Really, what else is there? Emotions, value and morality

The heart of every scholar is a deep well, whence deeply buried emotions bubble up to the surface in the guise of arguments.

– Natalie Clifford Barney

Let me first betray a dark secret of philosophy. Despite its commitment to reason and argument, the broad lines of every philosopher’s positions are determined by innate temperament. The arguments pile up later, to justify temperamental convictions. That’s why conversions are about as rare in philosophy as they are in religion. No doubt it happens; but personally I’ve never known a Kantian to switch allegiance to utilitarianism or vice versa. By temperament, I suspect, one sort of philosopher aspires to emulate literature; another looks to science for clues and especially to biology; and a third is keen to apply philosophical analysis to the question of how best to live. What makes contemporary work on emotions particularly exciting is that much of it embraces all three tendencies. To one who by temperament regards philosophy as a licence to poach, that’s a big attraction. We can’t know it all, but we can at least ask how it all fits together.

I shall suggest in what follows that the answer is, by reason of emotions.

1 Willard V. O. Quine’s demolition of the wall between conceptual and scientific issues emancipated philosophers to see what we do as continuous with science.¹ That implies that we should seek to understand even our normative practices and convictions in naturalist terms. Whatever we care about, including beauty, morality and religion, is what elicits emotion. Emotions are psychological states, now found with increasing precision to be correlated with activity in specific regions of the brain. They are well suited to provide a naturalistic explanation of our commitments, to morality or anything else. They are also notorious for their power to disrupt our rational and moral resolutions. Nevertheless, I mean to argue that morality, no less than art and religion, is but an elaboration of our emotional life.
This approach is a form of naturalism, the very idea of which as applied to morality will meet resistance. The naturalistic approach arouses panic because it threatens to explain morality away, undermining the absolute authority of ethical norms. Champions of human dignity charge that science ‘reduces’ us to something less than human. And yet, I shall argue, ethical norms do indeed originate in our somewhat haphazard stock of emotions, a hodge-podge bequeathed to us by the vagaries of natural selection. Although that wouldn’t seem to fit them for normative hegemony, I shall suggest that this is no cause for nihilistic despair.

Aristotle thought a virtuous person would greet all situations with the right feelings, to the right degree, in the right way. Far from claiming that emotions ground our morality in any way, that places emotions themselves under the aegis of morality. By contrast, the Scottish theorists of sentiment – Hutcheson, Adam Smith and Hume – identified certain positive emotions such as sympathy and compassion as intrinsically moral, since they motivated us to behave in socially beneficial ways. More recently philosophers sometimes labelled ‘neo-sentimentalists’, such as Allan Gibbard, have focused on anger, guilt and shame as emotions most apt to enforce moral prohibitions. They have qualified this, however, by insisting that no emotion is moral in itself: what makes it moral is being endorsed. But endorsement, as recently argued by Jesse Prinz in a radical defence of relativism, is nothing more than an emotion that takes another as its object. And that doesn’t seem to touch the normative problem: for how are we to know when an endorsement is correct? Perhaps emotions merely motivate us to be moral when we’re lucky, just as they motivate bad behaviour when we (or those around us) are not. That would leave us still in search of a foundation for our moral opinions. But where to look for that?

The right answer is: Nowhere. All foundations proposed by philosophers turn out to be no better than God’s command. And that, as Plato already showed in the Euthyphro, is no help at all; for we need to know whether something is good because God commands it, or God commands it because it is good. The same goes for any foundational value or principle: Should I make an emotional commitment to it because it is correct? Or do we deem correct and call ‘valuable’ whatever we happen to approve of?

What makes it awkward to answer this question is that one of morality’s tasks is commonly taken to be the overcoming of emotion. For some – call them Stoics, though Kant comes to mind as well – that is morality’s principal function. For Stoics, the challenge is to ground the rejection of emotion in some convincing principle. But the problem
arises with equal force for those who hold that sound emotional responses are a touchstone of virtue: how are we to distinguish sound from unsound emotions? If we justify our emotions on the basis of values which are themselves mere projections of our emotional responses, we are spinning in a circle.

Confucius, Aristotle, Hume and Mill all thought themselves acquainted with living paradigm-paragons, junzi, kalos k’agathos, or gentleman, to whom they could defer as impartial arbiters, responding correctly to every situation. That could settle the question, if not exactly answer it. But in a culture like ours, where narcissism takes the form of anxiety about its own parochialism, it no longer convinces.

A naturalist might still look for a foundation in something deeper than any individual emotional repertoire, such as human nature. That was Aristotle’s approach; but the pitfalls are obvious. Much of what is universal among humans is unlikely to be unique to us: as we are frequently reminded, we share most of our genes with chimps and disconcertingly many with worms. So should we look instead to what is uniquely human? On that basis, Aristotle thought it man’s job to philosophise. But even philosophy doesn’t make everyone happy. What is unique to humans is unlikely to be true of all humans. Should we seek Human Nature in some sort of blended average of human traits? Research into human beauty works like that.6 But what works for ideal beauty might not work for morality. One problem is that my emotional repertoire may not match yours. Another is that on the basis of the ideal blend, any proposal for improvement must be seen as merely deviant. Worse, we cannot assume that all emotion is morally right.

One answer to such worries is to ground morality in a priori principles of Reason, to which emotion is ruled irrelevant. Another appeals to some substantive but not emotional intuition about the ultimate good. A look at either strategy arouses the suspicion that what philosophy professors routinely sell their students is sophistry, if not intellectual child abuse.

Kant’s deeply puzzling ‘categorical imperative’ is the best known strategy of the first type. It enjoins us to ask whether one could consistently will what we intend (the ‘maxim’ of our act) to become a universal law of nature. That tells us, it is claimed, that we shouldn’t (unjustly) kill, because we couldn’t will it to be a law of nature that everyone unjustly kills. But what is the sense of ‘could’ that makes sense of that? Why suppose that any given act must be performed either always or never at all? What if my maxim is: Do it just sometimes?

For some rules, such as those governing property, even universal violation might be willed: communism isn’t logically impossible. Kant’s
rule seems to work better with lying, since *universal* lying would destroy the trust that make lying effective. But here again – unless it’s going to be on the exam – the uncorrupted student will dismiss this supreme principle of Reason as just daft. (‘Move along,’ said the policeman trained in Kantian logic, ‘if everyone just stood around, nobody could get by!’) Sensible persons settle for lying *judiciously*. Actually that’s already pretty much a universal law of nature: deception among plants and animals is rife, just as long as it isn’t too frequent. Biologists call this a frequency-dependent fitness-enhancing strategy. Monarch butterflies are poisonous, and viceroyos borrow their livery without bothering to make poison. The deception works so long as there are enough monarchs to sustain credible deterrence.

The appeal to some supreme indubitable foundational value fares no better, even if we could trust the fiction that everyone intuits it, regardless of culture or temperament. John Stuart Mill is often mocked by philosophy lecturers for strolling with aplomb over the abyss between fact and value. ‘The sole evidence it is possible to produce that anything is desirable’, he asserted, ‘is that people do actually desire it.’ He is often accused of exploiting the ambiguity of the suffix that indicates worthiness in ‘admirable aplomb’ but in ‘impassable abyss’ signals mere possibility. In fact Mill was quite right. If you think no inference can be valid from fact to value, then *nothing at all* can provide ‘evidence’. Which is not to say that desirability follows from desire. A claim about ultimate value can only rest on some unsupported intuition. Once you have claimed the ground, you give up on the possibility of further grounding. And there is nothing else a ground-level intuition of value can be than an emotional response.

Whoever deems that last assertion outrageous must confront an increasingly compelling body of empirical research showing, subject to the usual fallibility of science, that genuine moral judgements *consist* in part in emotion. This proposition can be detailed in three claims:

First, people’s moral concerns elicit specific emotional responses, derived from plastic dispositions bequeathed by natural selection, but shaped by cultures to enforce different norms.

Second, what makes a normative judgement moral (as opposed to prudential, aesthetic or conventional) is just the strength of the emotional responses it expresses, together with the second-order emotion of approval directed at that emotional response itself.

Third, the competent articulation of moral judgements is powerless to motivate behaviour in the absence of emotional response.

After sketching the evidence for these three claims, I shall briefly draw some implications for the anti-naturalist concerns I described at the beginning.
On the basis of some influential experiments, Elliot Turiel\textsuperscript{8} claimed to find a clear distinction between moral and non-moral norms. Even young children, he claimed, distinguish prohibitions based on the avoidance of harm from those based on convention. His subjects judged that violations of conventional norms, unlike those designed to prevent harm, would not be wrong if an authority – teacher or parent or legislator – decreed them permissible. Critics showed, however, that it is only among liberal Western educated subjects that the line is drawn where Turiel placed it. Even those subjects, when presented with ‘harmless taboos’ (using a chicken carcass for masturbation, or eating a pet dog accidentally killed by a car), persist in viewing the transgression as absolutely wrong, despite being totally unable to justify their judgement. Jonathan Haidt calls this ‘moral dumbfounding’.\textsuperscript{9}

In some of the strongest ‘moral’ convictions, then, emotions both precede and swamp reasons. The chair of the US president’s Council on Bioethics has taken this as a virtue of moral intuition, praising the unreasoning ‘yuk response’ as the ‘wisdom of repugnance’.\textsuperscript{10} And Jonathan Haidt has suggested that the liberal conception of morality, by restricting itself to the areas Turiel sought to delimit, distorts what most humans regard as the proper sphere of morality. If that is so, we can ask whether we ought so to restrict the moral domain.

These issues were neatly outlined by Steven Pinker in a recent essay in the \textit{New York Times Sunday Magazine}. Pinker noted, perhaps with tongue in cheek, that a ‘law of constant moralising’ seems to ensure that as we remove some items from the list of immoral behaviours, others take their place.\textsuperscript{11} In liberal circles homosexuality and marijuana are now off the list, but eating meat, driving SUVs, and tobacco have replaced them. What frames this surmise is the idea that for most of humanity morality involves five relatively distinct domains of intense concern, grounded in specific ranges of emotional dispositions.\textsuperscript{12} About each of those emotional ranges we can tell a plausible evolutionary story, but the innate capacities for response proliferate into widely diverse emotions that enforce divergent cultural norms. The liberal morality of Western secular cultures has tended to confine the scope of morality to just two of those five domains: the \textit{avoidance of harm}, and \textit{fairness} or \textit{reciprocity}. Both afford clear illustrations of the innate nature of their supporting dispositions as well as the variety of ways in which they are actualised to fit into different cultural expectations.

What counts as real harm varies across cultures, which makes the exact scope of harm avoidance unclear; but the emotional building
block of the desire to avoid harm emerges in very young infants. It is a primitive form of compassion which some neuroscientists ascribe to ‘mirror neurons’ that fire both when a specific action is performed and when it is observed. As obstetric ward nurses well know, a single shrieking infant can cue a chorus. Beyond infancy, empathy is modulated by numerous cognitive and cultural factors: deserved punishment or a medical procedure won’t elicit as much compassion as an accidental injury. And of course you won’t feel empathy for something or someone outside the group to which your concern extends.

Fairness and reciprocity illustrate the same mix of innate dispositions refined by culture into divergent practices. Aristotle unexceptionably defined justice as giving to each what they deserve. But who deserves what? Different societies obviously give different answers; but here again one can glimpse a primitive emotional ‘module’ which seems to motivate a simple form of the concern for fairness. In the ‘Ultimatum Game’, one participant is given a sum of which they can arbitrarily offer any portion to a second participant. If the latter accepts, the money is allocated accordingly. If the offer is rejected, neither gets anything. In a wide variety of cultures, offers below 25 per cent to 40 per cent have been found to trigger frequent rejection, although rejection violates economic rationality. Direct inspection of the players’ brain confirms that rejection has an emotional cost. It seems to be driven by a desire to punish an unfair offer. In this, humans may be unique: for chimpanzees are apparently immune to the temptation of costly spite.

For many people in the world harm avoidance and fairness do not exhaust morality. The domains left out are community, authority and purity. These too are enforced by emotions for which evolutionary antecedents – and homologues in other primates – are not hard to discern.

In many parts of the world, for example, what Westerners disparage as nepotism is regarded as respect for Community. Preferring a complete stranger to a relative is unforgivably heartless. The underlying emotional disposition here is predicted by kin-selection theory. This rests on the fact that natural selection promotes not the welfare of individuals, but the replication of heritable patterns. Hence the attachment to kin which can sometimes motivate sacrifices. Loyalty morphs in many directions, from innocent pride in a home team to suicide bombing.

The emotional dispositions that support authority can plausibly be traced to dominance hierarchies observed in other social mammals – or in every office and every kindergarten. It too has lost much of its
intensity with the replacement of private or tribal revenge by state power. We are expected to ‘respect’ our boss and obey the directions of our local policeman, but we regard the lack of such respect as imprudent rather than deserving of moral indignation.

Authority is also in tension with the Kantian ideal of autonomy, as has been forcefully argued by Robert Paul Wolf. Kantian autonomy transcends the distinctions I have drawn between different domains: whatever your standards of good and right, what matters to a Kantian is whether your act stems from an autonomous ‘good will’, in a sense that escapes any merely empirical tests of freedom or motive but belongs to the ‘noumenal’ world. In the noumenal world, everything is what it is absolutely and in itself, not relative to the perspective of someone experiencing it. By the same token, it is all radically unknowable, except insofar as the noumenal will pierces the metaphysical membrane that hides it from view. But once you get over the awe this idea inspires, it is hard not to view it as a hypertrophic manifestation of the cult of purity. In Kant’s Kingdom of Ends, as in the realm of Platonic forms, we are insulated from the coarseness of fleshly life. The impression is not lessened when one reads Kant’s rantings about the abomination that is sex, which can be redeemed only by a contract of reciprocal genital ownership called marriage. When the comedian Butch Hancock summed up Texan Christianity, he might have been paraphrasing Kant: ‘Sex is the most awful, filthy thing on earth and you should save it for someone you love.’ The emotional ground of purity is disgust, which we can assume arose by natural selection to protect us from harmful parasites and germs, but which, like the other emotions discussed, takes widely different forms. Extended to metaphorical pollution, purity takes on religious dimensions, which subsist in many supposedly secular invocations of the ‘sanctity of life’ and ‘human dignity’.

The fact that our moral convictions are conditioned by emotions is not surprising. What may seem unsettling is the complete irrelevance of emotions to some of the judgements they influence. In experiments conducted by Thalia Wheatley and Jonathan Haidt, for example, arbitrary words (‘take’ and ‘ever’) were first associated with disgust in hypnotised subjects. In a later phase, subjects without recall of the hypnotic episode read a simple anecdote, in one of two versions differing only in that only one contained the words ‘take’ or ‘ever’. The result was that the action described in the story was judged more severely by those whose versions contained the tainted words.
Even about entirely innocent protagonists the subjects confabulated some reason for reproach, saying, for example, ‘It just seems like he’s up to something’.18

In light of all this, motivation can hardly be confused with justification. Is morality just a projection of emotion, subjective and relative? These two terms shouldn’t be equated. One sense of subjectivity implies illusory projection, but another is merely perspectival relativity. In being perspectival, emotions are like perception, which informs us by means of ‘subjective’ experiences about ‘objective’ states of the world which we home onto by cross-checking among sensory modalities and making rational inferences. Like emotions, perceptions can mislead. Yet we would lose our grip on the world if we dismissed perception altogether as illusory; similarly, the rich landscape of value would flatten into universal indifference if we were to discount our emotions. It is not incoherent to say that objective moral facts exist on the ground that people actually make moral judgements that are relative to natural and cultural facts about human emotions. Still, if some value judgements are so manifestly determined by irrelevant emotions, how do we identify those that are objectively correct?

One promising method is the quest for reflective equilibrium.19 In its most restricted form, this requires that we assess a moral principle in terms of the acceptability of its consequences, and specific actions in terms of the principles they might come under. Conflicts of principle – say, whether to lie to save someone from harm – are handled in the same way. The method ensures that no single emotional response will determine behaviour; but it also means, as in the proverbial mending of a boat at sea, that no reconstruction from scratch is ever possible: we can’t question all our intuitions and assumptions at once.

It also brings to light conflicts that further illustrate the role of emotions. A widely discussed example is the ‘trolley problem’, first discussed over thirty years ago by Judith Jarvis Thomson20 and recently picked over by psychologists and neuroscientists. Most people would approve of pulling a switch to derail a trolley from a track where five people would be killed, onto another where it will kill just one. Yet most would not personally push (or approve of pushing) one fat man onto the track to ensure the same result. Peering into the brain with functional Magnetic Resonance Imaging (fMRI) machines confirms the obvious: the emotions that dissuade us from causing harm are stronger the more directly we are involved in bringing about the effect.21

Some subjects affected by lesions in the ventro-medial frontal cortex prove to be exceptionally consistent utilitarians.22 But it would be hasty to infer that utilitarians generally suffer from deficient brains. Plenty of
repugnant thoughts lodge, alas, in undamaged brains. There is no virtuous Greek Gentleman on whom to model oneself, and there is no shortcut to hashing it all out, starting from wherever we are, using all available arguments and considerations in pursuit of the broadest possible reflective equilibrium. This applies not only to questions of value, but also to apparently dispassionate arguments about facts. Every valid argument is a potential *reductio*, and every consideration is potentially relevant to deciding whether to accept a conclusion or reject some premise. That decision depends on the relative strength of the *feelings of rightness* or *feelings of knowing* elicited by each alternative.23 (Few have accepted Zeno’s conclusion that space is unreal just because the argument seemed valid and its premises true.) Everything goes into the lists; but what tip the balance or get brought into equilibrium are inevitably emotions: nothing else has the power to affect motivation.

This is not to say that confrontations among emotions are resolved by a simple dynamic of intensity. Emotions are moulded into widely divergent convictions, action tendencies and normative judgements by individual reflection as well as by the less transparent forces that shape cultural change. Among those forces is the power of words: conversation, debate, rhetoric and argument, all of which are bathed in passion, but all of which still allow for rational debate, providing we are willing to engage in it.

The quest for reflective equilibrium, in this inclusive sense, takes place within individuals as well as among them. Anything can be disputed. We cannot assume, in particular, that ethical values trump all others. On the contrary: any proposed way of marking out the domain of morality will itself appeal to a non-moral value – aesthetic, political, or even religious. Asserting the supremacy of aesthetics was the heroic route taken by Oscar Wilde, Baudelaire, Huysmans, and perhaps de Sade. Taking the religious to trump morality characterises fundamentalist Christians and Muslims. I will conclude this essay by sketching my own conviction that there are both political and aesthetic grounds for restricting the scope of morality in just the way conservatives like Jonathan Haidt deplore.

A political argument in favour of confining the scope of morality to the first two domains might go like this.24 Much of the State’s infringement on individual liberties occurs because politicians assume they are expected to champion morality. One way of promoting political liberty, therefore, is to curb the ambitions of morality itself. This will give politicians less scope to tell us we deserve to be punished for prostituting ourselves, using recreational drugs, choosing ‘inappropriate’ love objects, or letting a teddy-bear be named ‘Mohammed’.
We don’t first have to prove that our values are moral ones. On the contrary: we can adduce non-moral reasons for ceasing to count norms of purity as moral. Norms of etiquette can be reassuringly stable if they are aimed at avoiding disgust. But there is no reason to give them the extra heft and mystique of morality. Norms relating to community and authority can similarly be scaled down: the legal obligation to pay taxes and obey the law needs no support from moral sanctions. We should restrict the scope of ethics itself, in order that it may more closely match the scope of the law in the ideal liberal state.

That suggestion will persuade only inasmuch as it arouses the emotion of approval. But not all emotions serve our own interests. If they stem from ‘adaptations’, they were shaped, like parasites, viruses and bacteria, by biological and social processes blind to individual interests. The ecology of nature gets rid of most species sooner or later (except, it seems, for parasites, viruses and bacteria). The social ecology of moral persons, by contrast, may allow us to survive if it is sustained and modified by talking, reasoning, and reconfiguring our perspectives. The resulting edifices of thought and practice can be consciously designed to enhance our capacity to respond emotionally to the enhancement of our emotions; to select what is to count as worthy of being selected; and to promote the possibility of multiplying human possibilities.

Among the enhanced possibilities afforded by language and sociality is another set of values, the aesthetic. These are equally rooted in emotions, and can claim to compete on equal terms with moral ones. Human imagination serves the practical ends of planning, but it also affords us the capacity to attend to qualitative nuances the value of which cannot be reduced to the simple bivalence of good and bad. In the aesthetic attitude, we can focus not on the practical guidance for which our emotional repertoires evolved, but on their intrinsic quality.

The aesthetic stance comes closest to what Aristotle characterised as ‘divine’, namely pure contemplation. The cultivation of art thus shares an important feature with religion. Two factors, however, argue for the superiority of art over religion. First, art affords imaginative illusions, entertained as such; most religions, by contrast, insist on delusive commitments of belief, enforced with meretricious expectations of rewards and punishments. Second, aesthetic contemplation endorses the possibility of conflicting yet equally authentic values. Monotheistic religions must reject this, being committed to the view that all true values are compatible since they are united in God. The very structure of our emotional life, made manifest in classical tragedy, attests to the fact that conflicting values can be equally real.
Notes

2 Or reduces plants, apparently, to something less than plants. As I write, the Swiss government is reported to have promulgated guidelines requiring researchers to respect the ‘dignity of plants’. See Nature News, 23 April 2008.
7 John Stuart Mill, Utilitarianism (1863; Indianapolis: Bobbs Merrill, 1957).
14 Describing revenge in a tribal culture, Jared Diamond describes a man who had caused an enemy’s lifelong paralysis as perfectly happy – at least until a realignment of tribal alliances allows the victim to be reclassified as friendly. See Jared Diamond, ‘Vengeance is Ours: What Can Tribal Societies Tell Us about Our Need to Get Even?’, New Yorker, 1921; accessed online, 29 April 2008 at http://www.newyorker.com/reporting/2008/04/21/080421fa_fact_diamond?currentPage=all.
Really, what else is there? Emotions, value and morality


The term was first introduced by John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1977). It has been interpreted narrowly and broadly. Although I use it here in the broadest possible sense, I will suggest that what ultimately get pitted against one another are always emotional responses.


I am again indebted here to Ogien, *L’Éthique aujourd’hui* (Paris; Gallimard, 2007).