

## Truth, Authenticity, and Rationality

Ronald DE SOUSA<sup>†</sup>

### ABSTRACT

Emotions are Janus-faced. They tell us something about the world, and they tell us something about ourselves. This suggests that we might speak of a truth, or perhaps two kinds of truths of emotions, one of which is about self and the other about conditions in the world. On some views, the latter comes by means of the former. Insofar as emotions manifest our inner life, however, we are more inclined to speak of authenticity rather than truth. What is the difference? We need to distinguish the criteria of correspondence or appropriateness suitable for authenticity from those that embody the criterion of truth. Furthermore there is also a question about the transitions – among states of mind, and between states of mind and behaviour – that emotions encourage. This realm of transitions concerns rationality. After sketching the relevant distinctions, I will endeavour to justify the view that emotions should be appraised in terms of all three terms.

### *In, out, and about: three ways of appraising an emotion*

My emotions offer information about what is going on inside me. Thereby, they also alert me to opportunities and dangers about the world outside me. The information emotions provide in both these domains characteristically involves evaluation and appraisal. It is also, in itself, subject to appraisal, not only for its intrinsic character and quality, but also for some sort of correctness. Inasmuch as an emotion reveals my inner nature, we can speak of it as subject to an appraisal of correctness that pertains to *authenticity* – am I really such as my current emotion displays me? Inasmuch as it tells me about the world, we can think of it as subject to an assessment of *truth* – is the world really such as my emotion indicates? But emotions don't occur in isolation. A third dimension of assessment, without which the other two can make no sense, relates to the context of the emotions. Context frames *rationality*, which is judged synchronically when we assess the appropriateness of a given state in the context of all other co-existing states, and diachronically when we look at the transitions that have taken the subject from one state or set of states into another.

My aim in this paper is to show that the three criteria are deeply intertwined. I begin with some remarks about the claims of truth. Next I explicate the notion of truth as adequacy, by distinguishing between the *success* and the *satisfaction* of an intentional state. I argue that a deeper sort of adequacy, and derivatively of consistency, pertains to success, not satisfaction. This requires that we make, as

---

<sup>†</sup> Department of Philosophy, 215 Huron Street, University of Toronto, Toronto, Ont., M5S 1A1, Canada; Email: sousa@chass.utoronto.ca

clearly as possible, a distinction between compatibility and consistency of intentional states. That is reasonably easy for desire, but very complicated for emotions. The reason, I shall argue, is that each emotion has its own specific formal object, and its nature partly depends on each subject's individual nature. Finally, I shall look at two proposals for assessing coherence and authenticity in the temporal dimension.

### *The claim of truth*

The attribution of *truth* to emotions seems paradoxical at first sight. Emotions are sometimes triggered by experience of the outside world, sometimes they appear to arise spontaneously from endogenous promptings, and most often they are determined by a combination of both kinds of factors. How then could they be said to be either true or false? Elsewhere (de Sousa 2002) I have attempted to ground an answer on a distinction between generic truth, which requires only a certain relation of systematic informational correspondence, and the species of truth that pertains to linguistic formulations embodying propositions. I shall not rehearse that argument here. I will only call attention to some reasons for regarding as misleading the generalization of our conception of truth as applied to beliefs and judgments.

### *The model of belief*

It is tempting to regard beliefs as the paradigm truth-bearers. And since consistency is most easily defined in terms of the logical possibility of conjoined truth, philosophers often speak as if beliefs were the primary bearers not just of truth but also of consistency. That encourages the assumption that if propositional attitudes of any other sort are to be assessed for consistency, the crucial criterion employed must mirror the criterion of consistency for belief. That criterion of consistency (CC) can be expressed as follows for any propositional attitude  $A(p)$ :

$$(CC) \{B(p) \ \& \ B(q)\} \text{ consistent iff Possibly } (p \ \& \ q)$$

In themselves, however, logical relations hold between propositions; only derivatively do they hold of the attitudes of which propositions are the objects. Thus it is certainly true that a person has inconsistent beliefs who believes  $p$  and believes  $\sim p$ , because  $p$  and  $\sim p$  are logically inconsistent. But for reasons I shall now explain, the model of belief doesn't generalize to other propositional attitudes.

### *Success vs. satisfaction*

The reason is that the model of belief conflates two different properties. I call these properties *satisfaction* and *success*. Satisfaction is a purely semantic prop-

erty: a truth-valued entity (inscription, belief, desire, hope, regret, or what have you) is satisfied iff the proposition it contains is true. But success is a matter of whether the *point* or *aim* of the propositional attitude has been achieved, whether it can be rightly assessed as succeeding in its own terms. As Frege remarked long ago, ‘the word “true” indicates the aim of logic as does “beautiful” that of aesthetics or “good” that of ethics’ (Frege 1956, 289). It follows that for belief, satisfaction and success coincide: the truth of  $p$  is a necessary and sufficient condition both for the success and for the satisfaction of  $B(p)$ . But for the states that hold a place in aesthetics and ethics analogous to that of truth in logic, that will not be the case.

I’ll assume, as a first approximation, that the relevant propositional states for ethics are desires; it’s harder to find even an approximately satisfying term for what plays that role in aesthetics, but I will be assuming that any such state will be a certain sort of emotion. That does not imply that there is a single common criterion of success for all emotions. On the contrary, as we shall see, it is a crucial diagnostic fact about the role of emotions in our lives that there is no such common criterion.

*Consistency: the analogy of desire*

Desire is trivially *for* satisfaction, but its rightness or ‘success’ does not depend on truth. If it did, there would be no logical difference between belief and desire. A desire that is not satisfied is not thereby proved wrong. Desire aims at the good: lack of fulfillment does not show it to be defective. It fails only if its object is not good or desirable.<sup>1</sup>

This logical point lies behind the familiar moral thesis that not all values are compatible (Berlin 1981; Williams 1986). Obviously to say that not all values are compatible is not to say that they are not, compatibly, simultaneously *values*: such a view would refute itself. Rather it is to say that they cannot all simultaneously be *realized*. Their incompatibility in this sense does not preclude their all being true *values*.

It follows that if emotions are either propositional attitudes having values as their objects or – as I believe but won’t attempt to argue here – *perceptions* of value (Tappolet 2000), then the fact that not all values are compatible will frustrate their simultaneous pursuit, but it will not stop emotional responses from reflecting genuine values. Emotional responses to such values will be compatible insofar as they are mere apprehensions, just as the desire for two incompatible goods are, *qua* desires, compatible. These won’t count as inconsistent, any more than desires

<sup>1</sup> I’ve periodically harped on this, first in de Sousa 1974, and most recently in de Sousa 2002.

for incompatible goals should count as inconsistent. Like any such pair of desires, however, they will embody a potential conflict. The conflict will surface as soon as any *action-tendencies* associated with the emotions are set in motion. This may provide a reason for regarding the attitudes as irrational, and therefore inconsistent in a subsidiary sense. But we shall see that this offers only a thin and partial test of emotional irrationality.

### *Consistency and compatibility*

If two states are *compatible*, it must be possible for both to exist together. If they are *consistent*, it means that each has an intentional object, specifying some state of affairs, and it is those states of affairs specified by the contents of the respective states that must be able to coexist. In other words, the difference between the two – with their uneasy relations – only shows up in intentional states.

Consider loathing and love: are they not inconsistent emotions? (Neu 2000) It seems so for two distinct but easily confounded reasons, related to each of emotions' Janus-faces. One is that it's hard to see how what both emotions tell me about their target (the actual person at which the emotions are directed)<sup>2</sup> could both be true – that she is lovable, that she is loathsome (unless one is allowed to specify *respects*.) That is a question of consistency. The other is that it's hard to see how one could *feel* both emotions simultaneously. This relates to the emotions as feelings, considered in abstraction from their intentionality. Feelings have *compatibility* conditions: one might feel both hungry and cold, or hungry and warm. But one can't feel both hot and cold in the same respect at the same time.

The contrast between consistency and compatibility, as well as the conditions under which it breaks down, can be illustrated in terms of belief. To attribute contrary beliefs to an agent is to ascribe inconsistency. Thus if *p* and *q* are inconsistent (or incompatible) this will be because they can't be both true together. To ascribe to Arthur belief in both *p* and *q* is to tax him with inconsistency, and hence, in many cases, with irrationality. But note that by reason of the very fact that irrationality of this sort is possible, the two *inconsistent states of belief* involved must necessarily be *compatible*.

But the demarcation gets murkier if *p* and *q* are explicit contradictories – of the form *p* and *not-p*. It may be psychologically impossible, or perhaps even logically incoherent, for someone to be inconsistent because she actually assents to an explicit contradiction. Hence the ascription of such a deplorable state

<sup>2</sup> The term 'target' refers to one of several things that can be meant by the 'object of an emotion'. In de Sousa 1987, ch. 5 I have attempted a taxonomy in which I distinguish 'target', from 'proper target', 'focus', 'motivational aspect', 'propositional object', two sorts of 'causal conditions', and 'formal object'.

boomerangs to convict the ascriber (de Sousa 1971). The refusal to ascribe belief in an outright contradiction is sometimes referred to as a ‘principle of charity’ (Davidson 1982, 168-9). But there isn’t really anything particularly charitable about it, unless it’s charity to give up what you can’t hold onto. For my part, at least, I am unable to conceive of any evidence that could establish that someone actually believes an explicit contradiction.<sup>3</sup> For me, then, at least, the ascription of such a belief is doomed always to be less credible than the alternative hypothesis that I have somehow misinterpreted the subject of my attribution.

The case of mere inconsistency and the case of outright contradiction are relatively clear. But somewhere in between there must be hard cases. The case of *believing p and believing not-p*, simultaneously but not in a single act of assent, is a case in point. One temptation is to assimilate it to the ascription of an inconsistency; but another is to view it as amounting to the impossible ascription of a belief in an outright contradiction. Like the former ascription, *believing p and believing not-p* appears to presume nothing more unimaginable than a failure to confront one’s own thoughts with one another. As in the latter, however, it seems impossible to imagine how one could show that this correctly describes the subject’s predicament.

The problem is magnified when we try to apply this schema to emotions, since we can’t rely here on any preconception of what counts as the difference between an inconsistency that is reasonably ascribed, and one so extreme that it boomerangs to indict the ascriber. But we can still take away the following lesson from the case of belief: that a more interesting notion of inconsistency lies in the impossibility of simultaneous success, not in the impossibility of simultaneous satisfaction of the intentional states in question.

### *Success and formal objects*

We can generalize the notion of success for an intentional state by stipulating that it consists in the *attainment of the state’s formal object (FO)*, where the FO is that which gives the trivial answer to the question: *Why do you hold this attitude?*

Mention of the FO short-circuits a demand for justification: Why do you believe this? – Because it’s true! Why do you want this? – Because it’s good! Why do you love looking at this picture? – Because it’s beautiful!<sup>4</sup> The claim that ‘belief aims at truth’ can then be rephrased as ‘Truth is the formal object of belief.’ But

<sup>3</sup> But that may reflect no more than the limitations of my conceiving capability. See Priest 1998.

<sup>4</sup> This is, of course, where it starts to get complicated. Gertrude Stein is said to have remarked that she never bought a picture unless she hated it at first sight. A necessary condition, one assumes, not a sufficient one, but one that hints at the complexity of aesthetic appreciation.

the same question for emotion will reveal that *there is no answer that ranges over all emotions*. Why are you angry with him? – Because he did something unjust! Why do you regret doing that? – Because I shouldn't have done it! And so on.

*How do emotions get their formal objects?*

Each emotion has its own specific formal object. The reason is that each emotion plays a different role in our lives. Some emotions have functions that are plausibly enough related to common and vital situation types. Where certain scenarios are crucial to our survival, natural selection is likely to have bequeathed us pre-programmed psychological dispositions to deal with them in relatively standardized ways. These are biological functions,<sup>5</sup> of the sort for which evolutionary psychologists have claimed relatively stereotyped behavioural dispositions that sometimes transcend species boundaries,<sup>6</sup> and for which neuroscientists have proposed dedicated circuitry and specialized patterns of hormonal or neurotransmitter activity.<sup>7</sup>

But biological functions are not the only sources of goals and interests. Indeed, biological functions do not primarily serve the individual. They serve the genes, or (since talk of genes is unfashionable or even politically undesirable in certain circles), whatever complex of inherited developmental processes succeed in perpetuating themselves through the great show of generations of expendable individuals in which we play our ephemeral walk-on parts. And derisory as we are in that great and meaningless phylogenetic drama, we have our own aims, interests, and goals, in terms of which our emotional repertoire has proliferated far beyond the original set to which any biological function might be reasonably ascribed. The FO's of our emotions are multifarious, as are our emotions themselves. This fact has three important consequences.

(i) *There is no universal and complete set of emotional truths*. While physiological and evolutionary considerations suggest that many emotional mechanisms will be common to most humans, and even to our mammalian relatives in proportion to their evolutionary propinquity, my specific emotions have, like yours, been

<sup>5</sup> I use the term 'biological function' in the specific aetiological sense elaborated by Ruth Millikan and others (see e.g. Millikan 1989). A convenient and reasonably precise definition is the following, due to Paul Griffiths: '*Where  $i$  is a trait of systems of type  $S$ , a proper function of  $i$  in  $S$ s is  $F$  iff a selective explanation of the current non-zero proportion of  $S$ s with  $i$  must cite  $F$  as a component in the fitness conferred by  $i$* ' (Griffiths 1998, 442, italics in original).

<sup>6</sup> See for example de Waal 1998.

<sup>7</sup> According to Panksepp, 'basic emotional processes emerge from homologous brain mechanisms in all mammals' (Panksepp 1998, 51). Panksepp lists seven such command systems: 'seeking, lust, nurturance, panic, fear, play, rage and play'. But the 'Blue-Ribbon, Grade A emotional systems comprise just seeking, panic, fear and rage' (Panksepp 1998, 53).

shaped by individual experience on the basis of individual dispositions. Those experiences derive from a multiplicity of social and environmental factors that are unlikely ever to be identical for any two of us. Furthermore, despite species-dependent commonalities, the individual dispositions that constrain our responses originate, in part, in unique and irreproducible genomes. Beyond the 'core systems' described by Panksepp, the set of emotions that each of us is capable of experiencing are, like the beliefs we hold, both potentially infinite and, in their totality, unique to each individual.

(ii) *Specific emotional patterns may yield a conception of authenticity.* The variety and uniqueness of emotional repertoires imply both good and bad news for the notion of authenticity. The good news is that it seems possible to give some sort of meaning to the notion of authenticity; the bad news is that it will never be possible to discover, in any particular case, whether or not some particular emotion is a fitting expression of a person's individual nature.

(iii) *The notion of authenticity needs supplementation with a suitable notion of coherence,* deriving from the idea of success rather than from semantic satisfaction. Such a notion of coherence will therefore not necessarily be reducible to propositional consistency.

In the rest of the paper, I shall be elaborating on these points. But I want first to comment briefly on the difficulties raised by the idea that the assessment of an emotion's 'success' must be relativized to the subject's individual nature. The problem is that while individual natures are plausible posits, they are largely unknowable. Here, very sketchily, is why.<sup>8</sup>

### *The actuarial paradox*

To discover the likely truth about an individual, we could attempt to observe that individual in full detail. But if we allow that some of the manifestations of that individual's emotional repertoire may not be authentic, discovering what her dispositions are will not yield the right answer: observation will yield only the flat surface of actual behaviour, but authenticity demands the depth dimension of counterfactual possibility, which alone can discriminate causality from correlation. So we might approach the individual's own truth, like an insurance actuary assessing an individual's chances of an early death, in terms of the statistics that apply to the individual's various characteristics. The problem is that the various classes to which an individual belongs (nationality, class, ethnicity, education, income, place of residence, and so on) will not influence the result monotonically: each one may have a different bearing depending on the context of the others. In the absence of such steadily incremental information, we must treat the resulting

<sup>8</sup> For more details on what is summarized in the next section, see de Sousa 1998.

intersection of all relevant classes as alone able to provide pertinent information. But then the problem is that the accumulation of that sort of information about an individual will cease to be useful just when it reaches its fullest state. For at some point the intersection of all the classes to which the individual belongs is so small that the *margin of error* in the information about that class swamps the actual information collected about it. This is the actuarial paradox: in order to arrive at the best estimate of fair premiums, an actuary must obtain the greatest possible amount of statistically relevant information on the person to be insured. But an actuary who succeeds too well will have reduced the individual to a class of which her client is the only member; and in that situation, the utility of all that information will shrink to nothing. The perfect actuary, we might say, will reduce risk to zero, but that will defeat the whole purpose of insurance. The perfect actuary will have statistics on females, on Afro-Americans, on smokers, on PhDs, on those with long-lived parents, and on those with mutations *m*, *n*, and *o* in their genomes. But she will have no interesting statistics on Afro-American female PhD smokers born of long-lived parents with mutations *m*, *n*, and *o* in their genomes, *if there is only one such person*. We have no recourse, then, but to return to criteria that can be applied directly to an individual and that individual's genome.

### *Self-made inconsistency*

One more objection of principle must be confronted.

If emotions represent the world through the mediation of my body, as I have suggested,<sup>9</sup> does this amount to anything more than the trivial fact that I can infer something about what affects me from the nature of its effects on me? Someone might object that emotions are no more representational than a bruise, from which we can make the inference that the world affords a blunt object in the vicinity. If that is the sense in which emotions bring knowledge of the world, it is not interesting enough to be worth noting.

There is, however, a crucial difference between a bruise and an emotion. It is reasonable to suppose that the indicative functions of bodily states in emotion have been *selected* so as to provide information. A bruise is just an effect, arising when capillaries burst and the blood collects close enough to the skin to be visible. While it is visible and thus can afford information, it is unlikely that its visibility was selected for. Most likely, then, a bruise has no evolved signalling *function*. The bodily manifestations of emotion, by contrast, have *functions*: they are *supposed* to tell us something.

Still, if emotions carry information about the value of states of the world, they do so only by inference. Do the resulting judgments have any objective validity?

<sup>9</sup> And as illustrated by Damasio 1994; 1999.



For something to be about the world, it must depend on more than my responses. If something like Mackie's (1977) 'error theory' applies to the values to which emotions constitute responses, then those emotions will be in the odd position of never being true.

Perhaps. But maybe we are mistaken in demanding more. For consider this: if I respond in certain ways, it may not be up to me to determine whether my reactions truly reflect the world outside me, but it *is* up to me to decide whether I take or mean them to do so: for that, I need only acknowledge the possibility of being contradicted. Consider theologians trying to determine whether a candidate for sainthood has really performed a miracle. They may be impervious to the fact that their inability to discover a satisfactory scientific explanation is thin proof that none exists. But they do accept that the particulars of the case can be discussed, and that relevant arguments and evidence can be offered for and against the verdict that a miracle occurred. That does not validate their claims, but it serves to establish the point and content of those claims. This suffices for the concept of inconsistency to get a grip. It is a kind of inconsistency that depends not on the semantic satisfaction of a propositional content, but on the success of a given propositional attitude. We might call it 'self-made inconsistency'.

Here is the way this applies to emotions. To claim that an emotion tells me something about the world amounts to an acknowledgement that it might be mistaken. This is compatible with the truth of even a fairly extreme form of error theory of the justification of emotion. From this point of view, I can still ask what *counts* as being contradicted, or what it is for the deliverances of two different emotions to be inconsistent.

If this were the only sense in which emotions admitted of contraries, it would already provide some minimal sense for the notion, since it would account for the phenomenological fact that I am sometimes disposed to *think* that one emotion contradicts another. But it would raise the following questions: when I think that, what is it really that I'm thinking? What in fact is it for two emotions to contradict one another? In particular, does it mean that the emotions cannot in fact occur together, or that there is something irrational about their occurring together? And if it depends on my *thinking* that I am experiencing a conflict, can it ever happen when I am not even aware of my emotions?

#### *Another point of entry: aesthetic consistency*

A few years ago, I heard a talk by the philosopher Alexander Nehamas<sup>10</sup> in which he remarked that we don't mind if our friends have terrible taste, so long as we

<sup>10</sup> Now published as Nehamas 2002.

are satisfied that their taste is *consistent*. This was intriguing and appeared to make sense, but the appearance faded under scrutiny. What would be a combination of aesthetic choices that would *fail* of consistency? It seemed to me that any pair of choices could be made to seem coherent, providing we had the ingenuity to invent a story into which both would fit.

What then is it for a story to be coherent? If we knew, perhaps we could transfer the burden of explicating emotional consistency onto the notion of a consistent *story*. (I've long held that most of our emotions, in any case, unfold like stories.) Here again there was an idea as tantalizing as it was slippery. For it certainly seems to be true that some stories are more coherent than others. But the more obvious ways in which that is true don't help. Here are some things that would make for an incoherent story: I tell you that it happened on a dark night in the dead of a cold winter, and that the protagonist was suffering from the scorching tropical sun. Or I tell you that the heroine has long jet-black hair, and that her hair is blonde and short. Assuming that the references in both sentences are to the same protagonist at the same place and time, this is inconsistent. But the example doesn't help, because what needs to be explained is how some special way in which emotions can be inconsistent might be illuminated by a notion of coherence in stories, and the simple notion of logical inconsistency is just what we don't know how to customize to fit emotions. On the contrary: to make sense of the notion of coherent stories, we exercise our intuitive sense of what emotions are appropriately conjoined with provoking events and with one another.

In matters of taste, one can make the case for consistently liking the most *incongruous* things. (But what, if it comes to that, is incongruity?) Say you liked Palestrina and Acid Rock – well, that's easy: they're both meant to induce altered states of consciousness and sound especially good on dope. Or Palestrina and Rap? Harder, perhaps, but doesn't each express with passionate single-mindedness the spirit of a certain circumscribed but intelligible worldview? And so forth. If this were a game, where a challenger throws out putatively inconsistent tastes to be fitted into coherent stories, the challenger would always lose against an imaginative opponent. So what could this notion amount to? In the attempt to find examples, one is immediately drawn to cases that appear to present what are intuitively *aesthetic* oppositions. This indicates that we intuitively think there *are* such things. But it brings us no closer to figuring out what they might be.

#### *Intentional state inconsistency*

What might constitute inconsistency for intentional states other than beliefs? Following on the proposal that inconsistency must rest on success or on Formal Objects, one could try this:

Two emotions are inconsistent iff their Formal Objects are contraries.

Suppose that *admiration* is trivially justified iff its target is *of great worth*, while *contempt* is trivially justified iff its target is *worthless*. Assuming the provisos of ‘same target, same respect’, etc., and setting aside concerns about the univocity of *worth* – i.e. ignoring the fact that something might be worthy in one sense but not in another – ‘*X is of great worth*’ and ‘*X is worthless*’ are logical contraries. So admiration and contempt are inconsistent emotions.

This line of thought faces two problems. First, the number of emotions for which a nameable Formal Object can plausibly be assigned a judgment-based analysis of this sort is very small. It might work for contempt and admiration, for certain forms of fear and attraction, perhaps for secure contentment and jealousy. But it seems doubtful whether love and hate can be dealt with in this way, or joy, sorrow, depression, elation, sexual desire or disgust. Secondly, while there is a clear rationale avoiding inconsistent beliefs, there is so far no equally clear rationale for insisting that we *ought* to avoid feeling inconsistent emotions.

Both problems spring, I surmise, from the grain of truth in the traditional view that what emotions conflict with are not primarily other emotions, but *reason*. If emotions can be inconsistent, on that view, it is because they are intrinsically *irrational*: it is therefore idle to demand that they be anything else. The view draws comfort from some familiar facts of phenomenology. First, all proposed judgment-equivalents for joy, sorrow, depression, elation, sexual desire or disgust sound highly strained. Second, in so far as emotions do embody judgments, they may, like perceptual illusions, persist despite clear knowledge of their falsehood. Thus if fear embodies the judgment that something is dangerous, it can be notoriously impervious to the knowledge that one is not in fact in the presence of danger. The converse is also true: in the light of what we know about the relative risks, we should not only give up fear of flying but take up fear of riding in cars. Yet we often don’t. All this is unsurprising in the face of a good deal of neurophysiological evidence now available that Plato might after all have been right: that the ‘faculty’ that leads us to experience visceral fear and the faculty of judgment that tells us what is dangerous are indeed separate faculties (LeDoux 2000; Panksepp 2001).

Here is how these facts bear on the problems just raised. Suppose that at least some emotions are in part attributable to brain and somatic processes that elude control by the rational processes of judgment associated with the cortex. Then – if we are independently convinced that there are normative standards of consistency applicable to emotions – we should look for *sui generis* notions of emotional consistency rather than hoping to derive them from logical relations between judgments.

*Compatibility of resource use*

I argued above that compatibility and consistency merge at the limit. Perhaps we should look, then, at certain forms of incompatibility for clues about the nature of inconsistency.

In some lovely experiments, Marcel Kinsbourne showed some years ago that if you balance a pole on your right index finger you can't talk at the same time. If you balance the pole on your left index finger you can talk but not sing. On then-prevalent assumptions about hemispheric asymmetry,<sup>11</sup> this was evidence that pre-emption of brain resources in one hemisphere interferes with a task requiring resources in the same hemisphere, but not with one that mobilizes the other hemisphere (Kinsbourne and Hicks 1978). Could something like this be true of some pairs of emotions?

Insofar as emotions involve action-tendencies or physiological processes, they might compete for brain resources. We can get a glimpse of how it might work by thinking of the basically antagonistic organization of the autonomous nervous system into sympathetic and parasympathetic subsystems. The sympathetic system stimulates heartbeat, raises blood pressure, dilates the pupils, dilates the trachea and bronchi, stimulates the conversion of liver glycogen into glucose, shunts blood away from the skin and viscera to the skeletal muscles, brain, and heart, inhibits peristalsis in the gastrointestinal (GI) tract. The parasympathetic reverses these effects (Kimball 2003). These antagonistic effects obviously have consequences for the compatibility of basic emotions, insofar as some emotions tend to activate sympathetic effects, and others parasympathetic effects. (If I'm feeling *tight-assed* I won't *shit in my pants*.) But the brain processes underlying basic emotions are rather more specialized. According to Jaak Panksepp (2001), the brain contains a variety of genetically ingrained emotional systems for generating specific classes of emotional behaviors. . . . To all appearances, affective experience is a rather direct manifestation of the arousal of these systems. When these systems are electrically stimulated, humans report urges to act and describe emotional experiences that have a feeling of belongingness, as opposed to being alien to the self (Panksepp 2001, 146).

Panksepp shows that some of the same neurotransmitters are implicated in different emotions, as are some of the same brain regions. But information of this sort affords no easy inference as to the compatibility of the emotions in question. Panksepp does, however, hold out hope that 'a new discipline of experimental philosophy' might look into the 'laws of sentiment', citing a number of such

<sup>11</sup> These are: that motor functions of the left hand are controlled by the motor cortex in the right hemisphere, and those of the right hand by the left; that the right hemisphere also controls music (at least in non-musicians), and that the seat of language is generally in the left hemisphere.

'laws' proposed by (Shand 1920). Several of Shand's laws refer to compatibilities and incompatibilities. Here is a sampling:

17) 'The joyful temper lowers the threshold of sensibility for joy, hope, and confidence, but raises it for sorrow, despondency, and despair.' (p. 154).

38) 'Fear and anger tend always to exclude one another, where both are referred to the same objects.' (p. 254).

70) 'Sorrow tends to be increased by the knowledge that another rejoices at our suffering.' (p. 341) (quoted by Panksepp 2001, 151).

Instancing such 'laws', Panksepp suggests that 'the arousal of the various emotional command systems could be distinguished subjectively from each other by humans as being fundamentally distinct feelings . . .' (Panksepp 2001, 147). This encourages the thought that the most direct route to the question of the logical relations among those basic emotions that are least susceptible to cortical control might actually be through their phenomenology. It assumes that emotions necessarily have a phenomenology. This is inconsistent with the biological origins of our emotional dynamics, alluded to above (p. 7), insofar as those are likely to remain inaccessible to awareness. But for now we can evade that problem by focusing on the behavioural concomitants of emotion, as do these Shandian 'laws':

33) 'The universal end of Fear is merely to prevent the occurrence of some threatening event whether the danger be "real" or "imaginary".' (p. 215)

37) 'All varieties of anger tend to accomplish their ends by some kind of aggressive behavior.' (p. 250) (quoted by Panksepp 2001, 151).

### *Consistency of action tendencies*

Fear and anger typically involve dispositions to flight and fight. If both are activated, they can't both attain their respective behavioural aims, as specified in laws 33 and 37. Actually all of Shand's laws, were they not mitigated by the qualifier 'tends to', would probably be false. But they are suggestive in alluding to emotional states likely to be thought to be conflicted because of their associated behavioural dispositions. In addition to their bearing on compatibility, action-tendencies may provide a way to approach the question of emotional consistency. Shand's laws 33 and 37, for example, imply that if I feel anger and fear at the same time, I will be disposed both to approach aggressively and to withdraw from potential harm. And since I can't do both, we might infer that from a practical point of view it would be irrational simultaneously to hold that particular pair of emotions, in some indirect sense that is grounded in emotions' characteristic motivational force. This might yield a weaker but viable sense in which two emotions might be held to be normatively inconsistent – though Shand seems

himself inconsistent in also maintaining, in the particularly dubious ‘law’ 38, that such cases of unreasonableness can’t occur at all.

Here’s how this might work. Insofar as action-tendencies are dispositional, they remain entirely compatible so long as the conditions of their exercise don’t crop up. On the other hand, practical incompatibilities might suffice to ground a normative standard of emotional rationality if they are *felt* as inconsistent. It might seem *irrational* to remain simultaneously committed, even conditionally, to incompatible courses of action. For even if the experience of an emotion is, at least in part, about the world outside me, it is also typically experienced as having a certain felt quality in itself. Emotions that are *incompatible*, therefore, are likely to be *felt* as normatively *inconsistent*, felt, that is, as a problem requiring some sort of resolution. In this sense, just as too much inconsistency shades between beliefs into incompatibility, so conversely certain kinds of incompatibility may turn into inconsistency, by the ‘self-made’ criterion I offered above: if I *acknowledge* an attitude to inconsistent with one I hold, that suffices to validate the ascription.

### *Rationality through time*

So far, I have spoken only of synchronic consistency or coherence. But an accusation of inconsistency commonly imputes *inconstancy*: an irrational tendency to change one’s emotional attitudes from one moment to the next. Considered over time, authenticity, wisdom and rationality presumably rest on some principles governing preferable mixes of constancy and flexibility in our attitudes through time.

But how to find such principles? Grounding appraisals of emotional attitude through time on logical principles alone seems hopeless. Here are two illustrations, pertaining respectively to assessments of the future and the past, of how what turns out to be factually normal is often judged, in the abstract, to be irrational or even illogical despite the absence of clear standards.

George Ainslie (1992; 2001) has explored the many consequences of the fact that we tend to discount the future at a hyperbolic rate. Prominent among these consequences is that as we approach a pair of unequally distant targets, of which the more distant is rated as of greater value, their order of preference will get reversed as one gets closer to the lesser. (The effect resembles the perspectival effect that allows a smaller building to occlude the taller when one gets closer to the former.) We commonly regard this as irrational: if the more distant prospect is indeed more valuable absolutely, how could its value change merely by virtue of proximity? The problem is that it is hard to say why this is not a good idea without begging the question against the alternative view that value reversal is just a reasonable response to a change in perspective brought by the passage of time (Bovens 1999; de Sousa 2003).

Assessments of value can also illustrate the impossibility of making good on imputations of inconsistency about the past in a rationally satisfying way. Daniel Kahneman (2000) has found that when we look back on global past episodes that consist of several stages, our evaluation of the whole is not additively related to the evaluation of the parts. Instead we seem to use a ‘peak-end rule’, which takes into account only the ‘peak’ segment (registering the most extreme valence), and the last segment whatever its valence (the ‘end’). As a result, people rated a longer period of discomfort as the less unpleasant, providing the added period of discomfort was less painful than the last phase of the shorter (Kahneman 2000, 696–7). In this case, then, people rated *more pain* as significantly *less unpleasant*. This contradicts what seems an obvious truth, that if you take a string of unpleasant moments the addition of extra minutes of pain can only *increase* total discomfort. Kahneman regards this as involving ‘violations of logic, because the temporal dimension of experience is not directly included in the representations that are evaluated’ (2000, 707). He takes it to be inconsistent to respond to each component moment as indicated by the ratings assigned to those moments, but then fail to add them up when assessing the whole. But actually there is no violation of *logic* here at all. The inconsistency, if there is one, is of an altogether different sort, calling for some sort of *sui generis* emotional criterion of consistency over time, not reducible to any standard test of consistency.

### *Perceptual framings*

Karen Jones has advocated a ‘practical’ notion of emotional rationality that suggests one way of addressing the temporal dimension. On her view, to say that emotions are perceptions of value amounts to the claim that the objects emotion brings into salience are of practical importance to the agent’s projects. They are those to which the agent *ought* to be attending and responding. Jones suggests thinking of the rationality of emotions as rational perceptual ‘framing’, where the framing of a situation is what enables us to view relevant factors as suitably reason-giving. An agent:

... wants all and only such considerations to be salient to her – ‘all’ so as to be sure that she will not have occasion to regret her choice as having been made in ignorance of some important consideration, and ‘only’ so as to be sure that her deliberation will not be derailed by considerations that she does not think matter (Jones 2004, 340).

The normative force of this suggestion is clear: rational emotions facilitate the process of ‘latching on’ to the considerations that *should* form the basis of a practical decision. This is not equivalent, Jones contends, to the suggestion that emotions dispose us to form correct judgments. For the considerations to which emotions may sensitize us may be appropriate to our *individual* situation even though the *objective* beliefs they facilitate are less than probable. The emotion,

we might say, comes ready-weighted according to the degree of relevance it may have for an individual, rather than keyed to objective epistemic criteria.

On this interpretation, consistency of two different emotional states would depend on two factors. One is the compatibility of the strategies of action they 'frame'. The second factor concerns the compatibility of perceptual salience. Attention determines not just what is seen, but *how* it is seen and what it is seen *as*. To illustrate, take an often-cited case raised by Murdoch 1970: can a mother-in-law change her attitude to her daughter-in-law by construing the latter's *vulgarity* as *vitality*, her *juvenile* manner as *youthfulness*? If so, can she attend to both at once? (One can see the duck and the rabbit, but not at the same time. The two states are incompatible, and if they weren't they would be inconsistent.)

The compatibility of frames in both these senses is obviously much harder to assess than the compatibility of propositions. This leads us to expect three consequences.

First, it confirms our sense that two conflicting emotions can indeed be held to be inconsistent, where the whole orientation either of my attention or behaviour is monopolized by one emotion. In such cases any other emotion that would reconfigure those same elements in the service of a different perspective would be incompatible with the first, and might thereby be rated irrational and so termed in this sense inconsistent. This point in itself gathers together three ideas that have already been broached:

- (i) inconsistency of emotions is in part a matter of competitive resource use, but can't be reduced entirely to that;
- (ii) factual incompatibility cannot be clearly split off from normative inconsistency of content; and
- (iii) the normative standards we are seeking must be thought of as applying not just at an instant but across a stretch of time.

Second, Jones's suggestion also squares with the observation that the inconsistency of emotions, since it is closely related to incompatibility, can be used dynamically in order to move from one state to another: in therapy, one emotion is often pressed into service to dislodge another.

The third observation returns us to a fact already noted, which makes the analysis of emotional coherence particularly elusive: namely that the consistency of any two emotions cannot be assessed in the abstract, on general principles independent of the individual concerned. How are we to relativize such assessments to specific individuals? And how, if the notion must be relativized to a particular, can one make use of the normative or critical notion of emotional consistency?

Normative standards are generally assumed to be universal in their applicability. There is just one exception: the *aesthetic*. In matters of aesthetics, there are no formulable principles of coherent taste that apply across wide ranges of



properties.<sup>12</sup> Yet we stubbornly share a strong intuition that there is something to the idea of aesthetic coherence. But if, as I suggested above, we can make a coherent story out of just about any incongruous collection of fragmental anecdotes, then we still lack an account of what it means to ascribe aesthetic coherence or inconsistency to a set of emotions.<sup>13</sup> So aesthetic emotions appear to constitute a counter-example to Jones's thesis that emotional rationality is practical. On the other hand, they fit in well with her claim that such rationality must be relativized to the specific characteristics of an individual agent. Of course, aesthetic emotions may evade standards of rationality altogether. But in that case Nehamas's intuition about aesthetic consistency could not be made out at all.

### *Fractal patterns of emotion*

In their study of the three therapists Carl Rogers, Albert Ellis and Fritz Perls, Carol Magai and Jeannette Haviland-Jones (2002, henceforth 'M&H') offer a major extension of Jones's perspective on long-term practicality. They propose that each one of their subjects – and by induction the rest of us too – is driven in both life and work by a dominant pattern of emotion.<sup>14</sup> While all three of their therapist-subjects repudiate the traditional psychoanalyst's interest in their patients' past and particularly their childhood, M&H argue convincingly that in life, theory and practice, each exemplifies a unique pattern, largely set up in childhood, characterized by the dominant role of certain emotions as well as the inhibition or rejection of others. Ellis, for example, had devised in early childhood a number of stoic maxims by which he was able to keep anxiety and suffering at bay. His therapeutic practice consisted in large part in exploiting his dominant emotions of anger and contempt to browbeat his patients into using those same emotions to ward off their own negative emotions. Each of the three therapists displays a characteristic 'emotional signature', a dynamic pattern in which, in terms borrowed from Dynamic Systems Theory, some emotions function as 'attractors' and some as 'repellers'. This presupposes that there are indeed dynamic relations between emotions, enabling both subject and therapist to use one emotion in controlling or modifying another. Some seem to be based on incompatibilities akin to those in Shand's 'laws': in a simple 'oppositional' relation, for example,

<sup>12</sup> Arnold Isenberg (1949) has shown how to reconcile this fact with the intuition that criticism is nevertheless rational. His solution derives precisely from the contrast between the perceptual as opposed to propositional knowledge involved in critical appreciation.

<sup>13</sup> We can't, in particular, understand the sort of case Nehamas had in mind in terms of a requirement that the story into which the various apparently incongruous parts can be made to fit must be a common or familiar one. For that would disqualify all cases of originality and creativity, which in itself seems incompatible with our best understanding of the standards that govern the aesthetic.

<sup>14</sup> For a similar thesis, in a darker vein, see also Moldoveanu and Nohria 2002.

Ellis is seen as controlling fear both in himself and in his patients by mocking it (M&H 328). For Fritz Perls '[S]hyhness is a form of shame . . . an interrupter for intense excitement', and in turn 'excitement, when interrupted becomes anxiety.' Some dynamics are more elaborate: 'contempt can serve as an anti-shame strategy. . . . However, once the contempt defense is rendered impotent through the contagion of shame, renewed opportunities for self-awareness and self-evaluation are opened up' (M&H 95). Seesawing dominance of shame and contempt made Perls 'the master craftsman of humiliation':

By subjecting patients to his contempt for their manipulations and weaknesses, and by hounding them with his relentless scorn and derision, he was able to provoke the shame /rage spiral and the attendant release of tremendous emotional energy. In fostering the conversion of shame into anger and by supporting patients' tentative movements towards self-assertion, he taught a liberation politics of emotion for underdogs (M&H 185).

M&H's most original contention is that these emotional signatures have a 'fractal' structure, by which they mean that they are detectable at different scales of observation. From the broadest perspective, they can be detected in the patterns of the three therapists' life decisions. They also find themselves erected into principles embodied in the main tenets of the therapists' theoretical work. And they show up in the specific gestures, attitudes, physical postures and facial expressions revealed in the frame-by-frame analysis of individual therapeutic sessions recorded on film.

The attractors and repellers that emerge in these fractal structures are distinguished by various degree of fixity. Rogers, for example is found to be a life-long avoider of anger, whose main attractors are joy and shame. For Perls and Ellis, anger and contempt are the major attractors. Perls' dynamic is based on oscillation between contempt and shame. Ellis's pattern is much more rigidly fixed on avoidance of strong negative emotions. These differences account for the different potential for long-term transformation over the three subjects' lives.

As we increase the time span in terms of which we look at the emotions, the issue of consistency broadens out, from the narrow question of what emotions can counteract or generate other emotions, into a wider question about the leading structural and dynamic patterns in a life. At this level, talk of certain clusters of emotions as being consistent or inconsistent takes on a very different resonance. It can no longer rest on any kind of simple compatibility, since in the long term even the most incompatible states may supplant one another. Nor is it a question of whether someone can be assessed in terms of the sort of normative standards that ground ordinary charges of inconsistency. Yet the notion of consistency still seems usable at both the explanatory and the critical level, to aid our understanding of how the dynamics of emotions can shape a coherent and productive life. In that spirit, M&H speak of the dependency in all three of their subjects of their

'wisdom' on their 'passion'. And indeed, what emerges from their examination of these successful yet very different individuals, interestingly comparable since they are all in the same profession, is that all make the best of the widely different individual emotional configurations, around which their personal thriving and professional success is built.

M&H's book can be viewed as an argument in favour of a certain conception of authenticity. This authenticity is, in a sense, a reinterpretation of the old notion of *wisdom*. This is not, as in most traditional accounts, a matter of fulfilling some universal ideal based on a proper understanding of human nature. Rather it has to do with the achievement of a certain fit between the basic emotional configuration that defines an individual nature and that individual's choices and habits. That's why it counts as authenticity.

*Conclusion: the unity of the emotional virtues*

In this conception, the truth of emotion, their authenticity, and their rationality through time merge into a single complex requirement for the life of the emotions. Authenticity is truth to myself, and insofar as each of us is different, authenticity will also differentiate us from one another rather than lead us along the same path. In such a framework uniformity from one person to another is not to be expected. Furthermore, stability is not necessarily an advantage or a virtue. On the contrary, too much fixity, like too little, can impede development and prevent natural chaotic processes from generating a new equilibrium at a different place. Given the informational function of emotion, this is not surprising, since too much fixity is clearly a disadvantage in the search for knowledge as in much of life (Oatley 2000).

In illustration, M&H showed how of their three subjects Rogers was the only one to whom a mid-life 'crisis' brought a real sense of change, resulting in more openness in his therapeutic approach and a greater ability to form friendships with men in his personal life. Ellis, in the grip of an emotional pattern erected early in childhood to defend him against all strong negative emotions, changed very little. As for Perls, he constantly hovered between the polar attractors of shame and self-doubt on one side, and contempt and grandiosity on the other, so that the shifts in his mood and behaviour were not matched by long-term changes in that dominant pattern itself. Despite their reservations about the benefits of stability, M&H do claim that the fractal shape of their subjects' emotions endures throughout their lives, and that their subjects' ways of implementing this pattern constituted their way of grounding their own specific form of wisdom in their own characteristic passion.

It is a striking feature of the emotional dynamics to which M&H draw attention that they are mostly unavailable to their subjects' consciousness. Although all three are therapists, whom one might expect to have brought their own emotions

to a fine point of awareness, it is clear that the influence of the major emotional configuration that dominates their life, their theory and their practice is hidden from their own consciousness. How deep a fact is this? Is it the result of some kind of necessity? How would these men, had they had the benefit of M&H's insights, have used these insights? Could they have incorporated them into their own character in some way? Or must we suppose that, if they had been shown these truths about themselves, some other crucial fact about the patterns exhibited in their lives would have escaped them, so as to restore to their lives, as it were, a necessary background of unknowing? Perhaps there exists, at the larger scale of a whole life, something like the impossibility of being fully conscious of one's own processes when executing a skilled task. And yet it's hard to see why, once brought to one's attention, such patterns couldn't be focused on as an object of awareness.

Although M&H make much of this unawareness of their subjects' own emotional patterns, they don't answer or even raise these questions. Let us speculate a little on their behalf. Suppose Ellis were confronted with M&H's account, and came to see the role played in his life and practice by contempt and anger.<sup>15</sup> Might he repudiate it? Might he come to wish he could change? And if so could he change and remain true to himself, *authentic*? My guess is that M&H's prediction would be negative, on the basis of the rigidity of the structures in place throughout Ellis's life. But only a long-term perspective could tell. Suppose Ellis now decries his own behaviour. Either his protestations translate into changed of behaviour, or they don't. Whether they do or not, only the long term can tell. Either way, that too will become part of a pattern that can be discerned only over the long term. But by the time it proves to be stable, the new pattern, if any, will doubtless also have passed out of awareness, on the model of the stages of well-rehearsed routines. It follows that the relation of consistency or inconsistency that matters most at this level cannot be exhausted by the phenomenology of emotion, for the most important patterns will not show up as such in conscious experience at all.

Still, one should be able to take at least an aesthetic interest in the patterns formed by one's behaviour. Such an interest, I've argued, would not entail any attendant action-tendency. Taking an aesthetic interest in who one is at any given level of explanation and motivation would not seem to be excluded a priori from the life of a wise agent whose emotional life is authentic and rational, though it might not change her conduct. In part, then, it is according to criteria of aesthetic coherence that we might assess the authenticity as well as the long-term rationality of a life.

Just how one might do this is far from clear. The only rationales available to understand emotional consistency get no grip on the aesthetic, both because the

<sup>15</sup> Since Ellis is the only one of their subjects who is still alive, the question might actually find an answer.

aesthetic by definition restrains all action-tendencies, and because resource-based physiological criteria of compatibility are not guaranteed to surface into awareness. It is difficult to see how the aesthetic could lack a phenomenology, and hence what sense could be given to an aesthetic inconsistency or conflict that remained unconscious.

Perhaps, however, this failure should be welcomed: the aesthetic is precisely the domain in which we escape the forced choices imposed on us by the need to act. Hence it should also escape the straitjacket of consistency in the way that stories do. A story can always be defended by moving it to a meta-level or level of irony. More generally, we can often get away with a second-order coherence in all aesthetic matters: How do these elements fit together? *Well, I just wanted to show how incongruous these elements are.* The possibility of viewing oneself aesthetically affords a similar strategy, so that authenticity may be bought at the price of recognizing one's own incoherence: *Well, it's incoherent, but it's all me.* Walt Whitman flaunted his inconsistencies: *'I contradict myself: I am large, I contain multitudes'*. On a purely aesthetic level, I have been forced to conclude, Whitman won't be able to make good on his boast. We can now see a third reason for this: namely that on the aesthetic level apparent inconsistency can always take refuge in irony. But that leaves plenty of ways an art work can be criticized, and the incoherence of the notion of aesthetic inconsistency doesn't mean, despite my failure to discover purely aesthetic critical standards, that none are there to be found.

Insofar as emotions involve neural and bodily resources, I have argued, we can make sense of a notion of emotional compatibility. Insofar as emotional rationality is practical, we can also make sense of standards of consistency based on considerations of compatibility but guided in part by the needs of coherent long-term planning. The attempt to apply aesthetic standards to life, however, is complicated by two factors that, as I have tried to show, result in a proliferation of success criteria for emotions. One is the uniqueness of individual natures; the other lies in the variety of ways emotions acquire their formal objects. At the level of generic truth, there is no generally applicable standard of appropriateness that can be applied to different people, even in what appear to be similar objective situations. At the level of authenticity—as truth revealed about the subject – the relevant standards of appropriateness must both take account of the individual's sense of emotional truth, and look to the subject's individual nature. That, in turn, means assessing those standards themselves in terms of criteria of long-term coherence of which the most promising account rests on the moving sands of a dynamics of chaos. Such is the entangled complexity of the three ideals of emotional truth, authenticity, and rationality.\*

\* I am indebted to Aaron Ben-Ze'ev and another, anonymous, referee for useful suggestions.

## REFERENCES

- AINSLIE, G. 1992, *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*. Cambridge: Cambridge University Press.
- AINSLIE, G. 2001, *Breakdown of Will*. Cambridge: Cambridge University Press.
- BERLIN, I. 1981, *Concepts and Categories: Philosophical Essays*, ed. H. Hardy; introd. B. Williams, Harmondsworth: Penguin.
- BOVENS, L. 1999, 'The Two Faces of Akratics Anonymous', *Analysis* **59**, pp. 230–36.
- DAMASIO, A. R. 1994, *Descartes' Error: Emotion, Reason and the Human Brain*, New York: G. P. Putnam.
- DAMASIO, A. R. 1999, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, New York: Harcourt Brace.
- DAVIDSON, D. 1982, 'Rational Animals', *dialectica* **36**, pp. 317–328.
- DE SOUSA, R. 1971, 'How to Give a Piece of Your Mind, or the Logic of Belief and Assent', *Review of Metaphysics* **25**, pp. 51–79.
- DE SOUSA, R. 1974, 'The Good and the True', *Mind* **83**, pp. 534–51.
- DE SOUSA, R. 1987, *The Rationality of Emotion*, Cambridge, MA: MIT Press.
- DE SOUSA, R. 1998, 'Individual Natures', *Philosophia* **26**, pp. 3–21.
- DE SOUSA, R. 2002, 'Emotional Truth', *Aristotelian Society Proceedings Suppl.* **76(1)**, pp. 247–63.
- DE SOUSA, R. 2003, 'Paradoxical Emotions', in: S. Stroud and C. Tappolet, *Weakness of Will and Practical Irrationality*, Oxford, New York: Oxford University Press, pp. 274–97.
- DE WAAL, F. 1998, *Chimpanzee Politics: Power And Sex Among Apes*, revised edn., Baltimore: Johns Hopkins University Press.
- FREGE, G. 1956, 'The Thought: A Logical Inquiry', trans. A. Quinton and M. Quinton, *Mind* **65**, pp. 289–311.
- GRIFFITHS, P. E. 1998, 'Functional Analysis and Proper Functions', in: C. Allen, M. Bekoff, and G. Lauder, eds, *Nature's Purposes: Analyses of Function and Design In Biology*, vol. **44**, Cambridge, MA: MIT Press.
- ISENBERG, A. 1949, 'Critical Communication', *Philosophical Review* **54**, pp. 330–344.
- JONES, K. 2004, 'Emotional Rationality as Practical Rationality', in: C. Calhoun, ed., *Setting the Moral Compass: Essays by Women Philosophers*, Oxford, New York: Oxford University Press, pp. 333–52.
- KAHNEMAN, D. 2000, 'Evaluation by Moments: Past and Future', in: D. Kahneman and A. Tversky, eds, *Choices, Values, and Frames*, Cambridge, New York: Cambridge University Press, pp. 693–708.
- KIMBALL, J. W. 2003, *Biology Pages*, at: <http://users.rcn.com/jkimball,ma,ultranet/BiologyPages>
- KINSBOURNE, M. and HICKS, R. 1978, 'Functional Cerebral Space: A Model For Overflow, Transfer and Interference Effects in Human Performance', in: J. Requin, ed., *Attention and Performance VII*, Hillsdale, NJ: Erlbaum, pp. 345–362.
- LEDoux, J. 2000, 'Emotional Circuits in the Brain', *Annual Review of Neuroscience*, **23**, pp. 155–84.
- MACKIE, J. L. 1977, *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin.
- MAGAI, C. and HAVILAND-JONES, J. 2002, *The Hidden Genius of Emotions: Lifespan Transformations of Personality*, Cambridge: Cambridge University Press.
- MILLIKAN, R. G. 1989, 'In Defense of Proper Functions', *Philosophy of Science* **56**, pp. 288–302.
- MOLDOVEANU, M. C. and NOHRIA, N. 2002, *Master Passions: Emotion, Narrative, and the Development of Culture*, Cambridge, MA: MIT Press.
- MURDOCH, I. 1970, *The Sovereignty of Good*, Cambridge: Cambridge University Press.
- NEHAMAS, A. 2002, 'The Art of Being Unselfish', *Daedalus* **131(4)**, pp. 57–68.
- NEU, J. 2000, 'Odi et Amo: On Hating the Ones We Love', in his *A Tear is an Intellectual Thing: The Meanings of Emotion*, Oxford: Oxford University Press.
- OATLEY, K. 2000, 'The Sentiments and Beliefs of Distributed Cognition', in: N. Frijda, A. Manstead and S. Bem, eds, *Emotion and Beliefs*, Cambridge: Cambridge University Press.
- PANKSEPP, J. 1998, *Affective Neuroscience: Foundations of Human and Animal Emotion*, New York: Oxford University Press.

- PANKSEPP, J. 2001, 'The Neuro–Evolutionary Cusp Between Emotions and Cognitions Implications for Understanding Consciousness and the Emergence of a Unified Mind Science', *Evolution and Cognition* 7(2), pp. 141–63.
- PRIEST, G. 1998, 'What is So Bad About Contradictions?' *Journal of Philosophy* **XCIV**, pp. 410–426.
- SHAND, A. 1920, *The Foundations of Character: Being a Study of the Tendencies of the Emotions and Sentiments*, London: MacMillan.
- TAPPOLET, C. 2000, *Emotions et valeurs*, Paris: Presses Universitaires de France.
- WILLIAMS, B. 1986, *Ethics and the Limits of Philosophy* 2nd edn., Cambridge, MA: Harvard University Press.