

Against Emotional Modularity

Ronald de Sousa

Final draft of a paper now published in *The Modularity of Emotions*,
ed. Tappolet & Faucher, 29-50.

I. A Political Stand

How many emotions are there? Should we accept as overwhelming the evidence in favour of regarding emotions as emanating from a relatively small number of modules evolved efficiently to serve us in common life situations? Or can emotions, like colour, be organized in a space of two, three, or more dimensions defining a vast number of discriminable emotions, arranged on a continuum, on the model of the colour cone?

There is some evidence that certain emotions are specialized to facilitate certain response sequences, relatively encapsulated in their neurophysiological organization. These are natural facts. But nature, as Katherine Hepburn remarked to Humphrey Bogart, is what we were put in the world to rise above. I shall suggest that we can consider the question not merely from a scientific point of view, but from a *political* point of view. And so I will try to explain how to reconcile the evidence of emotional modularity – which, as some of the contributions to the present volume illustrate, is not devoid of a certain ambiguity – with a reasonable plea for an attitude of disapproval towards the rigidities of our taxonomy.

It may seem bizarre to speak of a political stance, since modularity is a scientific issue. And so it is: but we may have choices in the matter in two ways. First, it is far from clear just what it *means* to speak of modular emotions. So there is at least a choice of what version of the doctrine to focus on. Secondly, one can still ask whether thinking in terms of modular emotions in the relevant sense is a “good thing” or not. The facts don’t determine the attitude we take to them. There’s a long tradition that recommends accepting the Universe;¹ but one can find reasons to endorse some parts more full-heartedly than others.

In the history of philosophy, there are well-known examples of the politicization of factual and conceptual issues. Aristotle’s and Aquinas’s views on nature illustrate how one might make distinctions in the natural world between how things go and how they are *meant* to go. Aquinas can tell just by thinking, for example, what the sexual organs are *for*, and doesn’t have to be distracted by any facts about what people and animals actually do. And if that’s unreasonable in the light of the claim that what nature intends is what happens “always or for the most part,”² well, you can afford that in the Vatican, Mother of Theme Parks, because you have God on your side.

Without divine backing, the issue of what actually counts as a natural function needs to be somewhat more responsive to facts; but contemporary naturalistic philosophy can still distinguish in theory, among an organ’s actual effects, those that are its functions. Effects merely occur, from a variety of causes. But some effects are privileged by the role played by natural selection in securing the reproduction of the mechanisms that produce them, and only those count as functions (Wright 1973; Millikan 1989).

1 (“I accept the universe!” – Margaret Fuller; – “By God, she’d better!” – Ralph Waldo Emerson)

2 Aristotle, *Met.* VI–6.

This perspective raises a fresh problem, however, stemming from the fact that what evolution selects does not, even in the most metaphorical sense, have the goal of benefiting *me* – or you, or any other organism *as such*. Whatever view of genetics or fashionable stance on evo-devo you adopt, any particular organism remains a means for the transmission of heritable patterns. Indeed, the individual organism is strictly expendable where those larger “goals” might prove incompatible with the welfare of some particular vehicle such as you or me. So inasmuch as I can identify goals of my own, as distinct from those of natural selection, there is no guarantee at all that the pursuit of goals I identify as my own will necessarily be fostered by a policy of living according to nature (Buss 2000; de Sousa 2007).

Where emotions are concerned, then, it may be the case that certain forms of modularity are indeed in place as a result of natural selection, but that we have reasons to think that a *bad thing* and seek to correct it. That is what I mean to convey by speaking of the “politics” of emotional modularity.

II. ‘Modular’ and ‘Basic’: Multiple Senses

There are two ways of approaching the meaning of ‘modularity.’ One begins with Fodor’s fairly elaborate account (Fodor 1983). The other returns to basics, in the sense of taking its cue from the vernacular and in the sense of looking at what might be meant by “basic emotions.” The latter sometimes speaks of “Darwinian modules” and is associated with evolutionary psychology.³ Fodor laid down very specific criteria, on the basis of which he argued that the senses are modular but general cognition is not. Instead of quoting from that text,⁴ I’ll adopt Peter Carruthers’ (2003) summary of the defining features of modularity. In a strict sense, applicable without modification to sensory systems, modules are stipulated to be:

1. domain-specific
2. innately specified processing systems,
3. with their own proprietary transducers, and
4. delivering ‘shallow’ (non-conceptual) outputs; ...
5. mandatory in their operation,
6. swift in their processing,
7. encapsulated from and inaccessible to the rest of cognition,
8. associated with particular neural structures,
9. liable to specific and characteristic patterns of breakdown, and
10. developed in accordance with a paced and distinctively arranged sequence of growth.

On the face of it, emotions are candidates for criteria 2, 6, 7, 8, but do not meet 3 or 4. They may – at least in some cases – meet 9, though what counts as a breakdown of any particular emotional module is likely to be difficult to pin down. Criterion 10 is also probably satisfied for most common emotions, though modulated in important degrees by the role of individual experience in determining the characteristic triggers and typical manifestations of a given emotion in each individual. Criterion 1 is difficult to assess, for the modularity of emotions is supposed to relate

3 In the words of Edouard Machery, “by contrast with Fodorian modules, Darwinian modules need not be fast, automatic, cognitively impenetrable, or informationally encapsulated.” (Machery, forthcoming, 5)

4 Fodor’s own list is as follows: modular input systems or their operation are: (1) domain-specific; (2) mandatory; (3) limited in the access to the representations they compute by the central system; (4) fast; (5) informationally encapsulated; (6) shallow in their outputs; (7) endowed with fixed neural architecture, (8) have characteristic breakdown patterns, (9) and have characteristic pace and sequencing in their ontogeny. All this applies to functions that are or incorporate transducers, though Fodor does not explicitly limit his modules to systems incorporating transducers.

to output, whereas most of the modules on which Fodor originally focused were sensory modules, defined by the domain of their input. But insofar as a type of triggering situation for one sort of emotion or another can be identified, according to the hypothesis of evolutionary psychology, the difference between the range of possible inputs and the range of possible outputs is narrowed. Criterion 5 is also complicated by the distinction between input and output aspects of emotions. On the face of it, it demarcates two classes fairly clearly among the conditions we think of as emotions. Although emotions in general are often described as invading a merely passive subject with a power that is “*inescapable*” (cf. Ekman 2003, 65, citing Zajonc), that does not fit all emotions equally well, particularly if we think of the output manifestations of emotions such as facial and verbal expressions. Even in the case of the cognitive analogues of basic emotions – fear of giving offence, or fear of losing one’s job, for example – the outward manifestations of the emotion can generally be fully suppressed even if the feeling of the emotion can’t be evaded. Furthermore, insofar as the first and fifth criteria imply some form of innateness, this needs to be interpreted not as determining a fixed pattern, but as an “open” as opposed to a “closed” program in the sense defined by Ernst Mayr (1997, 694). It is not the full functioning of the mechanism that is held to be innate, but a learning bias likely to construct variants of such a mechanism in normal environments. Once the criteria are relaxed in this manner, the sharp contrast on which Fodor insisted fades, for this opens the way for cognitive modules that meet criteria 5, 6, 7 (partly), and 9. That would bring emotion modules very close to the cognitive ones, save for just one feature, namely the applicability of criterion 8. Cognitive modules are unlikely to involve dedicated neurophysiological structures of the sort now associated with “basic emotions.”

The less stringent characterization of modularity just sketched is closer to the vernacular uses of the word, as applied to devices, machines, the construction of buildings, and so forth. That yields the far more relaxed criterion, which has allowed people to construe cognition and “the mind” itself as modular.⁵

The “modularity of emotions” is an ambiguous phrase. It could mean that emotions are themselves modules. But it could also mean that emotions are put together out of modular constituents more elementary than themselves. Both notions are interesting, but for now I’m going to concentrate on the first. In a moment I’ll draw attention to one way of interpreting the second.

One way of taking the concept of an emotion module coincides more or less with “basic emotions,” though that usage too admits of many interpretations. Paul Ekman first identified as basic emotions a small set of emotions that proved universally recognizable in facial expressions (Ekman and Friesen 1975). In that sense, basic emotions seem to be both modular and innate. But basic also has several other senses: including what is *foundational*, or *atomic*, or (like “Basic English”) *frequent and important* or important because frequent.

This last conception would fit in well with that of evolutionary psychology, for trifling differences will leave no trace: *de minimis non curat selectio*. The structures that evolution has specifically erected are liable to have been worth the trouble, hence to have been sufficiently frequent and important. But as I have already implied (this will be crucial to my “political”

5 The expression often used with reference to cognitive processes is “massively modular,” implying that the mind is made up of literally thousands of modules in something like the looser (non-Fodorean) sense characterized above. (For an extended defence, see Carruthers 2006). Even where that position about cognitive functions has come under strong criticism, however, as in Buller (2005), it is generally conceded that a much smaller number of emotional modules may have a more robust psychological reality as well as being grounded in more specific neural circuitry.

stand) what is important to *me now* may have little in common with the concerns of my ancestors in the Environment of Evolutionary Adaptation (EEA).

I'll come to the functional approach in a moment. But for now let me mention a perspective on emotions in which they are viewed as (very loosely) modular rather than as modules. On this view, emotions are constructed out of elements or features that are not yet emotions.

I refer to one way in which Klaus Scherer's school of appraisal theory might be construed. Emotions might be seen as emergent phenomena that supervene on a number of appraisals, each of which defines a separate dimension.⁶ Some of the regions of this multi-dimensional state space are recognizable emotions. This sort of view need make no reference to the neurology of emotions; and no emotions are basic in the foundational sense, though attractors or "hot spots" in the space could be claimed to be "basic" in the sense of most frequent and important. Of course, since Scherer's methods tell us nothing about the underlying mechanisms that are responsible for the shape of the space, including the existence of hot spots, it is entirely compatible with the evolutionary psychology hypothesis that those hot spots are actually wired in, or wired-to-be-learned. In other words, it is not incompatible with the hypothesis of evolutionary psychology that affect programs, "inherent central mechanism[s] that direct emotional behavior" (Ekman 2003, 65) have been shaped by evolution to fulfill specific functions.

That hypothesis, however, leaves open the question of whether and how affect programs have been selected *as such*. That hangs on the answer to two questions. First, was the affect program selected *as a unit*? Second, was it *selected for*, and not merely *selected*?

To see the point of the first question, here is an analogy. Lewontin (1978) once remarked that *the chin does not exist*. What he meant was that changes in the shape of the chin resulted from relative lags in the extent of neoteny in two independent growth fields (alveolar and mandibular), subject to independent selective pressure. That makes the chin, however expressive of manly virtue or – in the "weak-chinned" – of lack of character, a mere accident, a "spandrel" of evolution. As for the second point, it was nicely illustrated by a toy conceived by Elliott Sober (Sober 1984, 99). The toy consists of a cylinder with graded sieves. Small green and large red balls are placed in the device. When the cylinder is shaken up, only the green balls get to the bottom. So they have been *selected* by the device. But since their colour plays no part in the causal explanation of that fact, what they were *selected for* is size, not colour.

An emotional module in the fullest sense, then, will be one that has been selected for, as a unit. That doesn't mean that an emotional module won't have been, like any other product of selection, cobbled together from all kinds of diverse sources. But it does imply that if it is legitimate to ascribe to it a function in the full sense, such a module will have been, at some stage of evolution, distinct enough to be favoured over some genetically and developmentally possible alternative.

To get the flavour of what it might mean in practice for some emotional characteristic to have resulted from selection without having been selected for, consider Paul Ekman's intriguing recent suggestion that moods, unlike emotions, fail the test of full functionality:

[...] emotions are necessary for our lives, and we wouldn't want to get rid of them. I am far less convinced that moods are of any use to us. Moods may be an unintended consequence of

6 (Scherer 1993; Scherer et al. 2001). I do not claim that the interpretation is the one Scherer himself intends. On the contrary (Scherer 2005) seems to favour a low-dimensional model which stresses "salient" emotions if not "basic" ones,

our emotion structures, not selected by evolution because they are adaptive. Moods narrow our alternatives, distort our thinking, and make it more difficult for us to control what we do, and usually for no reason that makes any sense to us [...]. I would gladly give up euphoric moods to be rid of irritable and blue moods. But none of us has that choice. (Ekman 2003, 50–51)

This passage is interesting in several respects. Notice first that Ekman’s argument appeals to the apparent utility of moods from the point of view of an agent in the present world. This has little bearing on the question of whether moods were selected as such.⁷ Moods may be a good example of a biological adaptation that for most people are experienced as mere nuisance. Thus Ekman’s remarks illustrate the distinction I want to stress, while at the same time underscoring the limited usefulness of making it. For let us suppose that moods are indeed more trouble than they’re worth *now*, but that they had, in fact, been carefully packaged as such by natural selection, in the sense just described. That would mean they could be adaptations in the full sense, but that fact would be irrelevant to the evaluation of their present value for us. It would make no difference to the wisdom of looking for some sort of generalized lithium-like chemical agent, where the mood’s triggering conditions cannot be evaded, to achieve the result that Ekman says he would prefer.

III. The Two-Track Mind

Moods are remarkably unitary, and share several features with paradigm emotions: they control salience in what is perceived or in preferred patterns of inference, and they lend a “style” to a whole lot of manifestations in body language, tone of voice, and dispositions to behaviour. But leaving moods aside, let us return to the small list of emotions for which the argument for modular affect programs has been made in the most compelling form. This turns essentially on two kinds of evidence. One is the identification by Joseph LeDoux (2000) and Jaak Panksepp (2001) of distinct neural pathways or “cell assemblies” and physiological profiles for affect programs. The other is the often-noted cognitive recalcitrance or “refractory” nature of emotions (D’Arms and Jacobson 2003). I’d like to set these arguments in a more inclusive context, viewing them as constituents of a larger argument for thinking of humans as language-using animals, as endowed with a *two-track mind*. By that I mean some version of the idea, broached half a century ago by Paul MacLean in several articles about the “Triune Brain,” that we have distinct “systems” in the brain that yield responses – sometimes mutually incompatible – to standard situations (MacLean 1975). The idea is compelling, but elusive, and it has taken many forms. In his recent book, Keith Stanovich lists twenty-three other versions of such “dual-process theories.” In summarizing his own version, he stresses, as central features of one system, “automaticity, modularity, and heuristic processing,” while the other is characterized as “rule-based, often language based, computationally expensive” (Stanovich 2004, 34–36).

The idea of the two-track mind has proved most controversial when applied to cognitive processes, particularly inference strategies, leading to a wide range of systematic irrationalities of the sort made notorious by Amos Tversky and Daniel Kahneman (Kahneman and Tversky 2000). Against them, Gerd Gigerenzer (2000) and others have taken what Stanovich (2004) calls the “Panglossian” view that natural selection in its wisdom has actually done everything for the best,

7 In a recent review, Nesse (2006) discusses various hypotheses, all speculative and uncertain, about the evolution of moods. If moods can be shown to result from integrated operations in neurotransmitters and neural circuits, the search for therapy in mood disorders would be facilitated. Thanks to the Editors for alerting me to Nesse’s paper. See also Charland (this volume) about depression as a possible example of a disordered modular mood.

and that the irrationalities alleged by Tversky and Kahneman are only apparent. Still others, notably Buller (2005) have rejected the arguments offered by evolutionary psychologists (Barkow et al. 1995) for the proliferation of cognitive modules under the influence of natural selection. My own view of that controversy is what Stanovich endorses as the “meliorist” view: that while the “fast and frugal” cognitive strategies favoured by natural selection (whether through hard-wiring or by the medium of learning biases) may well have been statistically optimal in the EEA, this doesn’t make them rational in every individual case. One reason for this is obvious: the EEA is not our environment, although, as I will shortly explain, this consideration may not be as powerful as it is often taken to be. Second, and more important, is the fact mentioned in section I above: that the benefits conveyed by natural selection are not measured by the success of any individual, but only by the reproduction of key patterns through generations of expendable organisms.

What holds for cognition holds also for emotion. I can concede the wisdom of nature, and equally the social wisdom that may be responsible for some of the constraints under which I live, as a member of a social group. Long tradition has the phenomenological force of given nature. Both biology and social context have lent to my emotional responses an appearance of compelling automaticity. And yet I can become aware that those emotional strategies may not be best for me. In other words, in the light of those goals that I self-consciously embrace as mine, many of my emotions may be simply irrational. The well-worn examples of fear of flying and road rage are obviously compelling here. But it can be refined and extended, by attending to the different ways in which an emotion can be appropriate or “fitting.”

Look at it first from the phylogenetic point of view. In the case of an affect program, we could think of the fittingness of an emotion as stemming from the conformity of the present situation to the situation type for which the affect program was originally selected. This sort of fittingness refers to the efficiency with which an emotion serves the function for which natural selection has designed it. As I have noted, that function is at best only incidentally beneficial to the individual in whom it is manifested. On that criterion, jealousy is fitting when it serves to eliminate actual sexual rivals or sequester a mate from possible rivals. Quite obviously, the behaviours typical of jealousy have been statistically successful in propagating the genes that favour their own manifestation. (I stress, once again, that this is not to say that these behaviours are innate, only that there is an innate learning bias that favours them.) Jealousy might be said to malfunction, however, when, in a different social environment, it results in driving the potential mate into the arms of a rival.

Even in the rawest cases of basic emotions such as fear or anger, however, there is a large range of possible triggering situations as well as a large range of possible responses. As noted above, the programs remain “open” rather than closed and the specific ways in which they are originally triggered may expand because of generalization, analogy, and association. This last, in particular, may result in idiosyncratic and pathological cases such as phobias. As for the responses, they will certainly include certain obligatory features such as those facial expressions that are seen in blind people no less than in sighted ones, and thus come with reasonable evidence of their innateness; they may also have in common certain characteristic patterns of physiological response. So much, at least, has been claimed for “anger, fear, sadness and disgust ... all marked by different changes in heart rate, sweating, skin temperature, and blood flow” (Ekman 2003, 26). This suggests that the criterion of universality picks out a class of emotions that comprise a number of physiological characteristics, measurable independently of cognitive content, and which can plausibly be attributed to mechanisms facilitated, if not entirely

controlled, by genetic factors honed by natural selection. We might say of those emotions that they are *physiologically modular*.⁸ Since some basic emotions can be combined, however, there may be some pairs of emotions that are phenomenologically “contrary” and yet physiologically compatible.

Even those physiologically modular emotions, however, will take forms that will reflect the specific scenarios in which the program was originally activated. Think, for example, of the range of functions that William Miller has claimed stem from that most visceral of our emotions, disgust (Miller 1997). These were aptly summarized in an *Observer* review of *The Anatomy of Disgust* by Anthony Storr, who remarked that “Miller rightly perceives that disgust helps to define our identities, create hierarchies, and order our world.” (Storr 1997, 16).

Such variations will shift the affect program from phylogenetic control to ontogenetic control. This will generate a different sense of “fittingness.” Where an emotion, linked to a social script, has its roots in some paradigm scenario of childhood, measures of fittingness will derive from several other criteria relating the present incident to the paradigm scenario. There will be, first, some sort of resemblance between the situation that elicits the present emotion to that original paradigm scenario. This is admittedly rather question-begging, for it is difficult to figure out what independent criteria of resemblance could be devised, besides the perceived resemblance signaled by the very fact that the emotion is elicited. But here’s the sort of thing intended: Jealousy is appropriate when a rival causes my lover’s attention to turn away from me, because that’s what was involved in my original experience of jealousy. It might be inappropriate if it is triggered by the mere fact of being on a boat, even if the defining episode took place on a boat.

But while the script fits, that implies nothing about it’s being useful, nice, or good. Whether the emotion is *morally justified*, then, is an entirely different question, pertaining to yet a third sense of “appropriateness” or “fittingness” (D’Arms and Jacobson 2000). I’ve argued before (de Sousa 2003) that the answer to that question is unlikely to be available without making tacit assumptions about emotions themselves and how they bear on the acceptability of conflicting moral judgments. But I will leave that question entirely aside here. Here is what I wish to stress instead: it follows from this sketch of different types of “fittingness” that there are various ways in which *my own emotions can impede my own goals*. On the other hand, my emotions also *set my goals*. What I care about, what I aim at, and what I desire are defined largely by my emotions, or sometimes at least by my beliefs – not necessarily true ones – about what my emotions would be in counterfactual situations.

The considerations adduced so far can be summed up as follows. While a limited number of “basic emotions” are plausibly regarded as governed by “Darwinian modules,” we can’t infer that the emotional dispositions concerned are currently adaptive from the biological point of view. Neither can we assume that they serve social cohesion, let alone the individual interests of any given person. The main reason for that is that all such modules belong to the “first track” mind, which is relatively impervious to language-mediated reasoning and to the individual and social goals elaborated on the basis of explicit deliberation. Considered from the philosophical point of view, moreover, our emotional capacities can be expected to yield a vast number of

8 Despite giving a barrage of arguments against cognitive modularity, Buller concedes that “Evolutionary Psychologists may be right about some of our more basic emotional adaptations, but nonetheless wrong in its claims that we possess a lot of *cognitive adaptations* ...” (Buller 2005, 143). The difference, he claims, is due partly to the fact that the development of the midbrain and limbic system, proceeding from the ventricular zone, differs from that of the neo-cortex, which proceeds from the sub-ventricular zone. The former, but not the latter, “appears to be under rather rigid genetic control” (Buller 2005, 131).

different experiences. What I have called the “political” approach encourages us to seek different criteria of fittingness, ranging from biological fitness proper to the promotion of subjective individual well-being. In matters of emotion, as in matters of belief, the only option is to seek in reflection a gradual harmonizing of ends and means, of long-term goals and short-term impulses, all of which are driven by different emotions. In the next section, I will ask what are the criteria that should guide our reflection in the search for reflective equilibrium about the most desirable emotional life.

IV. PEGGING Emotions

We will need, at some point in the process of reflection, a taxonomy of emotions, and that will require us to decide how important the existence of modular affects or other emotions should be to our taxonomy. A few years ago, Paul Griffiths (1997) insisted that a viable taxonomy could be grounded only in genuine homologies; but this rather dogmatic view has been fairly demolished by Louis Charland (2002), and it also faces a plausible alternative in Dick Boyd’s reconstruction of a functional concept of homeostatically stable kinds (Boyd 1999). Rather than embark on that debate, let me take a different tack and suggest a bifurcation of emotions based on the roles that they play in our lives. The two classes are non-exclusive, overlapping, and somewhat vaguely defined; nevertheless, the distinction is of paramount importance.

Among my goals and interests, some involve the cultivation of desirable emotions of the sort that contribute intrinsically to happiness. The pursuit of such emotions isn’t necessarily purely selfish. It may include the cultivation of desirable emotions for my children, my friends, and others. But the pursuit I have in mind under this heading aims essentially at the cultivation of emotional *quality*. My emotions, under this heading, matter for their own sake.

Contrasting with those, I have goals and interests that necessitate accurate *Predictions, Explanations, and Generalizations in Gossip*, or PEGGING of how others will act and respond.⁹ Here the *communicative* function of the emotional expressions comes into its own. It is in general to the subject’s advantage, at least for the class of emotions linked to affect programs, to allow others to see what they are feeling. Hence the hard-wired universal expressions that Paul Ekman has anatomized. Unlike language, the second-track device given to humans, as Talleyrand reportedly quipped, for the purpose of concealing one’s thoughts, facial expressions require a special effort to control. And perfect control is seldom achieved, at least for the first microseconds of any episode of emotional arousal. Furthermore, from the point of view of folk psychology, it is handy to have the simplest possible toolbox of explanatory concepts under which the behaviour of others and their states of mind can be PEGGED.

So here is the hypothesis I propose:

Hypothesis H: Insofar as I am interested in PEGGING the behaviour of others, a neat scheme involving relatively few elementary states (which could then be combined to make more complex states) will serve me best. It will be a scheme which as much as possible will involve a digital system of representation.

A digital system of representation is one in which every state of what is represented is forced into one or another of a finite number of pre-assigned possibilities. The alphabet is a good if imperfect example. Providing we assume that it was intended as a piece of English, any

9 The importance of gossip in the development of a larger and more powerful brain has been argued convincingly by Robin Dunbar (1996).

handwritten squiggle will be assigned to one or another among our 26 letters, 10 numerical symbols, and a few other marks standing for punctuation or mathematical operations. An ‘a’ may be confused with a ‘d,’ but there is nothing between an ‘a’ and a ‘d’ that an ambiguous mark can be assigned to. What such a system does, in effect, is to construe resemblance between any two letters not as a two-term relation, but as a three-term relation involving the two squiggles and the paradigm, the actual letter or symbol A, of which they are both meant to be instances.¹⁰

As these examples suggest, though, there are – unfortunately for the apparent neatness of the distinction – degrees of digitality when a system of representation is envisaged as a whole through a period of time. Language, for example, changes much faster than genes: linguistic mutations large or numerous enough to affect the entire population occur in timespans of the order of decades, not millions of years. And our conceptions of emotions, once we go beyond affect program modules, may be subject to gradual change, more like a traditional dance than a phoneme.

Our conceptions of emotions may be even more difficult to analyze than the evolution of linguistic forms. The elements we start with are more complex, and the combinatorial possibilities more numerous. Yet if Hypothesis H is right, there will be additional pressure to pare down our vocabulary of emotion in order to simplify our mastery of that complex domain. A corollary of the dominance of the context of PEGGING over the context of contemplation and experience is that we will, out of sheer sloth, likely come to use the same vocabulary in thinking of our own emotions, as if it comprised an adequate taxonomy of emotions for situations in which multiple reproduction is not at issue, but only *experience*.

Here is one more reason for the dominance of the needs of explanation and prediction over the value of quality of experience. Emotions are typically thought of as motivating, or as involving characteristic “action tendencies” (Frijda 1986). Insofar as emotions motivate behaviour, we can think of them as just pushing and pulling: swayed by the totality of your emotions at any particular time, either you act or you don’t. The way we *experience the world* can be as complicated or as subtle as you like, but in the end, where agency is in the offing, they have to be funneled into a single sequence of exclusive decisions. Each decision is a matter of acting or not acting; it is the black-and-white of yes or no. If we focus on the experience of emotions, on the other hand, they are so diverse as to constitute no single *kind* of thing at all. (Think again of the multidimensional space of appraisals which, on Scherer’s scheme, is the matrix of emotions.) Each carries a wealth of specific meanings enriched by an immensely large class of contrasts: call it the polychrome vision of the emotional field.

V. The Aesthetic Model

On such a full-colour view of emotions, geared not to the requirements of agency but to the realities of emotion as experience, it could be argued that 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 (Campbell 1998). This view seems particularly compelling when we consider aesthetic experience, a domain in which, as we have all learned from Kant, we are able to contemplate aspects of the world for their own sake, in abstraction from practical

10 This scheme is just a new version of Plato’s Theory of Forms. Plato failed to grasp its true significance, however, which is that it is invaluable in any case of *serial reproduction*. It pre-empts the Xerox effect, which is the degradation of copies of copies at the end of a long line of copying operations. For since reference is made to the paradigm at each stage, even the ten millionth copy is still just two steps away from the original. It is not surprising, then, that the two most spectacular examples of digital representation in the natural world are language and DNA, both of which are in the business of limitless reproduction requiring a staggering level of fidelity.

considerations. To that extent, the emotions aroused by works of art are not directly tied to any goals. (Or, if you prefer, they are tied only trivially to the goal of continued contemplation.) Reading poetry, looking at paintings, watching dance, or listening to music would be largely pointless activities unless they aroused emotion.

But what sort of emotions? The observation that art evokes emotion is threatened with two unequal and opposite forms of triviality. On the one hand, one may be tempted to think that the emotions expressed in art are “the grand emotions” that we can all list on demand: anger, fear, love, awe, jealousy, sadness, desire. But if that is the point of them, then why go to all the trouble of making (or consuming) new and original art? If works of art exist merely to evoke standard emotions, and if there are no significant differences between any two instances of “fear,” or “anger,” and so forth, it hardly seems likely that works of art in all their diversity should be sustaining our interest for their representation of *emotion*, rather than for some other reasons. On this view, it’s difficult to see why most art isn’t superfluous. (Most mediocre art is indeed superfluous in just this sense: once you’ve seen one scary alien-invasion movie, you’ve seen them all.) The alternative view is that each different moment in music, dance, or literature that evokes emotion is actually expressing an emotion *sui generis*, or better *sui ipsius*: not merely of a certain sort but in its own unrepeatable individuality.

But then the role of art is trivialized again: for if every piece of art necessarily expresses its correlative emotion, no more and no less, then that seems to remove the possibility that some forms of art might be better or worse at expressing emotion, or that the emotions evoked by some work of art might be more worthwhile, more interesting, more deeply felt, more authentic than others. For all those evaluative judgments seem to presuppose that there is something beyond the expression of emotion, in terms of which a given expression can be judged.

This objection can be overcome if we give up the assumption that each work of art or literature is seen as conveying a ready-made emotional “message.” If instead art is thought of as creating and embodying a particular emotion of its own, valued for its own sake, then there can be numberless emotions, and every work of art is more or less interesting in accordance with the quality of the unique emotion it conveys. There are then indeed literally *innumerable* emotions.

But now we face a problem of a different sort.

In the exploration of an aesthetic domain, we learn to discriminate, to compare, and to retain what most seems of significance. Present experience guides the superior refinement of future experience. Now if we concentrate on the unique features of each situation, on the specific qualities of each individual, and on the singularity of each emotion, it is difficult to see how they could provide any guidance at all. Comparisons require similarities and differences, classified and conceptualized in terms that necessarily return us from the particular to the general. Learning, in short, requires repeatable patterns. Taken literally, then, the suggestion that each particular situation in life (as well as each episode in a work of art) gives rise to an unrepeatably unique emotion is self-defeating. (Particular) percepts without (general) concepts cannot be refined.

If that problem seems rather unreal, it is because it takes the particularity of emotions too far. Experience, like all forms of cognition, is intrinsically *general*. It can be indefinitely *specific*, and it can have particulars as its objects, but it can’t contain the infinite properties of any particular as such. Even in the experience-centred utopia I am envisaging, then, non-conceptual contents of experience can be assessed along a variety of continua: valence, intensity, similarity to some paradigm scenario, and the set of associations that give them meaning. That field of emotional existence would not necessarily be devoid of privileged “hot spots,” but it has no need

of being digitized, that is, of being conceptualized in terms of a finite number of pre-defined emotions.

The last few paragraphs have described a utopia of liberated emotional experience. Yet the reality is that we are all too ready to settle for a simplistic taxonomy, modeled on modules, with its set stack of standard labels. Such a schema will not serve my own quest for a rich and nuanced emotional life. Yet as I have suggested, we are liable to assume that if the scheme works for us when PEGGING the emotions of others, it has to be right for ourselves. Unless I explicitly question it, I'll ascribe to myself the same limited set of possible emotions that I use to make sense of the lives of others. I may tend to take it for granted, furthermore, that the basic scenarios emotions are geared to are well understood and pretty much permanent. And in this, as in most things we are inclined to take as "natural," there may be a strong component of social pressure – possibly a mere special case of some general rule of conformism (Dugatkin 2001) – that forces anomalous cases into a marginal status.

For that reason, the effect of the difference between our own environment and the EEA is less important than it would seem. The way we interpret our objective situation depends on the reactions we observe in others. (The classic psychological experiment that shows this, though it doesn't show most of what it's usually dragged in to show, is that of Schachter and Singer 1962.) So as we watch others *interpret* the current situation in terms of atavistic scripts, we may assume that such scripts fit all. And that, in a sense, will *make* it fit, or at least will make it true that it fits in the minds of those agents PEGGING the interactions concerned.

Here is an analogy. We commonly take it for granted that there are just two sexes. On strictly biological grounds, Anne Fausto-Sterling (1993) has urged us to recognize at least five sexes, while noting that this is still a crude simplification of what in practice is really a *continuum*. "I would argue further," she writes, "that sex is a vast, infinitely malleable continuum that defies the constraints of even five categories" (Fausto-Sterling 1993, 21). But in the medical world as well as in the public consciousness, this is simply denied. Infants born androgenous are "assigned" or "reassigned" to one or the other of the two obligatory sexes. And, needless to say, in the public discourse surrounding the current debate about "same-sex marriage," there is little consideration of bisexuality, and virtually none at all of intersexuality or of the potential multiplicity of sexes and genders. To be sure, the biological and functional reasons for sexual dimorphism in both nature and discourse are fairly plain – though the details remain obscure enough to give rise to conflicts of medieval irrationality.¹¹ But the result, for those who don't easily fit either of the two obligatory options, is a brutal denial of the emotional truth of their experience, indeed of their very identity.

VI. Conclusion

I have suggested that there are two sources of potential conflict between the modularity of emotions, whether stemming from phylogeny or from social conformism, and our aspiration to a life of greater emotional richness. The first source is the existence of pre-programmed affect syndromes. These encourage our tendency to respond in ways that may frustrate our own goals and self-image. The second stems from the undeniable usefulness of a clear schema for PEGGING other people's behaviour. This falsely instills the conviction that the representational

11 As sadly exemplified by the outrage over some speculative remarks by the President of Harvard, even among sophisticated scientists presumably familiar with Bell curves, but who behaved as if an allusion to differences in *variances*, implying overlapping continua, connoted rigidly separate classes. See Murray (2005) for a discussion of the ramifications of that case.

scheme simply represents reality, and that (to put it excessively simply) the number of our emotion words is a sure guide to the number of emotions it is possible to experience.

There's not much we can do about either of these constraints on our capacity to enlarge our emotional repertoire, though some techniques have been plausibly recommended. Paul Ekman (2003, 76) reports, for example, that attentiveness, reappraisal, and "mindfulness meditation" have had some success in increasing the "impulse awareness" that may enable an agent to control the automatic responses of the intuitive track. But for my purposes, I am more interested in the way that literature and art can place us at one remove from our impulses, enable us to play down the importance of action tendencies, or simply perhaps view them as merely another aspect of intrinsically interesting experience. Reading Proust, D.H. Lawrence, or Henry James, or for that matter, if you can bring yourself to do it, the Marquis de Sade, we can be made aware of ranges of emotions lying outside the standard repertoire. This is a point made repeatedly by Martha Nussbaum, who has stressed the capacity of literature to present us with fully imagined emotions and particular characters, as contrasted with the bloodless abstractions customarily paraded in philosophical examples (Nussbaum 1992). The emotions evoked by art, music, and literature are typically enjoyed in abstraction from the practical considerations in terms of which standard emotions are deemed more or less appropriate. They need not carry with them any specific action tendencies. Instead, they constitute a "full-colour," multidimensional field of possible emotional experience that provides a model for the emotional richness that might be afforded by ordinary day-to-day existence.

Such a field would not necessarily be devoid of privileged hot spots but has no need of being digitized or conceptualized in terms of a finite number of pre-defined emotions. Armed with such examples, we might be better equipped to *resist* our own tendency to think of our own and others' emotions in terms of the limited vocabulary we use to PEGG people's motivations and behaviour.

And that – here in capsule is my political message – would be a Good Thing.¹²

12 My thanks to the editors and to anonymous referees for many suggestions on an earlier draft.

References

- Barkow, J. H., L. Cosmides, and J. Tooby. [1992] 1995. *The Adapted Mind*. Oxford: Oxford University Press.
- Boyd, R. 1999. Kinds, complexity and multiple realization: Comments on Millikan's "Historical Kinds and the Special Sciences." *Philosophical Studies* 95: 67–98.
- Buller, D. 2005. *Adapting Minds*. Cambridge (Mass.): MIT Press.
- Buss, D. M. 2000. The evolution of happiness. *American Psychologist* 55(1): 15–23.
- Campbell, S. 1998. *Interpreting the personal: Expression and the formation of feeling*. Ithaca: Cornell University.
- Carruthers, P. 2003. The mind is a system of modules shaped by natural selection. In *Contemporary debates in the philosophy of science*, ed. C. Hitchcock, 293–311. Oxford: Blackwell.
- Carruthers, P. 2006. *The Architecture of the mind*. New York: Oxford University Press.
- Charland, L. C. 2002. The natural kind status of emotion. *British Journal for the Philosophy of Science* 53(3): 511–37.
- D'Arms, J., and D. Jacobson. 2000. The moralistic fallacy: On the 'appropriateness' of emotion. *Philosophy and Phenomenological Research* 61: 65–90.
- D'Arms J., and D. Jacobson. 2003. The significance of recalcitrant emotion (or quasi-judgmentalism). In *Philosophy and the Emotions*, ed. A. Hatzimoysis, 127–45. Cambridge: Cambridge University Press.
- de Sousa, R. 2003. Paradoxical emotions. In *Weakness of Will and Practical Irrationality*, ed. S. Stroud and C. Tappolet. Oxford: Oxford University Press.
- de Sousa, R. 2007. *Why Think? Evolution and the Rational Mind*. New York: Oxford University Press.
- Dugatkin, L. A. 2001. *The Imitation Factor: Evolution Beyond the Gene*. New York: Simon & Schuster.
- Dunbar, R.I.M. 1996. *Grooming, Gossip, and the Evolution of Language*. London: Faber and Faber.
- Ekman, P. 2003. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Henry Holt.
- Ekman, P., and W. Friesen. 1975. *Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions*. Englewood Cliffs (NJ): Prentice-Hall.
- Fausto-Sterling, A. 1993. The five sexes: Why male and female are not enough. *The Sciences* 33(2): 20–25.
- Fodor, J. 1983. *The Modularity of Mind*. Cambridge (Mass.): MIT Press.
- Frijda, N. 1986. *The emotions. Studies in Emotion and Social Interaction*. Cambridge, Paris: Cambridge University Press, éditions de la maison des sciences de l'homme.
- Gigerenzer, G. 2000. *Adaptive Thinking: Rationality in the Real World*. New York: Oxford University Press.
- Griffiths, P. E. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Kahneman, D., and A. Tversky, ed. 2000. *Choices, Values, and Frames*. Cambridge, New York: Cambridge University Press.
- LeDoux, J. E. 2000. Emotion circuits in the brain. *Annual Review of Neuroscience* 23: 155–84.
- Lewontin, R. C. 1978. Adaptation. *Scientific American* 293: 156–69.
- Machery, E. Forthcoming. Massive modularity and brain evolution. *Philosophy of Science*.
http://www.pitt.edu/AFShome/m/a/machery/public/html/papers/Evolutionary%20Psychology%20and%20Brain%20Evolution_PSA_2006_machery.pdf.
- MacLean, P. D. 1975. Sensory and perceptive factors in emotional functions of the triune brain. In *Emotions: Their Parameters and Measurement*, eds. L. Levi. New York: Raven Press.
- Mayr, E. 1997. Behavior programs and evolutionary strategies. In *Evolution and the Diversity of Life : Selected Essays*, 694–711. New York: Belknap Press.
- Miller, W. 1997. *The Anatomy of Disgust*. Cambridge (Mass.): Harvard University Press.
- Millikan, R. G. 1989. In defense of proper functions. *Philosophy of Science* 56: 288–302.
- Murray, C. 2005. The inequality taboo. *Commentary* September: 13–22. Fully annotated version at
<http://www.commentarymagazine.com/production/files/murray0905.html>.
- Nesse, R. M. 2006. Evolutionary explanations for moods and mood disorders. In *American Psychiatric Publishing Textbook of Mood Disorders*, ed. D. J. Stein, J. Kupfer and A. F. Schatzberg. Washington (D.C.): American Psychiatric Publishing.
- Nussbaum, M. 1992. *Love's Knowledge*. Oxford: 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): 141–63.
- Schachter, S., and J. Singer. 1962. Cognitive, social, and physiological determinants of emotional states. *Psychological Review* 69: 379–99.
- Scherer, K. R. 1993. Studying emotion-antecedent appraisal process: An expert system approach. *Cognition and Emotion* 7(3–4): 325–55.
- Scherer, K. R. 2005. What are emotions? And how can they be measured? *Social Science Information* 44(4): 695–729.
- Scherer, K. R., A. Schorr, and T. Johnstone, eds. 2001. *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford: Oxford University Press.
- Sober, E. 1984. *The Nature of Selection*. Cambridge (Mass.): MIT Press.
- Stanovich, K. 2004. *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: Chicago University Press.

Storr, A. 1997. Why we don't like dog turds. Not even chocolate ones. Review of W. Miller 1997. *The Observer*, London, April 13.

Wright, L. 1973. Functions. *Philosophical Review* 82: 139–68.