

**EMOTIONS: WHAT I KNOW,
WHAT I'D LIKE TO THINK I KNOW,
AND WHAT I'D LIKE TO THINK.**

[penultimate draft of a paper now published in *Thinking about Feeling: Contemporary Philosophers on Emotions*, ed. R. Solomon, OUP 2004 pp. 61-75]

Ronald de Sousa
University of Toronto
sousa@chass.utoronto.ca

The Editor's invitation to contribute to this volume appeared to license telling more than I know. Accordingly this essay will move quickly from an all too brief survey of what I know to raise some of the increasingly speculative questions that currently preoccupy me.

I. What I know

On second thought, there's nothing I'm that sure of.

II. What I'd like to think I know.

1. *Reconstructing Cognitivism*. There has been much made in recent decades of the idea that emotions are "cognitive". The term is used in a confusing diversity of senses. Sometimes by 'cognition' one means merely to insist that emotions are not "merely subjective" phenomena. But that is hardly helpful, since there are by my count at least a dozen different things one can mean by 'subjective' (de Sousa 2002a). A more contentful thesis is that emotions are genuine representations not just of the inside world of the body but through that of the external world of value. As representations, they have a mind-to-world direction of fit. In this they are like beliefs rather than desires. At the minimalist end, Martha Nussbaum (whose own view, to be sure, is more robustly "cognitivist") has remarked that a theory is "already in [a] sense a cognitive theory [if] the transmission of information within the animal is central to it." (Nussbaum 2001, 114). But if that is true, it is difficult to see what ground is left for non-cognitivism to occupy. By contrast, Solomon's view that emotions are *judgments* (Solomon 1973)—call this *straightforward cognitivism*—remains refreshingly extreme. The burden of my own proposal will be to suggest that in so far as a "cognitive" view of the emotions can be sustained, it is better

construed on the model of perception than on the model of knowledge or judgment.

Experimental psychology as well as common-sense observation have accumulated an impressive range of facts that refute straightforward cognitivism (DeLancey 2002; Stocker and Hegerman 1996). Fear of flying is a classic example of a groundless emotion. It is experienced by some who are well aware that no method of transport is safer than flying. So while fear has something to do with danger, believing that X is dangerous is neither necessary nor sufficient for experiencing fear of X.

More controversial, because less conclusively shown to be irrational, is the fear of death. According to a notorious argument of Epicurus, no one should experience any fear of her own death *as such* (as opposed to its many unpleasant concomitants—the process of dying, the loss to one's loved ones entailed by one's own death, the unknown fate which those who don't really believe death is *death* contemplate in its stead, and so forth). For death is experienced neither while we are alive—since we're not dead yet—nor when we're dead—since we're no longer there to experience anything. The argument is neat; but many share Philip Larkin's impatience with "specious stuff that says *No rational being / Can fear a thing it will not feel*, not seeing / That this *is* what we fear.... ". Epicurus's argument assumes that to fear X in itself must be to fear *the experience of X*. And that is question-begging, though perhaps not wholly gratuitous.

Yet many who are well aware of these counterexamples continue to call themselves cognitivists. And indeed the dispute over Epicurus also establishes that it makes sense to criticize an emotion for *irrationality*, if only by impugning its appropriateness to its object. The pertinence of such a charge may suffice to justify the label of 'cognitivism'. More important, not all cognition is belief: *perception* are also a form of cognition, and so may provide a better model on the basis of which to think of emotions as cognitive. Compelling visual illusions that persist in the face of the knowledge that they are illusory constitute an immediate analogy to groundless emotions. (See Hoffman 1998 for numerous examples) I am sure enough that something of this sort is right; how the details of the analogy are to be worked out, on the other hand, belong under my third, most tentative heading.

2. *What is universal overflows the species.* Recently my five-year-old's teacher described the dynamics in the life of the day care's denizens. The two dominant children never fought or interfered with one another: each had his or her own sphere of influence. Among the others, most didn't have the strength actively to resist their dominant patron's demands, but some knew better than others how to evade it. Some were demanding yet quick to sympathy; some, on the contrary, seemed rather to relish others's distress. Some jealously demanded exclusive attention; others continually sought protection from the more powerful.... In short, all the commonplaces of human social interaction were already to be seen there. The picture she was sketching was a perfect match for the world scene of politics as well as any lesser realm of human life (cf. Moldoveanu and Nohria 2002). Whether it is displayed in the daily newspaper or dissected by a Hobbes or Macchiavelli, the dynamics of social life seem to turn on emotional dispositions equally powerful at all stages of life. These are merely more transparent in small children unable to articulate the confabulations with which adults dress them up. And as Frans de Waal (1996) has richly illustrated, while social life among other primates may not be quite as varied in its details as that of humans, the range of patterns present in other primate species are very much the same: affiliation and exclusion, rank, dominance and submission, sympathy and antagonism, reciprocity and revenge, compassion, jealousy, envy, anger, guilt and shame: none will be news to the primatologist any more than to my child's kindergarten teacher.

This simple observation raises two questions. The first might well arise from the resistance likely to meet any reminder of our kinship with our mammalian cousins. (For it is part of our primitive animal vanity to pretend we are not animals.) *If we are so much alike, how come we are obviously so different?* And does this leave room for any differences that can reasonably be ascribed *uniquely to our species?*

My answer, in a nutshell, is that human emotions, thanks to a gift of language which brings with it a particularly strong form of intentionality, are differentiated from the emotions of other animals as well as from one another by the acquisition of a *narrative form*. This adds to our

experience of reality a whole new dimension, or rather a whole new set of dimensions, which both transform and attempt to usurp the function of our primitive emotions.

That answer, however, must in all fairness be relegated to the category of things I don't quite know but would like to think.

3. *Varieties of intentionality.* Emotions face both in and out: they reflect facts about the subject, but refer also to something outside, to which they typically are responses. In this respect they offer both an analogy and a contrast with sensory perception. Perception, as recently pointed out by Alva Noë, are "transparent" in the sense that when you attempt to depict your visual field you just end up drawing a picture of the room you are in (Noë 2000, 124-126). By comparison, emotions are relatively *opaque*: often the effect of passion is precisely the reverse: when the angry man, or the joyful bride, or the jealous husband attempt to describe the world, they succeed only in describing their own state of mind, or perhaps even just the hormonal (or as it used to be said the *humoral*) balance in their body. Descartes may have implied something like this in his odd phrase about "passions of the soul" being affections "excited by the animal spirits" yet "referred to the soul" (Descartes 1989, Part I art. 25,27). Although his phrase is not altogether clear, it seems to reflect something of the complex ways in which emotion relates to many "objects": often to some specific target in the world (the "object of my affections"), but also to the state of my body, to my phenomenological consciousness, and to aspects and features of the target most naturally expressed in terms of adjectives or adverbs. Thus if I am *moved* by your performance, this has something to do with the tears welling in my own eyes; if I am *angry at* you, this may be *because* of some things about my own past, no less than because of something you have done, or *how* you have done it, or just what you are like. All this may be signalled, as Descartes (1649/1989), James (1884), and Damasio (1994) have pinned down with increasing accuracy, by some "perception" of the state of my own body.

4. *Emotions and Consciousness.* "Perception" needs to be put in scare quotes because, as all three writers realized, it need never reach awareness. This point is worth stressing. Freud

(1915, p.78) claimed that strictly speaking the notion of unconscious affect is a contradiction in terms. But much recent work on consciousness has established that any simple dichotomy between what is conscious and what is not (even when supplemented with slots for the "subconscious" or "preconscious") is hopelessly simplistic. The "Stream of consciousness" is something of a "grand illusion" (Noë 2002; cf. Churchland 1986). Insofar as emotions are construed on the model of perception, therefore, we should not expect them to be simply conscious or unconscious. A fortiori, since the body's role in emotions is more complex, Freud's remark seems even less likely to be true. In many ways, some of which I return to in a moment, an emotion can genuinely affect not just behaviour but our whole orientation to the world and the events of our lives, without the subject having any particular insight into either the identity of the emotion or the nature of its influence.

5. *Emotions and the Body.* It is often stressed that emotions are *bodily* phenomena. A contrast is presumably intended with "purely cognitive" phenomena, but in some sense perceptions and beliefs are bodily too, so the import of the claim is not clear. The insistence on the bodily character of emotion is presumably intended to underline two differences between emotion and other "states of mind". One is that in many cases specific emotions are associated with certain relatively gross physiological changes: blushing, accelerated heartbeat, skin conductance changes, in short, the kind of measures that impressed James, and that polygraphs are intended to record. But some of these involve highly specific and organized brain circuits such as those controlling fear, rage, euphoria, or disgust (Calder, Lawrence and Young 2001). Others appear to relate to emotional states that may altogether lie below the threshold of awareness. These last, revealed by such pathological states as Capgras syndrome—about which more below—are particularly interesting, because they testify to the existence of emotional states that are not experienced as passions, indeed that may not be consciously *experienced* at all.

6. *The power of teleology in emotions.* A further supposedly established fact about the emotions is that they appear to *motivate*. I say *appear* to motivate, because this apparently

obvious fact is infected by the obscurity of the notion of motivation. What is now increasingly clear is that despite their reputation for motivating bad behaviour, emotions are essentially implicated in our capacity to live a coherent and reasonably well-regulated life: unless you *care*, your life will be a mess; and whether or not you care is surprisingly neatly attested by your galvanic skin response (de Sousa 1987; Damasio 1994).

As amply illustrated in the works just cited, emotions are hugely variable in the ways in which they determine the course of our behaviour. But if, as I believe, complex emotions are learned in the context of "paradigm scenarios" (de Sousa 1990) from which they derive a dramatic or narrative structure, it is misleading to think of them as motivating in the same sense as desires motivate. Desire and belief are concepts tailor-made to enter into a certain neat theory of action, in which they constitute the two ready parameters of motivation. But emotions are more complicated, and play a more subversive and ambiguous role. Most complex emotions prescribe no specific behaviour. They also affect us at many levels, of which the body's readiness to undertake certain sorts of behaviour rather than others is only the simplest. More subtly, they influence decisions not so much by "motivating" as by orienting attention towards this or that among the plethora of considerations that might be thought relevant at any particular juncture. At all these levels beyond the first (and perhaps even there) the emotions play a determining role by virtue of the narrative structures of paradigm scenarios. These are elaborated mostly in the light of early experience, but also perhaps by exposure to art and literature, and they serve to define and differentiate each person's idiosyncratic repertoire of emotions. The resulting complexity means that the teleological structure of emotions often looks more like *fatalism* than like ordinary *determinism*. Rather as the Delphic oracle's predictions were realized largely because of the efforts made to evade them, the emotions often work like agents able to get around almost any attempt to foil their ends. But while the facts here are impressive—and widely described in the literature under the heading of self-deception, *akrasia*, and other pathologies of thought, inference and action—the mechanisms of this elaborate teleology remain obscure.

7. *The axiological hypothesis.* The role of emotions in ethics has long been debated, and the

history of philosophy has oscillated between two extreme views. One holds that emotions constitute a disruptive factor that stands between us and any possible redemption of our bestial nature. At the other extreme is the idea that emotions lie at the core of rationality and ethics. This latter view rests on three considerations.

First, by defining what we *care* about, the emotions set the ultimate ends of all deliberation.

Second, there is a whole range of emotions that are themselves intrinsically epistemic (Hookway 1998). Most notable among these is the feelings of certainty or doubt, as well as a more general feeling of "rightness" that validates our inferences and our conclusions (Mangan 2001). Such feelings, like the feeling of spiritual conviction frequently associated with temporal lobe seizures (Ramachandran and Blakeslee 1998, 177 ff.) are manifestly emotional experiences that come, as it were, bearing the marks of epistemic import.

Third, by controlling the salience of different considerations, the emotions also determine not just ends but also much about the means by which we choose to attain our ends.

On this last point the debate often crystallizes around the issue of *akrasia* or weakness of will. On one view, which I have defended in the past on the basis of an elaboration of the framework laid out in (Davidson 1980), *akrasia* consists in a failure to apply to a current decision the most comprehensive set of available reasons. Emotion is uniquely qualified to assume responsibility for these failures, because while it disrupts the path meant to lead from the "best" reasons to action, it is not the sort of brute cause that would remove the event altogether from the sphere of responsible action. Instead it brings about the failure of the best reasons to be channeled into action by virtue of its control of the agent's allocation of resources.

The work of (McIntyre 1990) and (Bovens 1999), however, has convinced me that my earlier view was tainted by a common prejudice. Many cases of *akrasia* can be viewed more benignly. The accusation of weakness of will is most often leveled on the basis of a clash between the operative reason and the vectorial sum of all available reasons as explicitly declared. But why assume that our best selves are to be found on the side of the most explicit of our

reasons? An accumulation of philosophical, psychological and even neurological research can be mustered to support the view that our explicit declarations are greatly contaminated by confabulation, and that much of what passes for rationality is actually rationalization. Thus, McIntyre and Bovens point out, if Huck Finn passes for akratic because he acts against his explicit principles, so much the worse for his explicit principles. Look instead to the deep emotional commitments that actually govern his inability to be ruled by his principles, and you will see that he was being faithful to his best self.

8. *Emotional cognition as exaptation.* The contrast between rationality and rationalization presupposes a real difference between cognition and illusion, between objective representation and projection—in short, between truth and illusion. I will therefore shortly need to say something about *emotional truth*. But the role of emotion in cognition may also be an auxiliary one, which we can speculate is the result of an "exaptation".¹

The role of emotions in learning, memory and recognition provides a good example. It is well known that memory best retains (or perhaps retains only) what is emotionally significant. (There may have been something pedagogically sound, if in other ways deplorable, in old-fashioned methods of instruction based on blows, humiliation and ridicule.) But let me focus on the more specific example of the "Capgras syndrome" (Ramachandran and Blakeslee 1998, 158-173). Subject affected by the Capgras syndrome insist when they see parents or loved ones that they are seeing an impostor. They have no problems of memory, nor do they suffer from prosopagnosia: on the contrary Ramachandran's patient Arthur conceded that the person before him looked exactly like his mother, and wondered about his mother's motivation in hiring an impersonator. What then could be going on? Ramachandran's suggestion is that a direct link normally exists between the facial recognition mechanism and the areas controlling the appropriate emotional responses (particularly the amygdala). The sight of a parent normally

¹ The term was coined in Gould and Vrba (1982) and refers to "a structure coopted for utility from different sources of origin ... and not directly built as adaptations for their current functions" (Gould 2002, 41).

triggers a—not necessarily conscious— affective response, which is itself subject to an evaluation for familiarity. Think of this on the analogy of an ID card that carries both a signature and a picture. Both have to match: if one of them fails to match, the other is automatically suspect. Thus the recognition of those who are close to us rests on two marks: the cognitive or visual representation, and the emotional "signature". If the signature fails, the visual presentation is deemed fraudulent. In Arthur's case, the affective response is missing. This sets up a discrepancy between the strictly cognitive familiarity check that applies to the face and the missing familiarity check applied to the expected affective response. The Capgras delusion is then no more than a perfectly reasonable inference: the person before me is not my mother, since I get a characteristic emotional marker when my mother appears and I'm not getting it now. On the other hand, she looks exactly like my mother. Therefore she is an impersonator.

The example is of more than merely anecdotal interest. It suggests that the emotions present us with an entire information processing system on its own, a parallel representational system for understanding the world. The questions raised by this possibility, including the question of how emotions might provide us with something we might call *objective representations* of values, belong firmly in the third of my subheadings.

III What I'd like to think

9. *Parallel Systems?* In the light of evidence that the brain's older systems of control continue to work in parallel—and not infrequently in conflict—with those that have evolved in the cortex, Paul MacLean elaborated a hypothesis consistent with the idea that emotions form not so much a set of functions that *contribute* to cognition and behaviour, but actually *a parallel system of control*. This fits in all too well with the evidence cited above from politics, chimpanzees and day care dynamics. MacLean puts it thus: "We might imagine that when a psychiatrist bids the patient to lie on the couch, he is asking him to stretch out alongside a horse and a crocodile". (MacLean 1960, 300).

Indeed there is reason to believe that the circuits of the limbic system—chiefly associated

with emotion—constitute the *original* system of behavioral control. If so, then we are again faced with the possibility that most of our rational thought is rationalization, and that it works well enough in practice because the robots that we have been well conditioned to survive regardless of the implausibility of our confabulations. But that, in turn, raises again the question of the nature of the compact between explicit reason and the ends and means of the emotional system. It has become fashionable to claim that there is not *really* any opposition between reason and emotion, but that may be nothing but a comforting myth. On the contrary: there is a deadly opposition between emotion and reason: it's just that "reason" can't set its own goals or do anything much about them without the connivance of its adversary.

If anything like this is right, it raises once again the question of the nature of our emotional representations and their claim to be giving us information about the world outside our own minds and bodies. Descartes warned that the representations of the senses might be fulfilling their natural purpose without giving us any information about the world as it is: "the proper purpose of the sensory perceptions given me by nature is simply to inform the mind of what is beneficial or harmful.... But I misuse them by treating them as reliable touchstones ... about the essential nature of the bodies located outside us.....". In particular, Descartes continues, "there is no convincing argument for supposing that there is something in the fire which resembles the heat, any more than for supposing that there is something which resembles the pain." (Descartes 1641/1986, 83)

Actually this last claim can be pretty confidently rejected, thanks to the existence of cross-modal confirmation of the reality of heat. What can be seen, heard, touched, measured with various instruments, and observed by means of different senses and by different observers to undergo the same constancies and changes must indeed be objective if anything is (Nozick 2001, 75 ff.). The point is especially pertinent to emotions: are they merely motivational guides to efficient living? or are they actually "cognitive" in the sense that they actually represent something that can be said to have objective reality? On this criterion, there will indeed be objectivity to its deliverances.

The speculative *axiological hypothesis* that currently engages me is designed to explore the prospects for emotional objectivity along these lines. "Axiological" is a term that has fallen into disuse. (There is no entry s.v. *axiology* in the recent *Encyclopedia of Ethics*. (Becker 2001).) It derives from the Greek work for *worth* or better *worthiness*. I revive it to express my conviction that, just as there is a gap between what we *prefer* and what we *decide to do* and another between what we *decide to do* and what we *actually do* (leaving room for two of many varieties of weakness of will), so the gap between what we *know or believe* and what we *prefer* allows for a level of *ascription of value* which is not yet preference yet more than mere assessment of factual truth. Poised as it is in that gap, the axiological shares some features with both epistemic and preference assignments. Yet axiology is distinct from both the epistemic and the level of preferences. We might say it is the locus of our capacity to be *moved*, and yet not *moved to any specific sort of behaviour*. Like preferences, axiological assessments involve a necessary component of subjective response, without, however, being committed to any unique dimension of valence such as attraction/repulsion. But like facts, the axiological aspires to a high degree of *objectivity*, which for present purposes might be roughly defined—in accordance with the remarks above—as the recognition of a need for and a possibility of *corroboration*.

This last trait suffices, however, to think of the axiological hypothesis as positing a special kind of "cognitivism" about emotions. The hypothesis involves three claims:

First, at least some emotions constitute *perceptions of value*, the content of which may at least sometimes meet something like the multi-modal requirement. If this can be made out, we should be able to make sense of a notion of *emotional truth* (de Sousa 2002b).

Second, the values in question are not limited to those valences that determine, along a continuum going from attraction to repulsion, the motivating vectors that drives our behaviour. Instead they are multi-dimensional or, one might say, in contrast to the black-and-white unidimensionality of behavioral motivation, *multi-colored*.

Third, most emotions are not intrinsically moral: it follows from the multi-dimensionality of

value that not all value can be moral value. Yet emotions are, in the end, inevitably the ultimate arbiters of all value including ethical value. At the meta-ethical level, this hegemony of emotion justifies the Wildean adage that ethics is a branch of aesthetics, but it could equally well be formulated by saying that ethics cannot exclude aesthetics from the scope of its vision.

In the rest of this essay, I will say a little more about each of these three points in turn.

10. *The concept of emotional truth.* Standard accounts of truth assume that truth bearers are propositions, however explicated. By adopting a perceptual model of emotion, I discard the idea that emotions are invariably propositional attitudes. (This does not imply that propositions can't be among the many sorts of objects to which emotions are, as I noted, variously related.) This immediately raises two related logical problems: if something can be true, it can be false, and we want to be able to say that its negation is true. We also want to be able to say that a set is *consistent*. How then can the notions of negation and consistency apply to emotions?

Although some named emotions seem to be related as polar opposites (love and hate, hope and despair, admiration and contempt, gratitude and resentment), there is no obvious criterion of contrariness. Two candidates suggest themselves. One is neurological; the other is phenomenological. The phenomenological criterion, however, suffers from extreme subjectivity. While many people would claim it is impossible simultaneously to feel certain pairs of emotions in regard to the very same objects, aspects and situations, disagreements are hard to settle. "Odi et amo", lamented Horace: and indeed love and hate are at once a paradigm cases of opposites, and a typical example of emotions that are frequently mingled (Neu 2000; cf . Greenspan 1978) On the neurological criterion, we might take some inspiration from some elegant work by Marcel Kinsbourne, in which he explored the likely contribution of areas of the brain to different tasks by measuring the degree of interference. He found, for example, that subjects could easily learn to balance a pole on the tip of a finger on the right or left hand. When asked to speak at the same time, however, the pole tended to fall off if balanced on the right index finger, while it had no disruptive effect on the left. Being asked instead to sing produced the opposite effect

(Kinsbourne and Hicks 1978). That sort of interference might yield one measure of consistency: those emotions that inhibit one another, like the activation of the sympathetic and parasympathetic nervous system, would then be rated opposites. Anxiety and sexual excitement may qualify on that score. But the criterion is feeble, as it may relate only to competition for the use of a given brain resource.

Anxiety and sexual excitement may not qualify on the phenomenological criterion. Some people report finding anxiety or fear to enhance sexual excitement. Does the physiology of those emotions work differently for them? or are there simply two incompatible methods of sorting emotions into compatible ones and contraries? However that may be, the phenomenological may, after all, be no more than a reflection of the neurophysiological, as it does in the case of the phenomenological colour cone. The structure of the phenomenological colour cone conforms perfectly to what is predictable on the basis of the opponent process theory of trichromatic colour sensation (Churchland and Churchland 1998, 166-172), which strongly suggests that there is *nothing more* to that phenomenological structure than the reflection of an underlying neurophysiological mechanism. Similarly we commonly assume that the qualities of warmth and coolness attributed respectively to red-yellow and blue-green are so experienced by association. But in fact there is evidence that our perception of colors as warm or cold is a direct reflection of the amount of activation of the low-level opponent channels that give rise to hue, rather than reflecting any learned association between colors and emotions. (Hardin 2000, 120).

For all that, the neurological criterion may often be inaccessible, and in many cases perhaps altogether unavailable. In the more subtle emotions, the "formal object" of the emotion—the attribution to its target of whatever features it is that *makes* it that emotion rather than another—is the only thing capable of defining a relation of contrariety for that emotion. How many such formal objects can there be, defining how many continua of appropriateness and contrariety? The answer hangs, I believe, on the emancipation of elaborate emotions from the practical purposes for which we can assume that many of them were originally selected (Neese 1990). Let me explain.

I argued in (de Sousa 1974) that an important confusion generally mars discussions of consistency of desire, resulting from the false presupposition that a single criterion of consistency is appropriate to both beliefs and desires. That presupposition confuses the *satisfaction* conditions of desire, that is, the condition under which the desired entity or state of affairs exists,² with its condition of *success*, that is, the condition under which the desire is *warranted*. For any two beliefs, compatibility coincides with consistency: satisfaction conditions just are success conditions. But for two desires to be consistent, it is not necessary that their contents be jointly satisfiable, but only that their contents be jointly *desirable*. So while a desire for *p* and a desire for *q* (where *q* implies $\sim p$) are clearly incompatible, it does not follow that they should be regarded as inconsistent. And while this raises difficult questions about how to cash in the claim that two desires are inconsistent, it makes room for inconsistent desires without requiring that consistent desires also be for compatible objects.

Emotions are similar, but messier. The reason is the one just noted: each emotion is linked to its own specific evaluative continuum, and so defines its own proper object, and thereby the dimension along which contrariety might be defined for that emotion. There is no single proper object of all emotions.

Consider a simple example. Allow that the formal object of fear—the norm of appropriateness that is distinctive of fear—is the Dangerous. *Fear that p* is *satisfied* iff *p* is true, but it is *successful* iff *p* is *actually dangerous*. Confining ourselves for simplicity to emotions admitting of a propositional object:

$E(p)$ is *satisfied* iff *p* is true

$E(p)$ is *successful* iff *p* actually fits E's formal object.

In all cases, the emotion's success is independent of its satisfaction. Fear of monsters is not semantically satisfied, but may be successful. The converse may be the case in fear of spiders.

² The word 'satisfaction' is used here in the sense in which it is used in traditional formal semantics. See e.g. Gupta (1998).

Emotional truth, then, refers not to semantic satisfaction, but to success. Fear's assessment of *p* as dangerous consists in some sort of *evaluation* of *p*. Success is tied to the correctness of that evaluation. The notion of opposition, if any, appropriate to a given emotion is internal to that emotion, and follows from the nature of its criterion of success. (Thus trust might be the polar opposite of fear, insofar as *dangerous* and *trustworthy* define a continuum). And while this provides no general criterion of emotional contrariety, it at least suggests a way in which such a concept might have application, as well as explaining why it is difficult to cash out in practice.

11. *The full-colour content of emotional experience.* Our experience of ourselves and the world gives rise to feelings, perceptions, desires and beliefs. These get funneled into a sequence of single decisions, where each requires acting or not acting, yes or no. Desires and preferences, then, form a single black-and-white dimension. Experienced emotions, on the other hand, are so diverse as to constitute no single *kind* of thing at all. Each carries a wealth of specific meanings enriched by a vast class of contrasts; each involves an evaluative response which may, but need not, exhibit a positive or negative valence or inclination to prefer some behaviour. On this, the full-colour view, there are no practical limits to the number of distinct emotions that can be experienced, any more than there are limits to the number of thoughts one can have. (Cf. Campbell 1998).

Consider the emotions aroused by aesthetic experience: watching dance or listening to music, for example. It seems obvious that emotions are involved in some way. But if the point were to express “the great emotions”—those we can list on demand: anger, fear, love, awe, jealousy, sadness, desire—then *why go to all the trouble* of creating and appreciating the subtleties of great art? It hardly seems likely that works of art in all their diversity should be sustaining our interest for their representation of *emotion*, if there are so few of them. The alternative view is that each different moment in art evokes a *sui generis* emotion. The experience of life affords an indefinite range of possible emotional qualia.

But why call such a view *cognitivist*? If cognition is defined to exclude non-propositional

content, then this view is non-cognitivist. If, on the other hand, the term is extended to include any acquired insight into the nature of an external reality, then perception, even where it admits of non-propositional content, becomes a paradigm case of cognition. Emotions, on the axiological view, can then claim an analogous place in cognition, broadly understood.

12. *The multivalence of narrative and the hegemony of emotion.* The idea that (some) emotions constitute apprehensions of a *sui generis* realm of values implies that they transcend any of the scripts for behavioral expression which some of them might originally have been selected to enact. Like other perceptions, they are reducible neither to beliefs nor desires, and I have urged that the world of values they reveal does not reduce to any single measure of positive or negative valence. That feature, I surmise, stems from the linguistic difference of humans: it reflects the possibilities for enlargement (and also the capacity for self-deception) entailed by the ability to elaborate our emotions in terms of a narrative framework.

One consequence has just been noted: although emotions are said to motivate, it might be better to say they channel and construct rationales for behaviour, in accordance with the characteristic narratives that define them. Another consequence is that while it is useful to recognize a limited number of named emotions for the purposes of social communication and classification, the actual experience of emotion is better modeled by aesthetic experience, and the range of values they apprehend better assimilated to the indefinitely many dimensions in which aesthetic experience can take us. This means that axiology should be as sharply distinguished from deontology and preference—classes of states concerned with choices and behaviour—as from value-free facts.

But now if every emotion sets its own standards of appropriateness or correctness, what can emotions contribute to ethics? It is tempting to appeal to Human Nature to set a standard of emotional correctness. Given any quale, an evaluative response that falls foul of the norm will lack appropriateness, and on that basis we can call it perverted, abnormal, or *false*.

The problem with human nature, however, is that there probably is no such thing (de Sousa

2000). Rather than taking human nature as a standard against which to judge emotion, I propose to do precisely the reverse: take emotions, rightly understood, as the ultimate arbiters of ethics.

Consider the classic thought experiment of Mencius: you see a child about to fall into a well, and your apprehension of the situation immediately moves you, and you want to save the child. In this instance, what is apprehended is the *need to intervene*. Or better it is the nature of the total situation, in which the need to intervene roughly sums up the supervenient valence. It is not impossible to witness the scene without being moved thus. Anyone who does so, however, may plausibly be said to lack an objectively appropriate emotion.

This way of describing the situation avoids simple projectionism: what I perceive is not merely the shadow of my own response, but something about the character of a situation as a whole in the context not only of my own singular responses but of the feelings and interests of others. I call this view *axiological holism*.

We do not apprehend value in discrete units but only in the light of a complex of factors that transcend individual experience. No single range of facts suffices for the overall fittingness or "success" of an emotional response. *Biological* facts will speak to its origins and may thereby assign it a proper function in the sense of (Millikan 1989), but they will not determine its relation to currently relevant norms. *Social norms*, in turn, are every bit as likely to be irredeemably nasty as biological ones. (To endorse social norms as the touchstone of normativity would be to condemn all social reformers.) *Individual biography* sets up paradigm scenarios in terms of which each individual understands the world, but this defines only a narrow sense of fit between a current response and a present situation. That fit cannot be identified with value in any comprehensive sense, still less determine what is morally right (D'Arms and Jacobson 2000). Instead it is the totality of all of these factors—biological, social, or personal, and more—that may properly be confronted with one another in the hope of arriving at something like reflective equilibrium. That holistic equilibrium of emotional responses is as close as we can come to reconstructing a notion of normative human nature: it will have to do, I suggest, as a substitute

for that bankrupt notion. Vision provides distal information about our surroundings, despite the possibility of visual illusions. Just so, despite the treachery of our emotional dispositions, emotions in general constitute apprehensions of axiological reality. We tell which is right and which is wrong much as we test the veracity of perceptual information: by appealing to corroborating evidence. This is, I believe, as close as we can get to the multimodal invariance that signals objectivity. It is not all that close, because we are looking for a coherent summing up without much hope of finding neat consilience along the way. But perhaps it is close enough.

References

- Becker, Lawrence C. and Becker, Charlotte B., eds., (2001). *The Encyclopedia of Ethics, Second Edition*. London and New York: Routledge.
- Bovens, L. (1999). "The two faces of Akratics Anonymous." *Analysis*, 59(4), 230-236.
- Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). "Neuropsychology of fear and loathing." *Nature Reviews Neuroscience*, 2, 352-363.
- Campbell, S. (1998). *Interpreting the personal: Expression and the formation of feeling*. Ithaca: Cornell University.
- Churchland, P.S. (1986) "Consciousness: the Transmutation of a Concept". *Pacific Philosophical Quarterly* 64:80-93.
- Churchland, P. M., & Churchland, P. S. (1998). "Recent work on consciousness: Philosophical, theoretical, and empirical." In *On the contrary: Critical essays 1987-1987* (pp. 159-176). Cambridge, MA: MIT Press.
- D'Arms, J., & Jacobson, D. (2000). "The moralistic fallacy: On the 'appropriateness' of emotion." *Philosophy and Phenomenological Research*, 61, 65-90.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: G. P. Putnam's Sons.
- Davidson, D. (1980). "How is weakness of the will possible?" In *Essays on actions and events*

- (pp. 21-43). Oxford: Oxford University Press, Clarendon. (Original work published 1970)
- DeLancey, C. (2002). *Passionate engines: What emotions reveal about mind and artificial intelligence*. New York: Oxford University Press.
- Descartes, R. (1641/1986). *Meditations on First Philosophy with selections from the objections and replies* (J. Cottingham, R. Stoothoff, & D. Murdoch, Trans.). Cambridge: Cambridge University Press.
- Descartes, R. (1649/1989). *The passions of the soul*. Indianapolis: Hackett.
- de Sousa, R. (1974). "The good and the true." *Mind*, 83, 534-551.
- de Sousa, R. (1987). *The rationality of emotion*. MIT Press, a Bradford Book.
- de Sousa, R. (1990). "Emotions, education, and time." *Metaphilosophy*, 21, 434-446.
- de Sousa, R. (2000). Learning to be natural. In Neil Routley (ed.), *Being humans*: (pp. 287-307, 313-316.) Berlin and New York: Walter de Gruyter.
- de Sousa, R. (2002a). Twelve varieties of subjectivity. In Jesus M. Larrazabal (ed.), *Proceedings of the Sixth International Colloquium on Cognitive Science*. Dordrecht, Kluwer.
- de Sousa, R. (2002b). Emotional truth. *Aristotelian Society Proceedings, Suppl.* 76(1), 247-263.
- de Waal, F. (1996). *Good natured: The origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.
- Gould, S. J., & Vrba, E. (1982). Exaptation: A missing term in the science of form. *Paleobiology*, 8, 4-15
- Gould, S. J. (2002). *The structure of evolutionary theory*. Cambridge, Massachusetts: Harvard University Press, Belknap.
- Greenspan, Patricia (1980) A case of mixed feelings: ambivalence and the logic of emotion. In *Explaining Emotions*. ed. A. Rorty. Berkeley and Los Angeles: University of California Press.
- Gupta, Anil. (1998). Tarski's definition of truth. In *Routledge Encyclopedia of Philosophy*, ed. Edward Craig, vol 9 pp. 265-269.

- Hardin, C. (2000). Red and yellow, green and blue, warm and cool: Explaining colour appearance. *Journal of Consciousness Studies*, 6(6-7), 113-122.
- Hoffman, D. D. (1998). *Visual intelligence: How we create what we see*. New York: Norton.
- Hookway, C. (1998). Doubt: Affective states and the regulation of inquiry. *Canadian Journal of Philosophy, Supplementary Volume*, 24, 203-26. C. Misak (Ed.), *Pragmatism*.
- James, W. (1884). What is an emotion? *Mind*, 19, 188-204.
- Kinsbourne, M. & Hicks, R.E. (1978). Functional cerebral space: A model for overflow, transfer and interference effects in human performance: A tutorial review. In J. Requin (Ed.), *Attention and Performance VII* (pp. 345-362). Hillsdale, NJ: Lawrence Erlbaum.
- MacLean, P. D. (1960). New findings relevant to the evolution of psychosexual functions of the brain. *Journal of Nervous and Mental Diseases*, 135, 289-301.
- Mangan, B. (2001). Sensation's ghost: The non-sensory "fringe" of consciousness. *Psyche*, 7. ([Http://psyche.cs.monash.edu.au/v7/psyche-7-18-mangan.html](http://psyche.cs.monash.edu.au/v7/psyche-7-18-mangan.html))
- McIntyre, A. (1990). Is akratic action always irrational? In A. Rorty & O. Flanagan (eds), *Identity, character, and morality* (pp. 379-400). Cambridge, MA: MIT Press.
- Millikan, R. G. (1989). In defense of proper functions. *Philosophy of science*, 56, 288-302.
- Moldoveanu, M. C., & Nohria, N. (2002). *Master passions: Emotion, narrative, and the development of culture*. Cambridge, MA: MIT Press.
- Neese, R. M. (1990). Evolutionary explanations of emotions. *Human Nature*, 1(3), 261-289.
- Neu, J. (2000). Odi et amo: On hating the ones we love. In *A tear is an intellectual thing*. Oxford; New York: Oxford University Press.
- Noë, A. (2000). Experience and experiment in art. *Journal of Consciousness Studies*, 7(8-9), 123-136.
- Noë, A., ed. (2002). *Is the visual world a grand illusion?* Special issue of *Journal of Consciousness Studies*, 9(5-6).
- Nozick, R. (2001). *Invariances: The structure of the objective world*. Cambridge, Massachusetts: Harvard University Press, Belknap.

- Nussbaum, M. (2001). *Upheavals of thought: A theory of the emotions* [Gifford Lectures].
Cambridge: Cambridge University Press.
- Ramachandran, V., & Blakeslee, S. (1998). *Phantoms in the brain: Human nature and the
architecture of the mind*. New York: William Morrow.
- Rorty, A. O. (1988). *Mind in action: Essays in the philosophy of Mind*. Boston: Beacon Press.
- Solomon, R. C. (1973). Emotion and choice. *Review of metaphysics*, 17(20-41).
- Stocker, Michael and Hegerman, Margaret. (1996). *Valuing Emotions*. Cambridge: Cambridge
University Press.