functions".

4. The ambivalence of emotional functions

On standard conceptions of their functions, the emotions effect appraisals of important aspects of agents' life-affecting situations, and prepare the body physiologically for some appropriate response.[3] While all or most emotions are generally acknowledged to be intentional states, the fact that they are grounded in more extensive physiological processes than other mental states is crucial to their capacity to afford "action readiness" (Frijda 1986). This utilitarian view of emotions was already quite pithily expressed by Descartes, who wrote that

> the utility of all the passions consists only in their strengthening thoughts which it is good [the soul] preserve and which could otherwise easily be effaced from it, and causing them to endure in the soul. (*Passions of the Soul* §74)

—and pointed out, in the next sentence, why that virtue of emotions also explains the trouble they get us into:

[3] Some of the relevant situations are concerned with at least three of the traditional "four F's": (Fnourishment, Fescape, Fcombat, and Freproduction), though the states most directly concerned with the first, namely hunger and disgust, are often deemed too simple to count as emotions.

> So too, – he continued – all the evil they can cause consists either in their strengthening and preserving those thoughts more than necessary or in their strengthening and preserving others it is not good to dwell on.

This is just one of the ways in which the assessment of emotions by moralists has tended to be ambivalent. To each silver lining, its cloud. I'll look at some specific examples later. But first I want to stress three general grounds for questioning the extent to which our native emotional dispositions should be expected to be beneficial to ourselves.

Here is the general point that arises out of the fact that our natures have been conditioned by evolution: there is only a largely adventitious correlation between what is biologically functional in the objective sense and what is of value to me as an individual, living in a certain social context at a certain place and time. While masters have an interest in the life of their slaves, that interest extends only so far as they need to live to do their job. That interest does not extend to the welfare, happiness, or subjective interests of the slaves themselves. Similarly, evolution didn't happen to benefit any individual, except accidentally.

(i) Why McDonald's is God's own food

It is a remarkable fact that only at McDonald's restaurant can you take any small child, raised in any culinary culture, and be certain of the child's approval. Clearly, McDonald cheeseburgers and fries are what God meant us to eat. That's because when natural selection, or, as it is known in some circles, the Intelligent Designer, was doing its work, just such a combination of fat, protein, salt and sugar as McDonald concocts was the most likely to secure survival.

I noted that certain emotions have earned the title of social or moral emotions. In terms of the contrast between norms and values, one might surmise that empathy and compassion are best thought of as axiological, while guilt is more plausibly described as an enforcer of deontology. It is worth remembering, however, that the emotional consequence of transgression is not always negative. Transgression of a prohibition can be more of a thrill than failure to act on a value. Do those who get a thrill out of doing something forbidden get only milder pleasure from doing something they consider merely worthless or even bad, but not actually forbidden? I don't know of research on this, but it's a promising thesis topic for a behavioural psychologist. It might explain why the consumption of alcohol actually rose during prohibition above the level attained just before (Warburton 1932, pp. 23–26, 72). It might also suggest cheaper and more rational alternatives to the current War on Drug Users.

In any case, the standard story obviously has plenty of room for analogues of cravings for McDonald's. Murderous male aggression is one example. Martin Daly and Margot Wilson (1988) have shown that while the rates of homicide vary by a whole order of magnitude across the planet, the proportion of homicides that are committed by males everywhere, in the most peaceful as in the most violent societies, doesn't stray beyond

two or three percentage point of 95%. It's hard to deny that this must result from the genetic success of that particular strategy for getting one's way. So while one standardly thinks of emotions of guilt, shame and compassion as liable to enforce the norm against killing, there's another way to think about it. From the strictly biological point of view, murderous rage enforces a kind of norm of its own. Though neither a social norm nor a consciously endorsed one, it's a kind of biological norm of its own. Lust for rape, xenophobia, religious zeal, and perhaps even homophobia[4] may be other examples of emotional reactions that might have roots in a fitness enhancing strategy, giving them an objective functionality that has become out of step with values endorsed by civilized 21 century people.

(ii) Frequency-dependent fitness

Both adaptationists and anti-adaptationists tend to assume that if a characteristic is under genetic control, it will either prove beneficial or harmful and tend, respectively, either to fixation or to extinction. But this is an important mistake. The success of anything you do depends largely on how many others do it too. Or putting it into jargon, the fitness of many strategies is *frequency-dependent*. What we should expect, then, is not elimination from or fixation into a species' nature, in such a way that exceptions can simply be dismissed as malfunctions resulting from a deleterious mutation, but *polymorphism*. As was first vividly illustrated by Maynard-Smith's example about hawks and doves, what we should expect in many cases is that whether there is a pair of alternative strategies, each will be advantageous only providing it isn't too widespread. Thus the genes of a female have an advantage if most members of the population are male, because they get to spread more widely than those of the males in the next generation. But the same is true of males. Regardless of the obvious wastefulness of dragging along so many useless males, therefore, the two sexes will reach an equilibrium of around 51%M to 49%F at the time of conception, which makes (given that males are the more fragile sex) for equality at peak reproductive age (Fisher 1930). This is an *Evolutionarily Stable Strategy* (ESS) (Maynard-Smith 1984; Dawkins 1976). Maynard-Smith originally defines an ESS as a *mixed strategy* that can't be invaded. Strictly speaking, what counts in measuring the crucial ratio is the number of acts conforming to one or the other strategies. They can be distributed among agents in any way that preserves that ratio. In the case of sex, each organism generally chooses a fixed strategy. But in the case of hawks and doves, it is compatible with every individual acting always in the same way, either as a hawk or as a dove; or with all individuals adopting a mixed strategy in the ratio specified by the ESS; or any combination of strategies in between.

[4] The reasoning here rests on a (potentially controversial) finding that most homosexual acts involve bisexual persons, and that bisexual individuals (human as well as of some other species) have a higher net reproductive rate than exclusively heterosexual ones (Baker and Bellis 1995, 116ff.). Paradoxically, then, homosexual men would be seen by heterosexual men as particularly dangerous rivals for reproductive success.

In sum, there is a virtually infinite number of combinations of possible strategies that might implement a given ESS. It would be interesting to explore how a particular implementation of a given ESS is actually arrived at in any concrete case. Another promising thesis topic. And in the case that particularly interests social science—the mix of strategies of cooperation, defection, and punishment—the range of possibilities is even greater.

It is widely recognized that kin selection and reciprocal altruism do not suffice to explain altruistic acts between strangers in large groups. Here again, the advantages of cooperation and defection depend in part on what others are doing. Peter Danielson has noted that although many of the classical discussions vaunt some specific strategy (such as TFT), we should expect populations to be mixed, consisting of unconditional cooperators, free riders, and different kinds of reciprocators (Danielson 2002).

Robert Boyd and Peter Richerson have explored, in a number of studies, the possible role of punishment in securing cooperation (Boyd and Richerson 1992; Boyd, Richerson, Gintis, et al. 2003). But punishment is itself costly, and so punishers must themselves display altruism and are subject to temptation to second order defection. There are therefore a number of possible strategies, including pure cooperation without punishment; cooperation at the first degree and also in punishment; "moralistic" punishment of non-punishers, and even punishment without cooperation. These give rise to a number of different possible ESS's. In work based on Boyd and Richerson's, computer simulations by Ernst Fehr and Urs Fischbacher (2003) suggest that only second order punishment (punishment of non-punishers) succeeds in securing cooperation in large numbers.

But that only pushes back the problem, for what, in turn, secures the motivation for punishment at the second level? At this stage, it is appealing to resort to some sort of group selection, favouring a propensity to experience a primitive moralistic emotion. But as noted in the very title of Boyd and Richerson's 1992 paper, punishment allows the evolution of cooperation *or anything else.* Moralistic emotions may succeed in promoting the replication of genes in a certain sort of group, but there is no reason to think that this must be, from the point of view of any individual members of such groups, a Good Thing.

5. Illustrative cases

(i) Epistemic norms and epistemic emotions

My first illustration is deliberately chosen to be as far removed as possible from the sort of norms that tend to arouse passionate intensity. It concerns the choice of epistemic strategies.

In a recent issue of *Nature*, Jonathan Cohen and Gary Aston-Jones (2005) look at some findings by Angela Yu and Peter Dayan (2005) on the application to science of the trade-off between exploration and exploitation. Exploitation of known resources is safe but likely to yield diminishing returns. On the other hand, giving up well-trodden paths for the sake of exploration may yield a jackpot of discovery, but is inherently risky.

That trade-off is well known to students of foraging. An ant faced with an established path may either follow it, in the expectation of finding food where many others have already found it, or else strike out in an unexplored direction. The latter option is risky but will pay off, if not for the individual at least for the colony, when the original sources of food are exhausted (Johnson 2003). This is a good example, then, of a mechanism first instantiated at the most basic level of foraging decisions. What is surprising is that it can be directly applied in the context of sophisticated scientific cognitive strategies, where it appears still to be controlled by a combination of chemical triggers.

Subjectively, the tension between the relative security of "normal science" and the excitement of a potentially fruitful paradigm shift is experienced as a struggle between cognitive fear and cognitive greed. The balance between the tendency to explore and the tendency to exploit in the cognitive domain are apparently regulated in part by specific neuromodulators, associated with two distinct issues of uncertainty. The first concerns the *bearing* of a signal, the second, the signal's *reliability*:

> Acetylcholine signals expected uncertainty, coming from known unreliability of predictive cues within a context. Norepinephrine signals unexpected uncertainty, as when unsignaled switches produce strongly unexpected observations. These uncertainty signals interact to enable optimal inference and learning in noisy and changeable environments. This formulation is consistent with a wealth of physiological, pharmacological, and behavioral data implicating acetylcholine and norepinephrine in specific aspects of a range of cognitive processes. (Yu and Dayan 2005, p. 681)

Yu and Dayan go on to remark that there seem to be "a class of attentional cueing tasks that involve both neuromodulators and shows how their interactions may be part-antagonistic, part-synergistic" (Yu and Dayan 2005). And of course those sorts of situations are typically experienced, in humans, as giving rise to emotional states: the "fear" in the risk of being wrong; the "lure" of the unknown; the "disappointment" generated by scientific prospects that don't pan out. What Yu and Dayan's discovery seems to be telling us is that a chemical mechanism underlies, in part, both the phenomenology of emotion and the process of what we assume to be high-level decision-making. This raises an intriguing question: can we think of the emotions more or less consciously felt in these epistemic dilemmas as *policing* the epistemic strategies of scientific exploration and invention? If so, with people as with ants, there may be one or more ESS's, which are not necessarily optimal from anyone's point of view. Under what conditions would

we have reason to believe that the balance between the two strategies, either in the mind of an individual person or a population of ordinary people, is the optimal balance? Again, I don't know: there's yet another promising thesis topic.

Boyd and Richerson (2005) have looked at a related problem: the equilibria relating to the tendency to act on the basis of what one has learned from the available cues, as against the tendency to imitate others. The equilibria achieved, and whether they are beneficial or maladaptive, depends on a number of factors. In a stable environment, if enough people have adjusted their behaviour to the objective cues available, it will pay for most members of the community to imitate rather than attempt to interpret the cues or devise procedures on their own. The advantages of one or the other strategy are also frequency-dependent. As in the ant foraging case, they depend for each on what others are doing. If nearly everyone else is just copying the majority, then it might pay me to be somewhat more discriminating. But in some circumstances, if I imitate *more* often than the majority, then I will require a higher threshold of difference before I risk diverging from the common behaviour. And that may make me *less* likely to make a learning error (Boyd and Richerson p. 87). On the other hand, the risk involved in too much imitation is that the originator of the behaviour one is imitating may have made a mistake.

(ii) Unconscious emotional drives

In our Western society, monogamy passes for an accepted social norm. Yet there is fascinating data presented by Robin Baker and Mark Bellis, suggesting that a surprising amount of sexual behaviour is influenced by a complicated teleology of "sperm wars". Some of this, they hypothesize, may be due to the advantage some females may derive in distinguishing one set of characteristics in the male from whom they solicit a genetic contribution, from another incompatible set in the partner that want for cooperative child raising:

> a female, by soliciting copulation with her partner at infertile stages of her cycle but with the most favoured male at the most fertile stages, could contrive to retain her partner's services as a parent while having offspring with the more favoured male. (Baker and Bellis 1995, p. 151)

In support of this hypothesis, there is evidence that

> females not taking oral contraceptives copulate more with their partners (IPC's, [intra-pair copulation]) during the infertile luteal phase but more with extra-pair males during the fertile pre-ovulatory phrase. There is a significant positive association between EPC incidence and probability of conception (P=0.018). (Baker & Bellis 1995, p. 161)

That makes it less surprising, then, that blood group studies and more recently DNA studies world-wide place the reported incidence of paternal discrepancy as high as 10% to 15% (Baker 2006, p. 64; Baker and Bellis 1995, p. 200). What particularly interests me here, however, is how to explain the emotions involved in bringing about these consequences. Where emotions influence a type of behaviour one can, of course, take the Fodor line: obviously (EPC) is more exciting and spontaneous, so you forget about birth control. Who needs more? If there is an *obvious* folk-psychological explanation for the fact, there is nothing more to be asked.

But that seems to me absurd. The fact that an emotion provides an intuitive sense of understanding doesn't dispense us from wanting an explanation of the emotion itself.

Why is it that (EPC's) have the greater emotional urgency? Again, it seems reasonable to suppose that there is a kind of objective biological "normative" value here, wholly disjoint, of course, in our culture, from the "social norm" that enjoins monoandry, and that the intensity of sexual emotion involved represent the enforcement of that discrepant value.

(iii) Psychopathy: the other side of an ESS?

My final illustration concerns psychopaths. My suggestion here is simple. In view of recent work on psychopaths, one can make a much sharper distinction between what I shall call true psychopaths and others manifesting a pattern sometimes labeled Anti-Social Personality Disorder (ASPD, sometimes also referred to as "sociopathy").

Blair et al. (2005) propose a three-factor model, comprising antisocial behavior, narcissism, and an emotional factor. A high score on this last is what sets true psychopaths apart. Anti-social behaviour in non-psychopaths tends to trigger high levels of anxiety; but in psychopaths, anxiety levels are lower. They fail to manifest the sweating reflexes generally elicited from subjects who are about to undergo an unpleasant experience such as an electric shock. In a game of cards in which it becomes clear to subjects that an initially winning strategy actually leads to consistent loss, they show no tendency to reverse their chosen strategy. These peculiarities form a cluster, a "callous-unemotional" pattern, which can be identified in children as young as seven (Blair 2005, pp. 113-115, 122).

At the neurological level, the crucial factor seems to be traceable to the baso-lateral nuclei of the amygdala (BLA). The amygdala is involved in three distinct patterns of learning. Psychopathy does not affect the association of unconditioned response to conditioned stimulus (UR-CS). So psychopaths are perfectly able to learn in the usual cognitive sense. By contrast, the BLA is crucial to associations of conditioned stimuli both with *affect representations* (fear) and with *valenced sensory properties* (such as images or perceptions of suffering). Psychopaths are simply unaffected not just by the suffering of others, but also by the prospect of their own. As a result, they are incapable of passive avoidance learning.

There is independent reason to think that genetic factors play an important part in psychopathy. Blair et al. (2005, p. 30) cite evidence from Viding, Blair, Moffitt et al. 2005, for high heritability of the callous-unemotional factor that differentiates psychopathy from ASPD.

Now if psychopathy is indeed heritable, why should we not regard it as an alternative strategy? Once again, like the distribution of the sexes, we might just be dealing with a *natural norm*, which is simply not one that is beneficial to the sort of people we like to think we are, or the kind of societies we'd like to live in. Psychopathy may be just one facet of an ESS.[5]

6. Conclusion

There is much work to do to figure out just what norms are, and how they relate to the broader concept of normativity. But what the preceding considerations suggest is that some norms appear to be entirely hidden or unconscious, and they come to light only through the emotions that promote, impede, or enforce them. This is the case for the Milgram inhibitions, and also in the sexual examples from Baker and Bellis.

As for psychopathy, it illustrates the three other central themes I have stressed, all of which, I have suggested, follow from taking seriously the perspective of natural selection.

First, while it is possible to make sense, on the basis of the aetiological model, of the notion of an objective natural function, we cannot assume that anything represents an

[5] In the absence of any direct measure, we could approach the question of whether psychopathy is heritable on the basis of the following fact. James Blair cites an intriguing statistic from a study of prison inmates in various countries, reporting a higher mean incidence of ASPD (50-80%) for male inmates, as opposed to 47% world-wide (Blair, Mitchell and Blair 2005, p. 19). They further note that only 25% of ASPD affected males in the US meet the more stringent criteria for psychopathy. In all, then, one can calculate that about 13%-20% of US prison inmates are psychopathic.
Now it's also a fact that the Land of the Free imprisons about 5 times as many people per thousand as the world average (6 times as many as China, that bastion of human rights, and 7 times as the EU). But, as we have seen, the proportion of ASPD affected inmates in the world as a whole is lower. So there are two possibilities about the proportion of psychopaths in the world population as a whole. If the proportion of psychopathy to ASPD in other countries is the same as in the US, we would have to conclude that the US somehow produces at least five times as many psychopaths as the rest of the world. Given that the US is ethnically diverse, that would suggest that this is an effect of culture, not of genetics. The alternative hypothesis is that psychopathy is indeed an innate strategy, which over time in the EEA has settled into something like an ESS. If that is the case, though, the proportion of psychopaths in the population as a whole should be the same in the rest of the world as in the US. And this would be confirmed, if one could establish that in the prison population world-wide (and particularly in Europe) true psychopaths were found five to eight times more often than in the US. But since we know that fewer than half of male inmates in Europe show ASPD, this would mean that virtually all of them would have to be psychopaths. In Europe, then, the ASPD affected inmates would all be psychopaths, whereas in the US the 75% of ASPD affected inmates who are not psychopaths would have been, somehow, manufactured by their environment. This is awkward for the genetic theory of psychopathy, but would cohere with the suggestion that there exist necessary genetic preconditions that require, for the production of the full-fledged psychopathic strategy, specific environmental conditions.

advantage, or is fitness-enhancing, in such a way that natural selection will gradually foster it to fixation in a population. Behaviour is only useful or deleterious in the context of what others do.

Second, even in the context of a strategy that can contribute to an ESS, a trait selected for the sake of such a strategy is not necessarily beneficial to its individual bearer. This holds regardless of the extent to which the trait has been elaborated at a great distance from what is genetically or otherwise inherited, by processes of cultural evolution. For cultures are, to put it bluntly, just as much liable to be irredeemably nasty as is nature. Submission to savage tyranny, or to bloodthirsty rituals demanded by cruel deities, is also a kind of altruism, and has all too frequently proved to be the outcome of the combined wisdom of nature and nurture.

Thirdly, while I have argued that the aetiological concept of function allows us to make perfectly good sense of *what Nature intends*, in an objective sense (though often, of course, epistemically inaccessible), we should be, on the whole, concerned with what Nature intends only in the sense that it is always a good idea to spy on your friends as carefully as on your foes. What we want, and value, and intend, collectively and individually, though ultimately causally traceable to the machinations of nature, is now largely unrelated to it. That something functions as nature intended does not show that we should value it. We are, as they say, children of nature, or products of nature. But we are also the grand-children of unicellular organisms. By now, nothing on Mars could be more alien from us than those ancestors, or than the beneficiaries of that long process by which we have become human, over three and a half billion years, one monstrous freak at a time.

REFERENCES

Amundson R. & G. V. Lauder (1994), "Function without Purpose: The Uses of Causal Role Function in Evolutionary Biology", *Biology and Philosophy,* 9, pp. 443-469

Baker R. (2006), *Sperm Wars: Infidelity, Sexual Conflict, and Other Bedroom Battles,* New York: Avalon: Thunder"s Mouth Press

Baker R. R. & M. A. Bellis (1995), *Human Sperm Competition: Copulation, Masturbation and Infidelity*, London: Chapman Hall

Blair J., Mitchell D. & K. Blair (2005), *The Psychopath: Emotion and the Brain,* Oxford: Blackwell

Blumberg H., Hare A., Kent V. & M. Davies eds. (1983), *Small Groups and Social Interaction* London: Wiley

Boyd R. & P. J. Richerson (1992), "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups", *Ethology and Sociobiology,* 13, pp. 171-195. Repr. in Boyd & Richerson 2005, pp. 166-188

Boyd R. & P. J. Richerson (2001), "Norms and Bounded Rationality", in Gigerenzer & Selten 2001, pp. 83-97

Boyd R. & P. J. Richerson (2005), *The Origin and Evolution of Cultures*, Oxford; New York: Oxford University Press

Boyd R., Richerson P.J., Gintis, H. & S. Bowles (2003), "The Evolution of Altruistic Punishment", *Proceedings of the National Academy of Sciences (USA)* 100, pp. 3531-3535. Repr. in Boyd & Richerson, 2005, pp. 241-250

Cohen J. D. & G. Aston-Jones (2005), "Decision amid Uncertainty" *Nature,* 436, pp. 471-472

Cummins R. (1975), "Functional Analysis", *Journal of Philosophy,* 72, pp. 741-764

Daly M. & M. Wilson (1988), *Homicide*, New York: A. de Gruyter

Danielson P. (2002), "Competition Among Cooperators: Altruism and Reciprocity", *Proceedings of the National Academy of Sciences,* 99, pp. 7237-7242

Dawkins R. (1976), *The Selfish Gene*, Oxford: Oxford University Press

Fehr E. & U. Fischbacher (2003), "The Nature of Human Altruism", *Nature,* 425, pp. 785-792

Fisher R. (1930), *The Genetical Theory of Natural Selection*, Oxford: Clarendon Press

Fodor J. (1999), "The Trouble with Psychological Darwinism. Review of Pinker, *How the Mind Works,* and Plotkin, *Evolution in Mind*", Http://www.lrb.co.uk/v20/n02/fodo01_.html

Frijda N. (1986), *The Emotions. Studies in Emotion and Social Interaction*, Cambridge / Paris: Cambridge University Press / Editions de la maison des sciences de l'homme

Gigerenzer G. & R. Selten eds. (2001), *Bounded Rationality: The Adaptive Toolbox*, Cambridge, Mass.: MIT Press, pp. 281–296. Repr. in Boyd & Richerson 2005, pp. 83-97

Greenspan P. (1978), "A Case of Mixed Feelings: Ambivalence and the Logic of Emotion", in Rorty 1978, pp. 139-161

Greenspan P. (1995), *Practical Guilt: Moral Dilemmas: Emotions and Social Norms*, New York: Oxford University Press

Johnson N. L. (2003), "The Development of Collective Structure and its Response to Environmental Change", *Semiotics, Evolution, Energy, and Development* 2, pp. 84-113

Maynard Smith J. (1984), "Game Theory and the Evolution of Behavior", *The Behavioral and Brain Sciences,* 7, pp. 95-126

Milgram S. & J. Sabini (1978), "On Maintaining Social Norms: A Field Experiment in the Subway", in: Baum, Singer, & Valins eds. *Advances in environmental psychology*, Vol. I, pp. 31-40. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Miller W. I. (2005), *Eye for an Eye*, Cambridge, MA: Cambridge University Press

Millikan R. (1984), *Language, Thought, and Other Biological Categories*, Cambridge, MA: MIT Press, A Bradford Book

Millikan R. (1993), "In Defense of Proper Functions", in Millikan 1993, pp. 13-29

Millikan R. (1993), *White Queen Psychology and Other Essays for Alice*, Cambridge, Mass.: The MTI Press

Mulligan K. (1998), "From Appropriate Emotions to Values" *Monist* 81, pp. 161-188

Ogien R. & C. Tappolet, Forthcoming. *Repenser les Normes et les Valeurs*, Paris

Oosterbeek H., Sloof R. & G. van de Kuilen (2004), "Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis", *Experimental Economics,* 7, pp. 177-188

Proust J. (1997), *Comment l'Esprit Vient aux Bêtes: Essai sur la Représentation*, NRF Essais Paris: Gallimard

Rorty A. ed. (1978), *Explaining Emotions*, Berkeley: University of California Press

Turiel E. (1983), *The Development of Social Knowledge: Morality and Convention* Cambridge: Cambridge University Press

Viding E., Blair R.J., Moffitt T.E. & R. Plomin (2005), "Evidence for Substantial Genetic Risk for Psychopathy in 7-year-olds", *Journal of Child Psychology and Psychiatry,* 46, pp. 592-597

Warburton C. (1932), *The Economic Results of Prohibition*, New York: Columbia University Press

Williams G. C. (1992), *Natural Selection: Domains, Levels, and Challenges*, Oxford Series in Ecology and Evolution. New York: Oxford University Press

Yu A. P. & P. Dayan (2005), "Uncertainty, Neuromodulation, and Attention", *Neuron,* 46, pp. 681-692

Received: October 4, 2006
Accepted: November 15, 2006

Department of Philosophy
University of Toronto
215 Huron Street
Toronto ON, Canada
M5S 1A2
sousa@chass.utoronto.ca