

Rationalities of Emotion - Defending, Distinguishing, Connecting

Sophie Rietti

Abstract: Claims that emotions are or can be rational, and crucially enabling of rationality, are now fairly common, also outside of philosophy, but with considerable diversity both in their assumptions about emotions and their conceptions of rationality. Three main trends are worth picking out, both in themselves and for the potential tensions between them: accounts that defend a case for the rationality of emotions A) by assimilating emotions closely to beliefs or judgements; B) in terms of the very features that traditional views of emotions as irrational/a-rational emphasized; C) by arguing that emotions exhibit a more sui generis kind of rationality, often one based on a narrative or dramaturgic “inner logic”.

Keywords: emotion, rationality, cognitive theory of emotion.

In the face of such diversity, certain responses are tempting, but should be resisted: 1) that the proponents of these views are just talking past each other; 2) that the approaches are necessarily mutually exclusive; 3) that if any of them can be true, it must be just one of them. Rather, the diversity reflects both the range of different ways in which we conceive rationality (which may not be legitimately reducible to some simpler unitary account), and the complexity and range of our emotional lives *qua* cultural thinking animals. A sufficiently fine-grained account of both rationality and emotion can – and should – accommodate these points. (All arguments on these issues, of course, being provisional in the light of potential further developments in our understanding of both emotions and rationality.)

Introduction

Claims that emotions can be rational, or even that they are so more often than not, are now fairly common in the literature on emotions, in philosophy and other disciplines, social as well as natural sciences included. There is also a growing literature, perhaps most notably in biologically oriented psychology and economics, that holds that rationality even *requires* the capacity for emotional responses. Such views cast themselves against a (supposed) tradition of viewing emotions as paradigmatically irrational or non-rational, a tradition to which philosophers are held as main though not sole contributors. Leaving aside that this is questionable in itself (usual suspects such as Hume and Descartes held rather more sophisticated views of emotion than they are often credited with, also in terms of the rationality of emotion), a striking point about the current views is their diversity, when examined more closely, both in their assumptions about what emotions are and in their conceptions of rationality. This may partly be due to the complexity of both topics in their own right, of course: combining was never likely to simplify matters. Nevertheless, particular views of emotion on the one hand and of rationality on the other often align in ways that may seem initially surprising.

A comprehensive overview would be a much too great task here (if it could be done at all), but three main trends are worth picking out, both in themselves and for the potential tensions between them. These appeal, respectively, to A) accounts of emotions as being more like (or even identical to) judgments than “brute feelings”, and therefore suitable for assessment on fairly traditional kinds of rationality-criteria B) ways in which the very features (“bruteness” and recalcitrance to reason and will) often invoked to argue that emotions are irrational or non-rational may play a part in making them rational, or in enabling agents’ rationality, albeit by sometimes revisionist rationality-criteria C) ways of tracing the “inner logic” of our emotional lives to point to emotions as possessing a *sui generis* rationality.

I shall argue that while it is important to keep these approaches distinct, it is also important to resist the too-easy temptation to assume that they are mutually exclusive or talking past each other, or that at most one of them can be true if any of them are. But to see this, it will first be necessary to examine what they say about both rationality and emotions in some more depth. In what follows, I shall start by distinguishing some

of the main relevant rationality-criteria at stake, and how these initially bear on emotions. I will then examine how these assumptions about rationality align, in the three accounts under discussion, with their different descriptive assumptions about what emotions are, as a basis for assessing the three contrasting approaches, in themselves and against each other.

Some main relevant distinctions in rationality-criteria

Minimal rationality

The literature on rationality is of course vast in its own right: what will be said here is primarily to provide background to how rationality criteria bear on emotions, and to introduce relevant terms as they will be used here – given that they are often used in a bewildering range of ways across the field. To begin with, then, some fairly traditional distinctions can be made first between phenomena that are viewed as not assessable in rational terms at all (the non-rational or a-rational), and phenomena that can be assessed, whether negatively or positively, by rationality-criteria. On a Davidsonian account, for instance, one might distinguish between phenomena that can be given a “merely” causal explanation and those that can be explained by reference to reasons – albeit noting that reasons (at least on Davidson’s own approach) are typically understood here as a particular subset of causes (see Davidson (1963)).

Reflexes, such as the blink reflex or the kneejerk reflex, would tend to be viewed as paradigmatically non-rational on this account. Intentional states such as beliefs, by contrast, will generally be at least *minimally rational*; that is, it is assumed that adequate explanations of them will need to invoke reasons, though those reasons may fall short of providing full justifications by more robust *evaluative* criteria by which we judge reasons good or bad *qua* reasons. Intentionality, though, tends to be viewed as a necessary but not necessarily sufficient condition of minimal rationality: the “aboutness” of beliefs and desires, by contrast with supposedly non-intentional states such as pain-sensations, is crucial for their rationality. This is not simply a matter of object-specificity: reflexes and pain sensations, after all, can have very specific kinds of “trigger” objects, without thereby qualifying as intentional. While the point is often left largely implicit, the intentional/non-intentional distinction, insofar

as it touches on the rationality of emotions (or other mental kinds) often turns on whether the state in question is thought to involve some form of mental *representation* of the object.¹

In relation to emotion, then, the immediate issue is that any account of emotions as being primarily “brute feelings” that track specific objects only in fairly brute, causally explicable ways, will tend to make their intentionality a contentious issue and thus undermine their claims to be even minimally rational. That emotions can at least seemingly “come over” us without clear reason would further undermine their claim to be based on reasons. An immediate challenge for any defender of emotion’s rationality, then, would be to argue that these issues are merely apparent, or at least atypical. An alternative tack, already indicated above, and which will be examined in some more detail below, is to revise, or diminish the importance of, the intentionality-criterion. A more frequent approach, at least in the philosophical literature, is to appeal to ways in which emotions *do* appear to be responsive to considerations that would count as reasons, and of characteristic kinds, as when for instance fear tracks threat.

Evaluative rationality

Assuming reasons *can* be invoked, then, these reasons can in the next instance be applied to assess the intentional state, for instance a belief or desire, as evaluatively rational (the reasons for the belief are good, and support it) or irrational (the reasons fail on either or both counts). Similar points apply, *mutatis mutandis*, to motivating states and intentional behaviours. Theorists vary as to what reasons count as good, and how strictly such criteria of rational warrant are applied: for instance, whether good-faith subjective warrant is enough to qualify as rational even if, unbeknownst to the agent, some of the premises are false, or the agent’s ways of forming beliefs and motivations are less than (evaluatively) rational.² The responsiveness of the intentional state to warrant-relevant considerations will also typically be crucial in assessing its rationality – a belief that does not change in the face of available counter-evidence (of sufficient strength) will by that token be irrational, and the agent may be

¹ See Deigh (1994) for some further discussion of intentionality and representation, as well as the relation of both to rationality, as it concerns emotion.

² See e.g. the discussions in Smith (1995), Arpaly (2000) and Jones (2003).

judged as irrational with it. That emotions seem prone to such recalcitrance – as when fear may not dissipate even where the agent himself believes there is no real threat – is of course a traditional reason to be skeptical of emotions here.

An ongoing background worry here, which also relates back to the initial distinction between the non-rational and the minimally rational, is that mental states, and perhaps especially emotions, may lack *transparency*: we cannot always reliably track whether there are reasons for them, or what those reasons are. And this problem is compounded, in the next instance, by our propensity to rationalize and confabulate.

For instance, in Schachter and Singer's (1962) famous social psychology experiment, subjects who were left ignorant of or misinformed about the physiological effects of an adrenaline injection were consistently more likely to engage in behaviour suggesting effective "mood contagion" of anger or euphoria from a stooge planted by the experimenters, and to self-attribute relevant emotional states. Subjects who had been primed to be accurately aware of the likely effects of the injection were less likely to be receptive to mood contagion or self-attribution of emotion than either of the other groups. (Subjects who had received a placebo injection containing saline solution, and no indication of what symptoms to expect, were more receptive to suggestion than informed "adrenalized" subjects, but less than ignorant or misinformed "adrenalized" subjects). So a state that is not even minimally rational, or intentional, can appear to the subject to be both, and lead the subject to falsely attribute its origins. We will return to the implications – and questions of their seriousness – further down.

Cognitive and strategic rationality

Another traditional distinction is that between cognitive and strategic rationality-criteria. These are distinguished primarily by the different goals at which they aim (De Sousa 1980). *Cognitive rationality* aims at truth-aptness (or, in the case of evaluative attitudes, to such truth-analogues as these are considered capable of, if any), and applies most obviously to beliefs. *Strategic rationality* aims at the satisfaction of preferences, and applies most obviously to desires (in the broadest sense of motivating states) and purposive behaviours. The cognitive/strategic distinction cross-cuts the minimal/evaluative distinction indicated above.

Accounts vary, again, as to whether cognitive and strategic rationality are seen as fully distinct forms, or whether either of them is ultimately reducible to the other (is truth-seeking just another form of preference?). I shall not attempt to resolve these issues here, but they will be of some importance regarding emotions, and so worth keeping in mind throughout.

The issue of which kind of rationality-criteria – strategic or cognitive – are applicable also arises with some particular force for emotions. For emotions typically have motivating force of characteristic types, as when fear inclines to avoidance or flight, and so it might seem most natural to assess emotions, and behaviours driven by them, for strategic rationality. A complicating factor here is the question of whether the motivating effect is “properly” part of the emotion, or rather comes from a desire to which the emotion is conducive but which can still be distinguished from it.

Another aspect of emotions, though, is more like belief: our experiencing of the object the emotion is directed towards in particular characteristic ways (as threatening, or offensive, or good, or loveable), which can at least be distinguished, if not fully separated, from motivations to act in particular ways towards it. And for these aspects of emotion, cognitive rationality-criteria, or some analogue, seem most applicable. The question of what would count as warrant for an emotion here also raises what De Sousa (1980, 1987) calls the Euthyphro-problem of emotions: are things e.g. frightening because we typically get frightened of them, or do we get frightened of them because they are frightening? These considerations also raise with some force again the question of what the emotion itself is understood to be: just the hedonic tone or “feeling” of an occurrent emotion? Desire, belief, or some analogue of either? Or a desire-belief complex with accompanying “tone” which may or may not be epiphenomenal?³

For immediate purposes, it is also worth noting that strategic and cognitive rationality may diverge, and that this may even raise a particular problem about emotions, if emotions tend to straddle the cognitive/strategic distinction. For instance, sadness or even despair about one’s circumstances might have cognitive warrant (the outlook may be genuinely bleak) but be strategically irrational relative to the goal of

³ See e.g. Shaffer (1983) versus Oakley (1992) on this point.

ameliorating or freeing oneself of the situation. Note that if both the construal of the situation *and* the action-tendencies are considered characteristic parts of the emotion-*type* (here, sadness), their not going well together from the point of view of rationality may have troubling implications, possibly also for other emotion-types, at least unless this kind of case can be dismissed as an anomalous (if unfortunate) token.

Instrumental and substantive rationality, antinomies of rationality

Emotions have also been argued, as we shall see in more depth shortly, to be crucial for rationality in ways that bear on long-standing antinomies in our views of rationality itself. As already suggested above, for one, we can also distinguish between instrumental and substantive rationality. A state or behaviour is instrumentally rational insofar as it is conducive, or plausibly believed to be so, to the attainment of a goal. Goals can, as suggested above, include truth-finding: from this point of view, cognitive as well as strategic rationality would be instrumental. The state or behaviour is *substantively* rational insofar as the goal is itself rational.

Again, there is considerable variation as to whether particular authors think there can be such a thing as substantive rationality: Humeans, notoriously, tend to deny that goals can in themselves be assessed for rationality. The limit of reason's capacities, on the traditional reading of Hume at least, is fact-finding and the working out of means to ends. Ends cannot themselves be assessed for rationality, except by such formal, instrumental criteria as consistency and compatibility – reason can point out to us that our goals fail in these respects, but no more. And crucially, it is emotion, instead, which sets the goals, and identifies things as good or bad for us: reason has at most an ancillary and instrumental role in this, important as that role may be.

Emotion has also been held by some (De Sousa 1987, Damasio 1994) to have a crucial role to play in such cases as decision between options that are otherwise equally ranked, choice between incommensurable options, and decisions under uncertainty. Some have also more generally emphasized the use of emotional reactions as heuristics, advocating the use of “gut feelings” to cut through deliberation processes. On a traditional view of rationality, these may not count as ways in which emotions in themselves are strictly speaking rational. Rather, similarly to the Humean account, emotion goes beyond what reason is capable of,

and may thereby serve to resolve the knots reason creates or leaves for us. Since this kind of position broadly corresponds to the second kind of account of emotional rationality to be discussed here, it will be examined in more depth below.

Emotional rationality: three approaches

With this background in mind, then, we can now turn to the three kinds of approaches to emotional rationality at stake here. The first makes a case for the rationality of emotions by giving an account of emotions themselves that promises to make them more easily assessable by – and even justifiable by – traditional criteria of rationality. The second allows for many of the charges, including bruteness, that traditionally underpin the rejection of emotions as non-rational or irrational, but appeals to ways in which just these features may also make emotions, if not rational in themselves, crucial for the rationality of the agents who have them. The third argues for a *sui generis* rationality for emotions, which are typically neither viewed as brute nor as like judgments or beliefs, but driven instead by an “inner logic” of their own. To see more of what is actually at stake for these accounts, then, we shall examine each of them in turn more closely.

Strong cognitive theories of emotion: judgmentalism

Within philosophy, a main way of arguing for emotion’s rationality against (supposedly) traditional assumptions of its irrationality or non-rationality has been provided by the rise of so-called cognitive theories, where emotions are understood to essentially involve cognitive appraisal of intentional objects. The further specification of “involve” however admits of a broad range of interpretations: cognitions are variously offered as causal antecedents and consequents of emotions (sometimes both), as their constituent parts, and as the whole of what an emotion is. “Cognition” also in itself covers a broad range of notions, from relatively weak senses such as construal to strong senses such as judgments, and

sometimes, rather confusingly, gets applied to what otherwise seems more like perception or perception-analogues.⁴

Since these theories, when used to make a case for the rationality of emotions, typically appeal to the cognitive element, however this is then further defined, it will be instructive to examine such theories in their most unequivocal form, that which both holds that the cognitive element in emotion is (a form of) judgment, and then identifies the emotion just with this judgment. For short, this position will be referred to as judgmentalism. Martha Nussbaum (1994, 2001) and Robert Solomon (1973/1993, 2006) are perhaps the most striking defenders of such a view. Under this heading we could also include, as inspirations for their views, respectively the Stoics and Sartre, though there is some contention as to whether the Stoics held a full-on identity-theory of emotions and judgments,⁵ and Sartre's account is primarily concerned with the *projective* nature of emotion-cognitions.

The starting-point for both authors is a case for emotions as states of rich intentional and cognitive content, and a related rejection of "brute feeling" views. It may be felt that the move from this to claiming that emotions are judgements, let alone based on and responsive to reasons in the way that beliefs and judgements are supposed to be is achieved a little too quickly – and possibly with some neglect of the ways in which beliefs and judgements may themselves fail in such regards. (Though both Nussbaum and Solomon emphasize this latter issue in defending their views against counter-arguments that appeal to recalcitrant emotions: beliefs and judgments, too, they argue, can persist in ways that are irrational or non-rational).

It is worth noting that neither Sartre nor – especially – the Stoics took a particularly high view of the *evaluative* rationality of emotion. Their accounts allow emotions to be assessed for rationality (that is, they are minimally rational), but also tend to give rather a derogatory assessment of them by these standards, once applied. The derogatory view, in the Stoic case, comes from a negative assessment of emotion's substantive

⁴ See e.g. Armon-Jones (1991, Chapter 1) or Deigh (1994) for further critical overviews on varieties of cognitive theories; see also Charland (1994) for the distinction between perceptual and cognitive accounts.

⁵ See e.g. Sorabji (2000) for further discussion of the Stoics' view on this point and the competing interpretations of their views in the secondary literature.

rationality: for emotional reactions attach importance for our flourishing to things outside of our control, and the Stoics hold this view of flourishing to be false, as well as a pernicious influence on behaviour. Also, such attachments are not imposed on us, but arise through voluntary (though not necessarily knowing) irrationality: natural impulses may incline us to attach to what is outside our control, but only the voluntary assent of our rational capacities allows the natural impulses to become full-fledged emotions.⁶ In other words, emotions as a class appear to be evaluatively irrational here.

In the Sartrean case (Sartre 1948), emotions *qua* judgements are viewed as strategically chosen ways of projecting one's own values and goals onto the world, often without taking due good-faith responsibility for the projection. Pretending to oneself that things possess features one has projected onto them is, from the point of view of rationality-criteria that aim at producing true belief, of course questionable – that doing so may (but need not always) be strategically, instrumentally rational from the point of view of helping to bring about one's goals while shirking inconvenient levels of responsibility for them is another matter. Given that Sartre's account of value-ontology is also more generally projectivist, though, the problems are primarily about bad faith and the generally self-serving nature of the projections than about their status of projections as such.

Solomon and Nussbaum both hold more positive views of the evaluative rationality of emotion than their inspirations, though: Nussbaum allows for, even requires, a greater degree of attachment to externals (including other people) in her notion of the good human life than the Stoics, and Solomon, to a greater degree than Sartre, takes an optimistic view of what our emotional lives can be like once we realize they are in our power to control. For present purposes, the main initial point is that judgmentalism holds emotions to be within the realm of rationality-assessment, whether the outcome of that assessment is negative or positive. The voluntarism about judgements in both Solomon's and Nussbaum's accounts also suggest that emotions are things we actively bring about, and so more like actions than beliefs (or passions): while cognitive theories generally emphasize a case for cognitive rationality in emotions,

⁶ At least in some of the later, Roman Stoics this is an explicit point: see for instance Seneca's discussion in "On Anger" (Seneca 1995).

this kind of judgmentalism suggests strategic rationality-criteria may also (or even preferably) apply. That this also suggests the judgments involved in emotions are a form of “deciding to believe” may even raise problems for assessing emotions by cognitive rationality criteria (Williams 1973), which traditionally hold that reasons to believe should be derived from evidence and inferential warrant, not that one wants to believe because doing so will serve one’s interests.

Some main difficulties facing this type of account come rather on two different fronts, though. First, from resistance to the descriptive account of emotion offered. For one, unless judgment is meant in a rather special sense, it seems that this account may exclude both animals and small children from having emotions, as opposed to some proto-version of them (judgmentalists vary as to which horn of this dichotomy they choose). More generally, there are worries about discrepancies between emotions (as ordinarily conceived by non-judgmentalists) and judgments, even if the relevant class of judgments is narrowed down to judgments on issues of value and importance. For the phenomenology of emotion seems rather different from that of judgments, notably in that emotions, or at least occurrent emotions, are typically *felt* (though emotional dispositions will usually not be). Also, emotions typically have a motivational force that “cold” evaluative judgments may lack, and it seems someone can have judgments on issues of value and importance that do not match what is implied by their emotions. Nussbaum’s response to this latter point is that people can demonstrably hold contradictory beliefs: she also suggests that agents in this position may be suffering a degree of inner division *qua* agents, and that the split between their emotion-judgment and their non-emotional judgment reflects this. But this still seems to leave unanswered the question of what the difference between the emotional and the unemotional version of the otherwise same judgment signifies, and whether it does not suggest some more fundamental distinctiveness of emotions.

Secondly, though, there is the worry that if judgmentalism were true, this might amount to achieving rational respectability for emotions at the cost of redundancy: *qua* judgments, it seems emotions are either contributing nothing “cold” evaluative judgments cannot match, or there must still be something important about the distinctively emotional aspect of emotions. And the worry, then, is that this “something else” may turn out to be a repository for much the same features of

emotions that gave them their bad reputation in the first place: in Simon Blackburn's (1998) phrase, the discovery of a Dionysian frog at the bottom of the Apollinian well. These difficulties are also, *mutatis mutandis*, inherited by other kinds of cognitive theory. Weakening the sense of "cognitive" at stake, one may also weaken the claim for intentionality and reason-responsiveness. Importing other elements than cognitions - for instance, desire and feeling - raises again all the traditional worries about the non-rationality or irrationality of these additional elements, and their potential distorting effects on the cognitive elements in emotion.

Emotional rationality through non-rationality, irrationality or trans-rationality

A second, contrasting approach is to defend the rationality of emotions partly or largely in terms of precisely the kinds of features - "quick-and-dirty" firing, lack of responsiveness to reason or will, lacking overt representational content, at least of a consciously accessible kind - that traditional rejections of emotions as irrational or non-rational emphasized, and which serve to set emotions apart from the more "respectable" mental kinds. While not all those who take this line hold emotions to be just "brute feelings", they do often tend to give descriptive accounts of emotion that come close to it. The attribution of rationality to emotion here, moreover, is often achieved by a somewhat revisionist approach to rationality-criteria, where the seeming non-rationality or even irrationality of emotion is in fact what helps emotions be rational or make positive contributions to the rationality of the agent that has them.

The most "brute" version of such accounts is one that basically assimilates notions of rationality to adaptive success: here, the rationality of emotion depends on the usefulness of having them. The usefulness in question may not even be in relation to the goals of the organism, as such, but may be about serving the goals of selfish genes. On this kind of account, emotions need not and typically do not involve much representational content, and even intentionality may strictly speaking be a superfluous consideration. The important thing is that having responses such as fear can aid us, for instance in producing predator-escaping behaviours such as flight, for which complexity of mental representation, reflexive awareness of one's own mental states, or conscious conceptual

processing and deliberation may be disadvantages rather than advantage (by the time you've thought it all out, you may be someone's dinner). Conversely, of course, emotions may be maladaptive (but in explicable ways) insofar as the agent/organism's current environment differs signally from the original context of selection: caveman instincts may serve us less well in the office than in the wild.⁷

This approach also comes in less brute versions, however, and a rather broad variety of them. It is also worth noting that re-emphasizing the importance of the body in emotion need not imply a view of emotional reactions as being hardwired and unchangeable by culture or choice. For instance, Jesse Prinz (2004, 2007) advocates a revival of the James-Lange theory of emotions as perceptions of bodily changes, where those bodily changes in their turn alert us to matters of importance to us, but he also criticizes evolutionary psychology-based approaches to emotions both on a descriptive level and as a basis for evaluation and norms about emotions. Full-on "biologism", including nativism, is also rejected by Antonio Damasio (1994), who argues that emotions, acting as "somatic markers" of what is of importance to us, help promote rational agency, primarily, it seems, by setting goals and value-schemes without which instrumental, strategic deliberation about what to do would lack direction. (Some of the neurologically damaged patients Damasio discusses are effectively incapable of decision-making because they lack a sense of when and how to cut deliberation off, for instance when asked when they would like their next doctor's appointments.)

The economist Robert Frank (1988), meanwhile, argues that emotions may solve the commitment problem, helping us to stay true to strategies whose rationality-warrant is long-term and at least initially higher cost rather than short-term and straightforwardly maximizing. For instance, if one's options are to get some benefit from an unfair deal, or nothing at all, it seems irrational to turn the deal down when the situation is considered in isolation. But resistance to such deals can help promote the negotiation and seeking out of, fairer terms and better deals, or at least, not get one stuck in an exploitative arrangement. In such cases, Frank argues, other agents' expectation that the agent will act in an apparently

⁷ For fuller discussion of such perspectives, see e.g. Barkow, Cosmides and Tooby's collection (1992): for critiques see also, apart from the works by Prinz cited just below, Griffiths (1997), Elster (1999) and De Sousa (2006).

irrational manner in the short term (“he won’t take the deal unless you offer him better terms than that”) can also help promote the agent’s longer-term rational goals. In these accounts, emotions aid rationality without necessarily being in themselves assessable for evaluative rationality – at least of a cognitive kind. Their claims to strategic rationality are through rather complex and less than straightforward routes, and often turn on their being, at least apparently or in part, irrational or non-rational – unresponsive to or even in direct contravention of reason.

Emotions may also track factors relevant to our values and flourishing that escape our conscious intellects, which may themselves be significantly warped by cultural and other distorting factors, or just badly used or developed. So, Nomy Arpaly (2000) argues, it may be possible to act rationally against (what appears to be) one’s own best judgement. For instance, the unhappiness someone feels in what seems on the surface to be a perfect career for them may be tracking points that are in fact good reasons to change careers, but not necessarily in a way that makes these factors accessible to conscious examination in thought.

Note that this need not involve attributing extensive but subconscious representational content to emotion: the representational content of emotion here could be fairly brute, although Arpaly’s account is at considerable distance from a revival of a full-on “boo-hooray” theory. It does, however, trade on the idea that rationality can be assessed in at least two distinct ways. The first Arpaly describes as a “manual” approach: here, the point of a theory of rationality is to tell the agent what the rational thing to do is, and assess her own rationality, based on the evidence available to her. But if what it is most rational for her to do is in fact tracked by, for instance, emotional reactions whose signal to noise ratio is strikingly poor, it may only be possible to assess her rationality from “outside” – from the point of view of an observer who has a fuller and clearer picture of her situation than she herself has access to at the time. (That one such observer might be her own later self may be cold comfort both at the time in question and later.)

Moreover, there may be no fully reliable way to tell inexplicable emotions that track real points from ones that do not, and the lack of transparency of even those recalcitrant emotions that *do* guide us right may make their guidance too unspecific to be of great help. Conversely, however, the implication is that we would not necessarily be better off if we did not have recalcitrant emotions, since our “higher” reasoning capaci-

ties can get things wrong, too – and if our emotions were more responsive to them, they might be in error along with reason.

The issues raised by Arpaly also point to ways in which the intentionality attributed to emotions may be significantly different, at least in this kind of account, than that of judgments. This, however, tends to bring in many of the kinds of worries that helped motivate judgmentalism. First, where an emotion is opaque, an immediate worry is that we may misidentify what object it is about – if it has a real object at all, and is not a mood or a “mere” biochemical state that has mistakenly come to be linked in the agent’s mind with a particular object.

While such a scenario might initially seem unlikely outside of for instance the Schachter and Singer experiment, consider the ways in which factors like fatigue, hunger and thirst (let alone caffeine, alcohol, and any still more psychotropic substances) can affect mood, and the ways in which those moods can, in the next instance, latch on to particular objects, without the subject clearly and consciously taking this into account. So lack of sleep can lead to irritability, and irritability can focus itself into anger with a specific person – say, a significant other that happens to rub one the wrong way on one’s morning-grumpy path. We do not always, in such cases, consider properly that the other person, *qua* emotion-object, is not so much the real cause of the emotion as a target the pre-existing affective state gets taken out on. Conversely, emotions that are in fact about specific persons and incidents may change the persons overall mood, or even the general tone of their affective life long-term, spreading out onto the world well beyond the original object (Armon-Jones 1991, Baier 1990).

And this points also to ways in which we can be unclear – or misguided – about the causes of our emotions, that is, the descriptions under which things qualify as their objects: why, for instance, in this case, one got angry with the significant other that morning. But if we cannot specify what or who an emotion is really about (or even be entirely sure it is an emotion and not a mood or a “merely” biochemical issue), it will be hard indeed to assess whether the emotion is warranted. This may or may not affect its strategic rationality – for instance, getting randomly angry or euphoric may, depending on context, have positive or negative outcomes in terms of one’s life goals and relationships with others – but

it will certainly make its cognitive rationality hard to assess, and will render its strategic rationality, if preserved, rather brute.⁸

That judgementalism promises this point is at least in principle one that can be settled, and that this contrasting approach effectively goes far towards undermining that promissory note, may explain a great deal of the appeal of judgmentalism, and more generally of cognitive theories. It may also explain the resistance to neo-Humean accounts of emotion, as well as revivals of the James-Lange theory of emotion as involving first physiological change, then awareness of this and *then* awareness of the event or object triggering the physiological arousal.⁹ For these accounts, by contrast with judgementalism, makes the opacity of some emotions not so much anomalous as a thing to be expected, and which may not even be in principle removable, since the intentionality of emotions may here be brute rather than cognitively fine-grained.

Moreover, an approach to rationality that appears relatively blithe about leaving the mental entities involved as black boxes, and willing to measure strategic rationality by what often looks suspiciously like luck in outcomes, may meet resistance on that front in itself. In the case of judgmentalism, then, there is the arguable problem of over-intellectualizing both emotions and rationality; this second approach might come across as going too far to the other extreme.

Emotional rationality as *sui generis*

Thirdly, then, some defend a view of emotions as rational by claiming emotions exhibit a more *sui generis* kind of rationality, often one based on a dramatic or narrative “inner logic”.¹⁰ Here, emotions “make sense” in terms of the expectations carried by “paradigm scenarios” (de Sousa) or larger, ongoing narratives, where people’s sense of their own identity and role, the role of other “players”, and the ways in which the

⁸ For concerns about Arpaly’s approach, partly motivated by these kinds of considerations, see also Jones (2003).

⁹ See James (1884), Prinz (2004, 2007); a discussion of James’ theory is also a crucial starting-point for Schachter and Singer’s (1962) experiment.

¹⁰ Versions of such views can be found in e.g. Rorty (1980), De Sousa (1980, 1987), Baier (1990) and Goldie (2000, 2003) – though it is worth keeping in mind that at least De Sousa and Rorty also hold views that in part tend toward the kind of account discussed in the previous section, albeit with an emphasis on the intentional richness of at least human emotion.

situations are construed, all play a part. So, too, will broader cultural narratives acquired in the process of socialization: de Sousa calls these “ideologies of emotion”. The meaning attributed to anger, for instance, may vary depending on culture, personal values, and one’s sense of who one is, in oneself and in relation to others. When it “makes sense” to get angry will vary accordingly, both in terms of what considerations are thought to warrant anger (which will roughly correspond to cognitive rationality criteria), and in terms of the likely outcomes of feeling and showing anger (which will roughly correspond to strategic rationality criteria).

Emotions, on this kind of account, have rather thicker content than “brute feeling”, but are typically conceived more on the model of feelings – albeit ones of and responsive to considerable representational complexity – than of judgements or beliefs, even evaluative ones. De Sousa’s account, while often classed as “cognitivist”, tends to view the representational content of emotions as more like perception, though not exactly analogous to it: specifically, emotions are (like) a form of axiological perception. While certainly a judgementalist version of this approach seems logically possible, it would have to be one that gave a clear sense of what is different about emotional judgments compared to other kinds: the insistence on a *sui generis* logic suggests simply importing standard rationality-criteria without also examining potentially important distinctive features of the phenomena they will be applied to will be problematic. We do, after all, already tend to apply rather different criteria to desires than to beliefs, and anyone wanting to disregard differences between the two kinds, in themselves or in the rationality-criteria applied to them, would have to make a case for doing so.

A notable feature of this type of account is that it tends to emphasize the distinctions between emotion-types more than the previous two. De Sousa (1987) for instance, argues that while beliefs have a common “formal object” at which they aim and to which they are supposed to be responsive, namely truth, and desires, while for different kinds of goods (drink for thirst, food for hunger, and so on), may at least in the abstract sense of “good” share a formal object, there is no such obvious shared formal object for all the different phenomena we call emotions. Rather, specific emotion-types – fear, anger, joy, surprise, and so on – have their

own characteristic formal objects,¹¹ and by extension, their own characteristic criteria for when particular tokens of them “make sense”, minimally or evaluatively. These may in turn be subject to constraint from other kinds of rationality-criteria: it seems desirable, for instance, that one have some sense of how emotions should align with beliefs and desires (so that we can tell, for instance, what beliefs would normally count as defeating the warrant for a particular emotion), and how to resolve cases of emotional ambivalence, where different emotion-types bear on an object in contradictory ways. The “inner logic” of emotion, though, appears on this account to be in a sense primarily a local phenomenon.

The special challenges faced by this type of account relate primarily to what might otherwise seem their big advantage, namely the idea of *sui generis* emotional rationality. For there is now the question, first, of where we get such standards from. If it is from observing, to paraphrase Aristotle’s approach in *The Art of Rhetoric* (itself arguably an early instance of this kind of approach), who typically feels what, towards whom, and why, and with what consequences for their later attitudes and actions, then the Euthyphro-problem of emotions threatens to rear its head again. Is an object offensive because it tends to offend people, or are they offended by it because of its offensiveness? What are the criteria for offensiveness, other than tending to get a specific response? Secondly, there is the question of what shapes these reactions in the first place, let alone the second or third: nature, so that the narrative/dramatic logic just tracks, at least in the first instance, an “animal” logic, albeit that it may then, for animals that can tell stories and create dramas, take on more intricate nuances? Choice? Culture? Or, as seems likely, varying combinations of any or all of these? And then there are the normative issues: if we want to determine criteria for when it is intelligible, let alone justified, for someone to be offended, or happy, or sad, or afraid, how do we choose these? Cultural standards, let alone individual ones, may be questionable in all sorts of ways, and appeals to nature raise their own (meta-)problems (see also De Sousa 2006). In setting out both their descriptive assumptions about emotion, then, and their normative

¹¹ Similar arguments are raised by Rorty (1980) and Baier (1990): for Rorty, this is also a reason to question whether the different kinds we call emotions form a natural class at all.

assumptions about the rationality of emotions, this type of account will have to navigate a host of nature-nurture issues, as well as meta-evaluative issues of naturalism, (social) constructivism and response-dependence. This may, of course, count as an acceptable – and even necessary – cost of avoiding reductivism about emotions themselves – of resisting aligning them too closely to either judgments or “bruter” phenomena such as reflexes or sensations. But it also points up, again, the continuing elusiveness to definitive accounts of both emotions and rationality.

What (not) to do about this diversity of views

In the face of such diversity, certain responses are tempting, but should be resisted. It would be easy, for instance, to assume that the proponents of these views are just talking past each other. But this ignores the fact that there are substantive issues at stake between them, both as regards rationality and as regards the nature of emotions, let alone any connections between the two.

For instance, a “judgementalist” account of emotions such as Nussbaum’s or Solomon’s is not easily reconcilable with a view of emotions that sees them as largely innate “affect program” responses built into our biology and more susceptible to explanation in terms of evolutionary psychology than appeals to ancient philosophy or existentialism. Nor are the kinds of rationality emotions appear to be (favourably) assessable by the same for the two accounts. It is also worth noting, again, that both Nussbaum and Solomon hold fairly strongly voluntaristic accounts of emotions, where an account that sees them as in-built responses, biologically programmed to be responsive to certain types of stimuli rather than others, and to produce quick-and-dirty responses and output dispositions of fairly stereotyped kinds, will tend to indicate less voluntary control over emotions, especially on a case-by-case basis. Broader, long-term dispositions may still be viewed as open to a degree of modification by habituation, though rarely by means of the kinds of rational persuasion strategies, such as explicit argumentation, that beliefs and judgements are (supposed to be) responsive to. But to hold that emotions are open to culture- and nurture-effects, or even to a degree of choice in shaping them, need not mean that they are judgments: as we have seen, some authors who hold such views, like Prinz, offer a descriptive account of

emotion that is closer to the James-Lange theory that at least some judgementalists (notably Sartre and Solomon) were reacting against.

This need not mean, on the other hand, that the approaches are necessarily mutually exclusive: elements of all three can often be found, albeit in varying ratios and strengths, in any one author without this necessarily implying that the author in question is simply muddled, conceptually speaking (Elster's 1999) views, for instance, seem to take in elements of all three). It seems possible, and plausible, after all, that emotions can take forms that range across a spectrum from "brute feelings" to highly conceptually complex, representationally sophisticated responses, and their degree of responsiveness to voluntary control and argumentation might vary accordingly. In other words, the views of emotion implicit or explicit in the three accounts distinguished above may be viewed as due to focusing on different kinds of cases as paradigmatic of emotions. Whether the views are also exclusive will depend on whether their proponents take their own accounts to be exhaustive, so that whatever does not fit them will not count as an emotion. Some authors do take such views – but it is not clear that such a view is inherently compelling, and there is considerable reason to resist such "nothing-but-ism", under all three headings.

A third kind of temptation to be resisted, then, is towards the view that the differences must mean only one of the approaches can be right – if any of them can be, that is. (Since the views above combine particular views about emotion and particular views about rationality that could be paired in rather different ways, we can relatively easily imagine other combinations: and only the third, narrative-dramatic option has really close ties between its assumptions about emotions and its assumptions about their rationality.) This assumes, too easily, that rationality is just one thing, when in fact the (vast) literature on rationality suggests a number of diverse criteria, not easily reducible to one another, and themselves capable of being in tension with one another, in practice or theory. Consider, for instance, the instrumental rationality that certain kinds of false beliefs may have, for instance when people sustain social confidence through an exaggerated sense of their own abilities and likeability: from the point of view of rationality-criteria aimed at truth-aptness, such false beliefs are irrational, but this need not make them irrational in terms of their strategic aptitude for achieving particular goals. The rationality of pursuing those goals in themselves, moreover,

may be a matter for yet another kind of rationality-criterion, one which sceptics about substantive rationality might reject, for instance on Humean grounds. (While having unachievable-in-principle, or incompatible, or intransitive, preferences might still be assessable for (ir)rationality on such accounts, goals *qua* goals would not typically be.) The rationality of constrained maximizers, moreover, will be different from the rationality of straightforward maximizers, and the relative rationality of constrained over straightforward maximization may itself be a thorny issue, both in theory and practice. Given this variety of criteria, then, it is easy to see that emotions (whatever they in their turn turn out to be) will score variably depending on what type of rationality they are assessed for.

The “they can’t all be right” view may also assume too easily that emotions as a class cannot (or cannot coherently, if the category is itself to hold up) display significant internal variation. For instance, the temptation to subdivide the emotion-category e.g. into innate dispositions (such as Griffiths’ (1997) use of Paul Ekman’s concept of “affect-programs”) on the one hand and cultural artefacts on the other, should be viewed with some suspicion, given the real difficulties of reliably distinguishing, let alone separating, the animal and cultural aspects of our natures, which seem likely rather to be importantly continuous with and mutually shaping of each other. Moreover, even those emotions most plausibly innately preprogrammed to appear in us – fear, anger, joy, and so on – are open to significant acculturation as to what will elicit them, how and whether they will be expressed, and what the consequences of expressing them will be. Nor do we currently possess a particularly advanced understanding of the connection between putative “basic” emotions, biologically based or otherwise, and other aspects of our emotional lives, so some caution in drawing broad conclusions is in order. (This also in light of the often rather unedifying tendency of debates on the issue to turn into nativist/social constructivist trench battles.)

(Issues of the rationality of what Griffiths calls “culturally sustained pretenses” – effectively, presenting oneself as being in the grip of emotion that one does not feel, or of an emotion-type whose existence is in fact an invention, as some have argued is the case for romantic love – are another matter, and complex in their own right: the strategic usefulness of such moves seems likely to depend crucially on surrounding ideolo-

gies of emotion, which will determine, for instance, the extent to which others cut one slack for being in the throes of a feeling. Around existentialists of a Sartrean persuasion, the move seems likely to backfire.)

Given the range and complexity of at least human emotion, it seems finally unsurprising that our emotions could meet – and come apart from – a diverse range of rationality-criteria, in varying ways according to the particular case. Nor, given our evolutionary history, is it all that surprising that our emotions may in significant respects still be relatively brute, and brute in such rationality as they are capable of. But we are also thinking animals, self-redefining animals, cultural animals, and this too makes a difference to our emotional lives, for all the continuity there may be between them and the emotional lives of animals that lack these features. But a sufficiently fine-grained account of both rationality and emotion can – and should – accommodate these points. All arguments on these issues of course, are provisional in the light of potential further developments in our understanding of both emotions and rationality.

Some ways of going forward do suggest themselves, but on fairly broad strokes lines, and of kinds unlikely to be covered by any one discipline, let alone any one author. First, more empirical and conceptual work is needed on emotions as a group, and for that matter, whether they form one in the first place, in order to assess those parts of the competing accounts under discussion here that depend on what we understand emotions to be in themselves. Secondly, we need to acknowledge, and keep clearly in mind, the differences and tensions between the diverse ways in which we understand rationality – if we can't agree on one criterion (or set of criteria) to assess the rationality of emotion by, we should at least stay clear about there being a diverse set of criteria, and that much of the source of debate stems from lack of consensus about rationality, not just about emotion.

*Department of Philosophy
University of Ottawa,
Ottawa, Canada
srietti@uottawa.ca*

References

ARISTOTLE (1991): *The Art of Rhetoric*. Trans. H.C. Lawson-Tancred. London: Penguin Books.

- ARMON-JONES, C. (1991): *Varieties of affect*. Toronto: University of Toronto Press.
- ARPALY, N. (2000): On Acting Rationally against One's Best Judgement. *Ethics* 110, 488 – 513.
- BAIER, A. (1990): What emotions are about. *Philosophical Perspectives* 4, 1 – 29.
- BARKOW, J.H.L. – COSMIDES, L. – TOOBY, J. (1992): *The adapted mind: evolutionary psychology and the generation of culture*. Oxford: Oxford University Press.
- BLACKBURN, S. (1998): *Ruling Passions: A theory of practical reason*. Oxford: Clarendon Press.
- CHARLAND, L. (1997): Reconciling Cognitive and Perceptual Theories of Emotion: A Representational Proposal. *Philosophy of Science* 64, 555 – 579.
- DAMASIO, A. (1994): *Descartes' Error: Emotion, reason and the human brain*. Grosset/ Putnam, New York.
- DAVIDSON, D. (1963): Actions, reasons and causes. *The Journal of Philosophy* 60, 685 – 700.
- DEIGH, J. (1994): Cognitivism in the theory of emotions. *Ethics* 104, 824 – 54.
- DE SOUSA, R. (2006): Restoring emotions' bad rep: the moral randomness of norms. *European Journal of Analytic Philosophy* 2, 27 – 45.
- DE SOUSA, R. (1987): *The rationality of emotion*. London/Cambridge MA: MIT Press.
- DE SOUSA, R. (1980): The Rationality of Emotions. In RORTY, A. O. (ed.): *Explaining Emotions*. Berkeley: University of California Press, 127 – 151.
- EKMAN, P. – DAVIDSON, R. (eds.) (1994): *The Nature of Emotion*. Oxford: Oxford University Press.
- ELSTER, J. (1999): *Alchemies of the mind: Rationality and the emotions*. Cambridge: Cambridge University Press.
- FRANK, R. H. (1988): *Passions within reason: The strategic role of the emotions*. Ithaca, NY: Cornell University Press.
- GOLDIE, P. (2003): Narrative and perspective; values and appropriate emotions. In HATZIMOYSIS (ed.), 201 – 220.
- GOLDIE, P. (2000): *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.
- GRIFFITHS, P. E. (1997): *What emotions really are*. Chicago: University of Chicago Press.
- HATZIMOYSIS, A. (ed.) (2003): *Philosophy and the emotions*. Cambridge: Cambridge University Press.
- JAMES, W. (1884): What is an emotion? *Mind* 9, 199 – 205.
- JONES, K. (2003): Emotion, weakness of will and the normative conception of agency. In HATZIMOYSIS (ed.), 181 – 200.
- NUSSBAUM, M. C. (2001): *Upheavals of Thought*. Cambridge: Cambridge University Press.
- NUSSBAUM, M. C. (1994): *The therapy of desire: theory and practice in Hellenistic ethics*. Princeton, N.J., Chichester: Princeton University Press.
- OAKLEY, J. (1992): *Morality and the emotions*. London: Routledge.
- PRINZ, J. (2007): *The emotional construction of morals*. Oxford: Oxford University Press.

- PRINZ, J. (2004): *Gut reactions: a perceptual theory of emotion*. Oxford: Oxford University Press.
- RORTY, A. O. (1980): Introduction. In RORTY A. O. (ed.): *Explaining Emotions*. Berkeley: University of California Press, 1 – 8.
- RORTY, A. O. (1978): Explaining emotions. *The Journal of Philosophy* 75, 139 – 161.
- SARTRE, J. P. (1948): *Emotions: The Outline of a Theory*. New York, Citadel Press.
- SCHACHTER, S. – SINGER, J. (1962): Cognitive, social and physiological determinants of emotional state. *Psychological Review* 69, 379 – 399.
- SENECA, L. A. (1995): *Moral and political essays*. Edited and translated by John M. Cooper and J.F. Procopé. Cambridge: Cambridge University Press.
- SHAFFER, J. (1983): An Assessment of Emotion. *American Philosophical Quarterly* 20, 161 – 73.
- SHERMAN, N. (1997): *Making a necessity of virtue: Aristotle and Kant on virtue*. Cambridge: Cambridge University Press.
- SMITH, M. (1995): Internal reasons. *Philosophy and Phenomenological Research* 55, 109 – 131.
- SMITH, M. (1987): The Humean theory of motivation. *Mind (New Series)* 96, 36 – 61.
- SOLOMON, R. C. (2006): *True to our feelings: what our emotions are really telling us*. Oxford: Oxford University Press.
- SOLOMON, R. C. (1993): *The passions: emotions and the meaning of life*. Indianapolis: Hackett Publishing Company (originally published 1973).
- SORABJI, R. (2000): *Emotion and peace of mind: from Stoic agitation to Christian temptation. The Gifford lectures*. Oxford: Oxford University Press.
- WILLIAMS, B. (1973): Deciding to believe. In his *Problems of the Self*. Cambridge: Cambridge University Press.
- WOLLHEIM, R. (1999): *On the emotions. (The Ernst Cassirer Lectures, 1991)*. New Haven, CT: Yale University Press.