ARE EMOTIONS NECESSARY AND SUFFICIENT FOR MAKING MORAL JUDGMENTS?1

AS EMOÇÕES SÃO NECESSÁRIAS E SUFICIENTES PARA FAZER JUÍZOS MORAIS?

MARCO AURÉLIO SOUSA ALVES
(University of Texas at Austin / EUA)

RESUMO

Jesse Prinz (2006, 2007) afirmou que emoções são necessárias e suficientes para julgamentos morais. Primeiramente, explicito qual é de fato a sua tese. A teoria que ele chama de emocionismo é, então, criticamente avaliada. Prinz arregimenta várias descobertas empíricas para defender uma série de teses cada vez mais fortes acerca de como as emoções são essenciais para os julgamentos morais. Argumento aqui que o suporte empírico sobre o qual se assentam seus argumentos é não apenas insuficiente, mas até mesmo sugere o contrário, se devidamente interpretado. Minha crítica é, então, estendida à sua teoria sentimentalista, que se ocupa de como as emoções são integradas em julgamentos morais. O problema central é que a perspectiva de Prinz não consegue apreender o aspecto racional do julgamento moral. Tal incapacidade é explicitada, e por fim reivindico que alguma forma de neossentimentalismo é a via mais promissora.

Palavras-chave: Jesse Prinz; emoção; emocionismo; teoria sentimentalista; neossentimentalismo

ABSTRACT

Jesse Prinz (2006, 2007) claimed that emotions are necessary and sufficient for moral judgments. First of all, I clarify what this claim amounts to. The view that he labels emotionism will then be critically assessed. Prinz marshals empirical findings to defend a series of increasingly strong theses about how emotions are essential for moral judgments. I argue that the empirical support upon which his arguments are based is not only insufficient, but it even suggests otherwise, if properly interpreted. My criticism is then extended to his sentimentalist theory, that accounts for how emotions are integrated into moral judgments. The central problem is that Prinz’s view fails to capture the rational aspect of moral evaluation. I make this failure explicit and defend that some version or other of neosentimentalism is a more promising route.

Keywords: Jesse Prinz; emotion; emotionism; sentimentalist theory; neosentimentalism

1. On necessary and sufficient conditions

Classical debates in metaethics have been increasingly influenced by empirical sciences. Jesse Prinz (2006, 2007) believes that these contributions suggest a strong connection between emotional reactions and morality. But what does this strong connection amount to? Prinz...
elaborates on the following terms: emotions are both necessary and sufficient for moral judgments.

However, the claim above, as it stands, is quite unclear: is he proposing some sort of conceptual analysis of moral terms? Prinz (2006, p. 30) is perfectly aware that conceptual analyses are thorny and hopeless. He is not after something like the analysis of 'bachelor' as an unmarried adult male. The claim, instead, is that having emotions is necessary and sufficient for possessing ordinary moral concepts (concepts like right, wrong, bad, etc.). This thesis is labeled epistemic emotionism (Prinz 2007, p. 21). Concerning the nature of morality, there is also the metaphysical claim that emotions are essential for determining moral properties. This amounts to what Prinz (2007) calls metaphysical emotionism. If both theses are combined, the result is called strong emotionism, and that is the view that Prinz advocates.

The twofold aspect of strong emotionism corresponds to the dual aspect of meaning: sense and reference. The metaphysical thesis regards moral properties, or the referents of moral terms. The epistemic thesis concerns the sense, understood roughly as moral concepts, or “psychological tools by which we come to understand morality” (Prinz 2007, p. 17). An ordinary moral concept, according to Prinz (2006, p. 30), is “what goes on in our heads when we use moral terms like 'good' and 'bad' or 'right' and 'wrong'”. Hence, the necessary and sufficient conditions for having an ordinary concept can be regarded as minimal conditions for linguistic competence in using moral terms.2

The abnormal use of a term, after a certain point, indicates the possession of a different (or idiosyncratic) concept, and not merely the same concept being used in a non-standard way. For instance: the reference of 'tiger' is a certain natural kind. But its sense, defined as its ordinary concept, may involve the association with a stereotype such as 'large, orange, striped catlike creature'.3 Obviously, the stereotype gives no necessary or sufficient condition for being a tiger: tigers may have other colors, they may not be striped, and so on. These conditions individuate concepts, or mental files, not referents. So understood, strong emotionism is a semantic theory about both the reference and the (ordinary) sense of moral terms.

2. The empirical evidence

Prinz marshals empirical findings to defend a series of increasingly strong theses about how emotions are essential for moral judgments. In this section, I initially consider some
experiments that allegedly support the claim that emotions are sufficient for moral judgments. Following that, I consider a case that supposedly suggest that emotions are also necessary for moral judgments. The mere fact that judgments are influenced by emotions - or that emotions distort judgments - is trivial. What matters here is the move from this platitude to the sufficiency or the necessity claim.

Let’s first consider two experiments showing that emotions influence moral judgments. Wheatley & Haidt (2005) hypnotized subjects to feel sudden disgust when they hear some trigger words (‘take’ and ‘often’). The hypnotized subjects tended to evaluate stories containing these words much more severely. Even when the trigger words were used in morally neutral stories, they tended to condemn the characters.

In another experiment, Schnall et al. (2008) asked subjects to evaluate stories while sitting at a filthy desk (with an old greasy pizza carton, a chewed up pencil, used tissues, and a dirty beverage cup). Those at the filthy desk were much more severe on their evaluations than those sitting at tidy desks. They also found out that an unpleasant odor – a fart-spray in the surroundings, for instance – had a similar impact on people’s judgments about ethical topics.

Empirical support for the sufficiency claim also comes from the so-called dumbfounding phenomenon. Murphy et al. (2000) studied moral attitudes towards consensual incest and observed that even though 80% of the subjects judged the behavior as morally wrong, they had great difficulty explaining why. Every justification provided by the subjects was appropriately replied, and even after conceding to the counterexamples, only 17% of the subjects withdrew from their initial judgments. Prinz interprets this case as showing that even in the absence of rational justification, one can have a moral attitude. The dumbfounding phenomenon is supposed to reveal that rational arguments play no essential role in arriving at or sustaining moral judgments. According to Prinz (2007, p. 32), “basic values are implemented in our psychology in a way that puts them outside certain practices of justification”.

The claim that emotions are necessary for moral judgments is supposedly supported by studies on psychopathy. Psychopaths tend to engage in chronic antisocial behavior, and they do so without emotional cost. Their lack of emotional response co-occurs with a deficit in moral competence, which suggests that the moral deficit results from the emotional one. Blair (1995) observed that criminals diagnosed as psychopaths, in opposition to ordinary criminals, respond to moral and conventional wrongs in the same way, failing thus to draw the moral/conventional
distinction. Based on this study, Prinz claims that psychopaths do not grasp moral concepts, since they cannot tell the difference between being morally wrong and being conventionally wrong. Prinz (2006, p. 32, and 2007, p. 42-7) then concludes that psychopaths only seem to understand morality. They use moral terms in deviant ways, and therefore they do not possess ordinary moral concepts.

3. On the conditions on moral concept possession

Prinz (2006, p. 31) claims that the empirical findings above “suggest that we can form the belief that something is morally wrong by simply having a negative emotion directed towards it. In this sense, emotions are sufficient for moral appraisal”. But which sense is that?

The empirical data, I claim, only suggests that emotional reactions are causally sufficient for putative moral judgments. Causal sufficiency, however, is not the kind of condition that we were after. As a semantic theory, emotionism is in business of determining conditions on ordinary moral concept possession and, by extension, for making ordinary moral judgments. Whatever is caused by filthy desks, fart-sprays, or hypnotic suggestion, it cannot be assumed to be a genuine moral judgment, on pain of begging the question. The relevant behavior to be explained involves more than mere locutions that sound like moral judgments. Prinz (2006, p. 32) himself explores the distinction between genuine and pseudo moral judgments when he talks about psychopaths: “they give lip-service to understanding morality, but there is good reason to think that they do not have moral concepts”. By the same token, or so I claim, one may wonder if subjects under the influence of hypnosis are in fact making full-blown moral judgments.

Consider the following case: John is subjected to post-hypnotic suggestion and is later informed about his condition during the experiment. Even after the debriefing, he refuses to withdraw his judgment. John is perfectly aware that his judgment was produced by a hypnotically induced feeling of disgust towards triggering words. If all that matters to making a moral judgment is the induced feeling, then it is unclear why he should be expected to take back his judgment. He felt the “necessary and sufficient” emotion, and that is what his judgment is all about.

Jones (2006, p. 48) considers the case above and sees two possible interpretations: either all there is to making a moral judgment is having an emotional response, or John is not making a
moral judgment, despite appearances. If his refusal was not based on any independent reason or assumption, but in the mere fact that he felt the way he did (no matter if hypnotically induced), then there is some strong evidence that John is not making an ordinary moral judgment. What makes his refusal so surprising is the fact that he takes his emotional state *per se* as justifying the judgment.

Now consider another subject, Mary, who holds a moral judgment on the basis of emotional reactions. She may be a perfectly competent moralizer, insofar as she recognizes that her judgment is answerable to reasons. Mary may assume that she is justified, that her emotion is responsive to some appropriate reason, even though she can’t articulate it. Mary may be stubborn, or too confident in her intuitions, but she is not conceptually confused. John, however, takes hypnotically induced responses as justifying his judgments. This attitude, I suggest, indicates a lack of competence. Sauer (2012, p. 12) proposes the following criterion: a proper standing in the space of reasons is required for moral concept possession. Given that constraint, John’s judgment does not count as genuinely moral because he is happy to hold a judgment that clearly does not respond to morally relevant features of the situation (assumed that he is aware of the method used to form the judgment).

Theories about concept possession are based on minimally competent use of terms. This kind of evidence indicates which stereotypes are associated with certain terms. John’s refusal to withdraw gives us grounds to doubt his moral competence. A normal moral agent should be responsive to obviously relevant considerations. To take any emotional fact (like the feeling of disgust) as warranting a moral judgment indicates not merely a mistake, but a lack of grasp on morality. John uses moral terms in abnormal or idiosyncratic ways. He can’t share conversation about morality with normal people: they will talk past each other. According to Sauer (2012, p. 19), incompetent moral concept users are just trying to make moral judgments, but they fail to do so, because “they cannot be said to engage in the practice of moral judgment in the right way. Constantly failing to live up to these standards renders a subject morally incompetent, and ultimately deprives it of its status as a morally judging subject altogether”.

Minimal requirements for moral competence involve, at least in part, a rational stance. That is why ordinary concepts of taste are not like ordinary moral concepts. In order to possess a moral concept, one must have a minimal understanding of the justificatory commitments that come with it. Irrelevant situational features do not bear on the moral wrongness of an action: an
emotional reaction induced by fart-spray is morally irrelevant because its cause is not justification-conveying. What is needed is a sufficient condition for moral judgment that does not threat its standing in the space of reasons. Emotionism lacks a way of ruling out cases of morally irrelevant causes constituting the very essence of moral concepts.

Sauer (2012, p. 12) proposes that the essential components of sufficiency of moral judgments can be spelled out as counterfactual idealized conditions: an emotional reaction is sufficient if “it is causally sufficient in a way that can be reflectively endorsed under conditions of full information and rationality”. Whether this analysis is right or not goes far beyond my purposes here, but it illustrates very well what kind of condition emotionism lacks.

4. Reinterpreting the empirical evidence

Studies on psychopathy, as seen above, are used to ground the necessity claim. Having the appropriate emotional reaction may indeed be necessary for making the moral/conventional distinction. As a mere causal necessity, this seems true: that is how our psychology works. I suggest, however, that the conceptual deficit of psychopaths may be due to a justificatory deficit (in opposition to a mere lack of emotional disposition).

Blair (1995, p. 13) observed that psychopaths’ justifications for why a certain norm is wrong are “less likely to make references to the pain or discomfort of victims”. Jones (2006, p. 51) stressed the psychopaths’ failure to grasp the normal kinds of reasons: considerations like 'killing with a clean gunshot wound', for instance, are taken by psychopaths as morally excusing. Kennett (2006, p. 76) enumerates a list of rational shortcomings of psychopaths: “weak capacity to stand back from and evaluate their desires, to estimate the consequences of their actions, to eschew immediate rewards in favor of longer term goals, to resolve conflicts among their desires, to find constitutive solutions”.

It is likely that the psychopaths’ emotional deficit causes their justificatory deficit. But it is not clear to me which deficit shall be used to explain the failure in concept possession (or if both deficits are equally relevant). One way or another, it seems clear that psychopaths fail to grasp essential elements of moral judgment. Moral judgments are made by morally competent subjects, and psychopaths do not pass this test.
Dumbfounding, on the other hand, purportedly supports the sufficiency claim. But once again, a rational constraint seems to be doing the relevant job. The fact that we often make moral judgments without being able to spell out the reasons is uncontroversial. The rational aspect of morality does not have to take the form of an explicit belief accompanying each act of judging. All that is needed is a tacit justificatory assumption. Most of the times, moral judgments are not causally regulated by reflection. That is why rational constraints on concept possession are better understood in counterfactual terms. Typically, moral attitudes are emotionally triggered. Since we are equipped with emotions that reliably track moral features, we are often right in relying on them. Emotions, in this sense, are very important psychological tools: they guide our actions without the need of any conscious deliberative processing. The urgency of real life situations and the importance of having entrenched moral responses explain why such ability is so important.

Given the method by which dumbfounded people arrive at their judgments, it is no surprise that they can’t put their fingers on the reasons justifying their judgments, even after careful scrutiny. But it would be surprising if they did not assume that there must be a reason somewhere. That is why they tend to come up with biased and ad hoc justifications. Normal dumbfounded subjects recognize that moral judgments are answerable to reasons. They are, perhaps, stubborn or too confident to reject their emotionally triggered responses, but they are using moral concepts in a perfectly normal way. They are stubborn (even after having their arguments refuted) because they still hope that some unknown reason is hidden somewhere. Unless the method by which they arrived at the intuition is properly brought into consideration and undermined, they tend to keep stubborn. Given their appropriate stance in the space of reasons, they are perfectly competent moralizers.

Prinz (2007, p. 49) is well aware that empirical evidence can only suggest metaethical theories. He knows that alternative views are viable. He only claims to have the best one on the market. However, or so I claim here, his argument to the best explanation fails. Emotionism is not the best candidate available for accommodating the relevant data. When properly understood - as (at least in part) a theory about conditions on concept possession - emotionism loses its attractions and alternative views fare better.
5. Against Prinz’s sensibility theory

The same difficulty of accommodating the rational dimension of moral evaluations can be found in the kind of sensibility theory advocated by Prinz. Emotionism claims that emotions are essential (necessary and sufficient) to morality. The sensibility theory is a more elaborated metaethical view that spells out the precise way in which emotions are integrated into moral judgments.

Prinz (2007, chap. 3) develops increasingly more elaborated versions of his favored sensibility theory. Roughly speaking, he ultimately claims that to judge that something is wrong (or right) is to have a sentiment of disapprobation (or approbation) towards it. *Sentiment* is a disposition to have a certain emotion; it is seen as the “possession condition on the normal concept RIGHT (WRONG)” (Prinz 2007, p. 87). Properly speaking, moral judgments do not involve emotions directly. Judgments express sentiments, and sentiments refer to the property of causing certain emotions in us. The ordinary concept WRONG is a reliable detector of the property of wrongness; it comprises a sentiment that disposes us to experience emotions in the disapprobation range.8

In what follows, I will not be concerned with the details of this theory. My sole objective is to show that the same difficulty concerning the rational dimension renders his sensibility theory equally implausible. The central challenge for any version of sentimentalism is to preserve the idea that morality is grounded in sentiments, while at the same time making sense of the rational aspect of morality. Most contemporary sentimentalists, the so-called neosentimentalists, defend that something must be added to the simple sentimentalist thesis in order to capture the distinctive normative force of moral terms and judgments.9

Contrary to simple sentimentalism (like the one proposed by Prinz), neosentimentalism takes the rational endorsement of our emotional responses to be constitutive of moral concepts. Neosentimentalists introduce some kind of metacognitive constraint: right (wrong) is not just that towards which we have a sentiment of approbation (disapprobation), but that which we take to warrant such sentiment. To judge that something is wrong (right) is to endorse a sentiment of dis(approbation) towards it. Once again, I won’t be concerned with the subtleties and variations of neosentimentalist theories. The core idea is that emotions and sentiments are not necessary and sufficient for morality: some rational constraint must be added.
Prinz explicitly rejects any metacognitive constraint. The main charge against simple (non-metacognitive) sentimentalism is that it fails to account for error. Prinz replies that the rational aspect of moral evaluation can be fully explained by conative (as opposed to cognitive) states. He claims that the possession of sentiments that are governed by norms is normative enough: our emotional feelings can be criticized, and eventually corrected.

Prinz (2007, p. 112) considers the case of a person raised to think homosexuality wrong, but who now thinks otherwise. Though the former homophobe still has the disposition of having feelings of disapprobation towards homosexual acts, he does not take the act as morally wrong. Prinz grants that there is a sentiment (normal and reliable) eliciting the feeling of disapprobation, and that this sentiment is not morally meritorious. However, this intuition that merit is essentially involved in moral judgments derives solely from the fact that emotional responses can be controlled to some degree. This can be explained by a second-order emotional level. According to Prinz (2007, p. 115), moral emotions are merited by their causes because they are “regulated by meta-emotions”. The former homophobe, for instance, is still a (partial) homophobe so long as he still has negative sentiments towards homosexuality. His meta-sentiments are, however, trying to change the first-order sentiments, and the battle is still on.

Nonetheless, I claim that the distinction between first and higher-order sentiments cannot do the expected job. Is the first-order judgment of the homophobe correct? If it is merely about whatever causes his reliable disposition to have a feeling of disapprobation in the appropriate range, then Prinz is committed to saying that the first-order judgment is correct, though not meritorious, given his meta-sentiment towards it. This strikes me as very implausible: merit and correctness are not together anymore.

Furthermore, the dispositional embellishment does not change the crucial point: dispositions can’t capture the normative aspect of judgments. Prinz’s sentimentalism does not capture ordinary judgments. Not every kind of moral wrongness can be accounted for by appealing to normal responses. The normal isn’t normative. In Prinz’s view, moral properties do not make demands upon us. A disposition explains an ability, but it does not demand anything. As Joyce (2009, p. 516) puts it, “if we cannot make out how a property can be construed as making a practical demand, then what business have we identifying that property with moral rightness or moral wrongness?”. The strictly causal-conative approach of Prinz may explain how
we humans come to experience the world as making demands upon us, but it cannot explain the semantic nature of those demands.

Whether a subject has arrived at her judgment using a reliable method or whether she is prepared to reconsider her judgment by reflective scrutiny are two different things, and both of them are essentially involved in making a genuine moral judgment. Consider, for instance, the distinction between moral and conventional rules. According to Prinz (2006, p. 37), we tend to moralize mere conventions if we learn them through a process of emotional conditioning. If someone is emotionally conditioned to obey certain religious dietary rules, for instance, these rules will be treated moralistically, even if the subject recognizes, in a second-order attitude, that they are merely conventional. The striking consequence is that the first-order moral rule is about whatever causes a strong emotion. The moral character of the first-order attitude, however, cannot be altered by the meta-attitude. Moreover, the intensity of the emotion seems irrelevant to the moral character of something: moral transgressions can be relatively non-severe without being less moral (copyright infringement, for instance, as well as certain policies of copyright enforcement). To be treated moralistically is not the same as being moral. The blurring of that distinction puts Prinz’s theory in a position similar to psychopaths: it fails to capture the relevant moral/conventional distinction; it fails the test of moral concept possession.

The considerations so far lead naturally to the defense of neosentimentalism. This view, however, has a serious problem on its own. D'Arms and Jacobson (2000) argue that any form of sentimentalism, including neosentimentalism, faces the conflation problem: it yields systematically wrong answers to evaluative questions. This is so because (neo)sentimentalism “offers no resources to differentiate (and hence to preclude conflating) moral and prudential reasons for feeling a sentiment” (D'Arms and Jacobson 2000, p. 732).

Consider the case of a joke: what makes a joke funny is not the appropriateness of being amused by it. An offensive joke may still be funny, though it does not merit amusement. Things can be deemed appropriate in various ways (socially, morally, comically), and there is no way of singling out moral appropriateness without being circular (being ‘right’ is being morally appropriate).

Prinz believes that he avoids the conflation problem because he has no need of singling out any particular sense of appropriateness. The extension of moral concepts, he claims, is
determined by whatever sentiments reliably respond to. Prinz (2007, p. 107) hence believes that, in his view, “moral properties come for free”.

However, I believe that Prinz only puts himself out of the challenge imposed by the conflation problem because he misses the relevant desideratum. His theory fails to be a theory about morality, in the relevant sense. Neosentimentalism has indeed a great challenge ahead, but it at least has a slot to be filled in by the appropriate explanation. Neosentimentalism does not offer all the answers, but it is at least a promising research programme. Prinz’s project, by contrast, is not only incomplete, but it is incapable of getting to anything like an appropriate solution to the conflation problem. The problem itself finds no place in Prinz’s framework. This is so because Prinz’s view fails to capture normativity altogether: asking it to differentiate kinds of appropriateness is to make a demand that is completely out of its domains.
Notes

1 This paper was funded by Capes Foundation, Ministry of Education of Brazil.

2 This clarifying suggestion was made by Schroeter (2006, p. 340).

3 This example is from Putnam (1975). Putnam's notion of stereotype corresponds to Prinz's (2002) prototype.

4 This same point is made by Nichols (2004).

5 The emphasis is mine.

6 Sauer (2012:11) stresses the same point when he says that “by naming the causally sufficient conditions for something to alter somebody's judgment, one misses the point of a conceptual decomposition of the notion”.

7 The psychological importance of emotional devices give support to the claim that the moral deficit of psychopaths derive, at least in part, from an emotional deficit, since it is virtually impossible to have the appropriate moral behavior without automatic reactions to morally relevant features.

8 Prinz (2007, chap.2) develops a detailed theory of emotions as mental calibration files functionally unified by their capacity to track concerns. The “disapprobation range” names a mental file calibrated to represent a certain concern.

9 The most sophisticated versions of neosentimentalism were developed by Blackburn (1998), Gibbard (1990), Wiggins (1987), and McDowell (1998).
References


