

Critical review: *The Emotional Construction of Morals*

Erick Ramirez

Jesse Prinz's The Emotional Construction of Morals is an ambitious and intriguing contribution to the debate about the nature and role of emotion within moral psychology. I review Prinz's recent claims surrounding the nature of emotional concepts as "embodied representations of concern" and survey his later arguments meant to establish a form of cultural relativism. Although I suggest that other theories of emotional representation (i.e. prototype views) would better serve Prinz's aims, the underlying meta-ethical relativism that results is well defended and represents a significant advance for constructivist Sentimentalists.

Keywords: Emotion; Meta-Ethics; Moral Psychology; Sentimentalism

1. Introduction

The late Robert Solomon is quoted on the back cover of Jesse Prinz's last book on emotions, *Gut reactions* (2004), saying that it was "an exciting book, I couldn't put it down, but I fought with it every inch of the way." One might very well say the same about Prinz's companion book, *The emotional construction of morals*. Prinz's arguments are sophisticated, empirically informed, and his conclusions enticing. Still, I fought with it every inch of the way. Prinz's latest project is at its heart an attempt to accommodate empirical data on the emotions with our intuitions about morality. Prinz tries to *explain* our folk notions, not just of emotion but also of our moral practices and, in giving us those explanations, draws intriguing conclusions about the nature of both. His project is audacious in that he wants to give an account of emotion and ethics that preserves many seemingly incompatible aspects of these folk

Erick Ramirez is an Instructor of Philosophy at the University of California, San Diego
Correspondence to: Erick Ramirez, Department of Philosophy—0119, University of California, San Diego,
9500 Gilman Drive, La Jolla, CA 92093-0119, USA. Email: ejramire@ucsd.edu

concepts by showing that emotions, understood as perceptions of somatic states calibrated to “concerns,” are necessary and sufficient to account for moral judgments and moral values as we (the folk) understand them. In doing this, Prinz thinks that the resulting meta-ethical position, speaker relative Sentimentalism, is not only true but can accommodate intuitions about what moral judgments are, why we make them, and why they are justified.

In what follows I address specific questions at only one critical juncture of Prinz’s arguments: his theory of emotion. I will also briefly touch upon his account of moral disagreement, and the meta-ethical relativism he ultimately endorses. Prinz argues that although relativism is thought to imply moral nihilism, his view has the resources to retain most of our moral beliefs. I question the strength of these arguments, especially his claims about relativism and tolerance.

2. What is an Emotion?

One aspect of Prinz’s work that ought to be praised is his making his theory of emotion central to his normative theory. In other words, Prinz wants to make clear that if we’re going to be committed Sentimentalists—committed, that is, to the belief that our emotional responses are what separate the domain of the moral from the non-moral—it matters what our theory of the sentiments actually is. The content of morality, its scope, and its normative authority should crucially depend on the underlying theory of what sentiments are.

Prinz focuses much of the early parts of his book on getting straight just what he thinks emotions are, their origins, and how they impact us. This turns us toward questions in the philosophy of mind, theories of concepts, and questions about motivational internalism. Prinz draws quite heavily on the Neo-Jamesian theory of emotion he developed in *Gut reactions* (2004), though not without important differences. Before addressing those differences, however, I want to pause to address Prinz’s arguments against cognitive theories of emotion, and raise questions not only about those arguments but about his own positive arguments in favor of non-cognitivism.

Much as in *Gut reactions*, Prinz divides the wide-ranging theories of emotion along one dimension: whether propositional attitudes are a necessary feature of an emotional experience. The division then amounts to a debate between emotional *cognitivists* who argue that emotions necessarily involve propositional attitudes (beliefs, desires, etc.), and *non-cognitivists* who deny this. Prinz is a non-cognitivist and amasses impressive support for thinking that any theory of emotion that requires propositional attitudes is false.¹ For example, propositional attitudes are clearly not *sufficient* for having an emotional experience (e.g., I can judge that I’ve been wronged without feeling angry either at the moment of judgment or afterward), but Prinz’s arguments go farther than this. He argues that propositional attitudes of any kind are not *necessary* in order to have an emotional experience. If propositional attitudes are

neither necessary nor sufficient to account for genuine emotional experiences, then cognitivism is false.

One criticism that Prinz makes early on is that cognitive theories seem too demanding to be a realistic account of emotional experience (p. 51). Cognitivists, Prinz argues, require us to make judgments/appraisals of our feelings, when in fact we often find ourselves in an emotional state without any mediation through prior judgments. In other words, although appraisals do at times seem to play a role in getting us to enter an emotional state (“Is that man insulting me? Yes, he is. That makes me furious!”), they don’t seem to accurately account for emotional experiences in which we first find ourselves feeling a certain way and only afterward determine the causal antecedents that put us in that state. As a first pass this doesn’t seem to attack cognitive theories that posit non-conscious judgemental states as essential to emotions. Solomon, in *Thinking about feeling*, for example, allows that:

An emotion is rather a complex of judgments and, sometimes, quite sophisticated judgments, such as judgments of responsibility (in shame, anger, and embarrassment) or judgments of comparative status (as in contempt and resentment). Emotions as judgments are not necessarily (or usually) conscious or deliberative or even articulate, but we certainly *can* articulate, attend to, and deliberate regarding our emotions and emotion-judgments, and we do so whenever we think our way into an emotion, “work ourselves up” to anger, or jealousy, or love. (2004, p. 11)

So, if the criticism here is simply that we don’t necessarily have a conscious judgment that is an essential component of emotion, then this criticism, by itself, misses its mark. Whatever states Solomon is referring to here as “not necessarily (or usually) conscious or deliberative or even articulate” seem compatible with Prinz’s argument that we don’t often find ourselves making judgments when we experience an emotion. However, Prinz isn’t yet done with his assault on cognitivism.

Prinz’s argument against the necessity of propositional attitudes is a bit more complex than I have initially framed it. Prinz contends that how one conceives of thoughts is crucial to determining whether they are necessarily a part of an emotional experience. Thoughts, at a minimum, are defined by Prinz as “mental representations that contain concepts;” accordingly, he claims that “to say that emotions are necessarily cognitive is to say that one cannot have an emotion without possessing and tokening certain concepts” (pp. 56–57). This definition thus makes clearer why a theory like Solomon’s would still count as cognitive on Prinz’s account. However, having added this much, what reasons are we given to think that cognitivism, understood now as a theory of emotion that necessarily requires the tokening of concepts, is false?

Prinz’s appeal is again to emotional *immediacy*. Emotions don’t seem, phenomenally, to require judgments. Cognitivism, he claims, places emotions “too high up on the phylogenetic and ontogenetic scales” (p. 57), i.e., it denies emotions to nonhumans and very young children, which Prinz argues is a mistake.

Appeals to emotional immediacy, again, leave untouched versions of cognitivism that posit non-conscious appraisals as essential to emotion; neither is the appeal, by

itself, a very powerful argument against cognitivism in general. There needs to be some reason for thinking that emotional experience actually *lacks* cognition (not just that it *seems* so to us) and that our phenomenal experience with emotion is not only essentially non-cognitive but also that it should serve as arbiter in the debate between cognitivism and non-cognitivism.² Solomon, to use a familiar foil, is all too happy to include animal emotions within his taxonomy despite his commitment to a form of cognitivism about emotions.

The real meat in Prinz's arguments against cognitivism comes from his extensive use of psychological data on emotion to support his conclusion that cognitivism is false. He argues that if experimental data can show that emotions can be elicited by bypassing cognition altogether, then we have good reason to conclude that cognitivism—at least a cognitivism that states that *all emotions* must *always* include propositional attitudes—is false. At best, if the data point in this direction, cognitivism would find a home in some “mixed theory” of emotion. How does Prinz propose to show this?

Prinz appeals to studies that he believes show that emotions can be induced without judgments. Here he refers to studies purporting to support the facial feedback hypothesis, which simply states that moving one's facial muscles into specific emotional configurations (smiles, frowns, and so on) can *induce* the emotions corresponding with their facial expressions—smiling tends to make us feel happy, frowning makes us feel sad, etc. Research on facial feedback is quite robust, and I here don't want to take issue with the data or the hypothesis so much as I want to take issue with what Prinz believes facial feedback can tell us about emotion (for more on facial feedback, see Buck, 1980; Laird, 1974; Levenson, Ekman, & Friesen, 1990; Rourangeau & Ellsworth, 1979).

Facial feedback, Prinz tells us, does the following:

1. Causes distinctive patterns of bodily changes in subjects
2. Causes subjects to report feeling emotions. (p. 58)

Prinz wants to use facial feedback data to claim that “somatic signals” are sufficient for emotion—but how good is this argument? If, as it is impossible to deny, facial feedback can cause affective change, what could one say in response?

Here I want to focus on the second claim about facial feedback. Without the subject's self-reports, without subjects saying to the researcher, for example, “I'm feeling a bit happier,” after being manipulated into smiling (e.g., by holding a pen in the mouth) the argument against cognitivism would fail because, crucially, self reports require *conceptualizing experience*. They require deploying emotional concepts in order to arrive at a judgment about the emotional state a subject is in. When a subject responds, after having her facial muscles oriented in a certain way, that she feels happier or sadder, she is reporting on her internal states via some understanding of her phenomenal experience and knowledge about what her personal emotional repertoire is, her introspective and intersubjective history using emotional concepts, and how her current affective state relates to that background information.

These reports, therefore, shouldn't be taken as an indicator that facial feedback is eliciting non-cognitive emotional responses from subjects. The assumption here seems to be that the measurable physiological change is being immediately reported on by the subjects without the aid of conceptualization—but this seems unlikely. If cognitivism only requires the tokening of concepts, however, then it is unclear why a subject's reports would be useful evidence against cognitivist theories generally. But Prinz has another argument available here. Prinz gets out of this dilemma by instead focusing on the facial feedback hypothesis' first claim and arguing that emotions can be separated from self-reports and identified physiologically to at least some degree, because there are distinctive patterns of physiological states that correspond to different emotions. This gives him room to claim that the physiological changes that result from facial feedback can be used to identify the emotional state that a subject is in, independent from a subject's own reports about her emotional states (thereby bypassing the self-report problem I mentioned above). This claim, call it the "independence claim," is also false. In the next section I offer a sketch of an alternative theory of emotion that would explain this data equally well without committing itself to non-cognitivism *or* cognitivism about emotions in general. For the moment I turn to Prinz's arguments regarding the representational content of emotion, before returning to the independence claim and its relation to what Prinz calls an emotion's "calibration file."

Prinz argues that emotions are mediated by relatively primitive brain structures like the amygdala and the thalamus, and that these brain structures are clearly not harboring or tokening concepts when they operate. Recall that Prinz ultimately wants to identify emotions with particular physiological changes that have evolved to represent universal human concerns. Since these physiological changes are not mediated by judgments they are immediate, they are non-cognitive. Departing from the traditional James-Lange theory of emotion, Prinz does believe that emotions have representational content. On the classical James-Lange theory, emotions are simply perceptions of internal states, states of bodily change. The intentional contents of emotions, what it is that they represent, are, on this view, these states of bodily change; Prinz argues that the representational content of emotions is far richer than this (p. 51).

Prinz wants to develop a theory of representation that can capture the richer content of emotional episodes but without requiring the tokening of concepts (p. 61). He argues that emotions do not represent bodily states *per se*, but that they instead represent those things that, in virtue of their structure, bodily states reliably detect. On his view, "emotions represent concerns," and concerns are defined as organism-environment relations that bear on well being (p. 63). Each emotion will have a corresponding concern that is represented by particular somatic states. To have an emotion is to become aware of being in one of these somatic states. Prinz therefore needs some way to explain how emotion can represent these concerns. A theory of representation is needed that can handle this task non-cognitively.

Prinz appeals to Dretske's (1995) theory of mental representation to ground a theory of emotional representation without needing to appeal to the tokening of

concepts. On this account, in Prinz's words, "a mental representation, M, represents that which it has the function of reliably detecting. Roughly, M represents that which it was set up to be set off by" (p. 61). Each emotion then is reliably set off by a specific concern, and the bodily states (the emotions) represent these concerns in virtue of their causal relationship with concerns. That is, emotions are set up to reliably detect specific concerns, and in virtue of being set up in this way they can be said to properly represent them.

For example, Prinz argues that sadness might have plausibly evolved as a loss detector. If losses reliably put us into the somatic state of sadness, then sadness can be said to *represent* losses (pp. 62–63). Loss seems like a plausible candidate for an organism-environment relation that bears on well being. Losing something, especially something one finds valuable, seems to be the kind of thing we would think is important for us to keep track of, and the emotion we call "sadness" is the emotion that tracks these losses.

Here we might pause again to consider Prinz's commitment to non-cognitivism, and pose a question for a non-cognitive theory of emotion that hitches itself to Dretske's theory of representation. Loss, we might plausibly say, is a relatively complex relation between an individual and his or her environment. It seems to require an agent to be able to detect that they had something of value and that that thing (a lover, money, tenure, religious beliefs, etc.) is no longer in their possession. Furthermore, not all losses make us sad. So it seems like the conception of loss that is tied to sadness must be somewhat more complex than Prinz's initial gloss that sadness is a loss detector simpliciter.

Is it plausible to say that emotions represent losses before an agent has a concept of loss? If conceptions of loss or, more specifically, conceptions of losses that are meaningful, are required in order to be able to register something as a loss, then it seems as if some relation to the tokening of concepts might be necessary for emotional experience after all. Prinz of course wants to avoid this conclusion. His example here is of a smoke detector's alarm representing smoke. Obviously a smoke detector doesn't need to token any concepts to be reliably caused to beep in the presence of smoke, even though we might say that the smoke detector, when it beeps, is representing smoke. How close is the smoke detector case to an instance of sadness? There is an important disanalogy between the two cases. *Any* smoke would set off a smoke detector. Presumably the reason we're interested in detecting smoke is because where there is smoke there is fire, and it's the fire we're really interested in detecting.

The same is not true of sadness. Unlike the smoke detector which beeps in the presence of smoke, not all losses make us sad. We can try, for example, to imagine becoming sad at the loss of a stranger's pocket lint, but it is highly unlikely to elicit sadness from us. At the very least, some story needs to be told about how complex representations like sadness can be represented by a somatic state without the tokening of any concepts at all. Losing a nickel doesn't make us sad even though it is a genuine loss. Losing a friend however does tend to make us feel sad, indeed profoundly so. In this sense, then, sadness isn't simply a loss detector in the same way

that smoke alarms are simply smoke detectors. Sadness, we might say, is meant to detect the loss of things we think have value. Simply assuming that the somatic state associated with sadness is set up to detect losses of value without the tokening of concepts is a bit less plausible than the story about the smoke detector. To detect genuine losses the relationship between our somatic states and losses of value must make an end run around conceptions of value, and detecting value is less obviously non-cognitive than Prinz' smoke detector.

I don't intend for this to be taken as a knock down argument against non-cognitivism. My point is only that emotions and their concerns can be construed as related to one another in ways more complex than a smoke detector is related to smoke. Although we can set up a smoke detector to sit in a causal relationship so that it beeps in the presence of smoke (smoke \rightarrow beep), the same doesn't appear to be true of sadness and lamentable losses (or fear, or happiness, or any emotion). What we need is a theory that explains why only merited emotional responses count as genuine instances of representation for emotion. Recognition of this problem leads Prinz to make an interesting claim about where this extra information about emotions resides.

Prinz's move in response to the seemingly cognitive features of emotions is to argue that the conceptual machinery typically associated with emotions—the kinds of complications I've been explicitly appealing to thus far—are not part of emotions proper, but instead are a part of an emotion's "calibration file." Calibration files "contain a wide range of representations, both cognitive and non-cognitive, and these representations change over the course of cognitive development" (p. 63). These files are not, however, part of the emotion itself. The emotion itself remains the somatic state that represents the concerns, which the calibration files contain all of our knowledge about. It isn't clear, however, that Prinz should be entitled to make this move. Why think that calibration files aren't necessarily a part of an emotional episode itself? Can we experience any emotion without having a calibration file for it?

If the content of an emotional experience resides in the calibration file, then we need at least some independent motivation (aside from a commitment to non-cognitivism) to think that calibration files should be kept apart from emotions proper. I find this move motivated more by a pre-commitment to non-cognitivism than it is by an attempt to do justice to the phenomenology of emotional experience itself. To make this point clearer I want to highlight Prinz's specific response to the "somatic similarity problem" (SSP).

Prinz's position, as originally conceived, suggests that each emotion has a particular somatic state that corresponds with it, i.e., a particular physiological state that can be used to identify the emotion in question. The SSP, as posed by Prinz, is that there simply aren't enough unique somatic states to account for all of our emotions.

In light of this problem, Prinz proposes to again modify his position. He argues, via analogy, that the same somatic state can come to represent *different* concerns by a re-calibration of an emotion's file. If this is true, then the same somatic state could come to be associated with different emotions, and hence could get around the SSP. How does he do this?

Recall that for Prinz an emotion is a somatic state that reliably detects something—a concern—and it is this concern that the emotion represents. It is what the emotion is “about.” Prinz argues that the SSP can be overcome if different mechanisms come to put us into the same somatic state. For example, consider anger and indignation: in such a case “one mechanism is a collection of impressions and ideas pertaining to injustice” (pp. 66–67); the other mechanism would be a collection of impressions and ideas pertaining to whatever the concern for anger turns out to be. But here again the calibration files are doing all the work in differentiating emotions from one another. In what way can we be said to be angry without a calibration file for anger? Put more aggressively, how can we get angry without accessing our calibration file as an essential component of an angry experience? Indeed this move undermines some of his earlier arguments for the independence claim (i.e., that the emotional state a person is in can be identified independently of self-reports about his or her emotional state). If we need to be able to conceptualize our emotional experience (by accessing our calibration files in order to determine the mechanism that put us into a particular somatic state) in order to be in one emotional state rather than another (given a shared somatic state), then it seems like calibration files, and conceptualization, are after all a necessary part of our everyday emotional experience.

Prinz argues that if emotions work this way—that is, if we are able to detect which mechanism was operating to put us into the particular somatic state that we find ourselves in—then we will know whether we are angry or indignant, sad or guilty, and so on. The causal elicitors are, as Prinz has already had occasion to note, part of each emotion’s calibration file. Different calibration files can put us into similar somatic states, but the calibration files, as part of the causal chain leading up to an emotional experience, determine the particular emotion being experienced.³

I raise these questions here only to note instances in which the dialogue is very much alive and in dispute. Prinz’s sustained and interdisciplinary argument in favor of his embodied emotional non-cognitivism is persuasive and passionate, and has already left its mark on the field in the short time since the book was published. Although my focus in this review has been on Prinz’s theory of emotion, I want to now suggest an alternative theory of emotion, a prototype theory, and argue that this kind of theory better fits the facial feedback data, better explains the nature of emotional “calibration files,” and better accommodates Prinz’s argumentative purposes than his own professed view. In the final section, I end by touching on the consequences that Prinz believes his theory of emotion has on meta-ethics (his chief aim in the book).

3. An Alternative Hypothesis

The theory of emotion I propose in this section has a rather short philosophical pedigree. de Sousa (1987) has come closest to proposing the view I have in mind, when he suggested that emotions are partially biologically based and partially socially

constructed via what he called “paradigm scenarios.” A paradigm scenario, for de Sousa, is a learned script that helps us determine when to experience specific emotions and what those emotions require us to think and do. The general family of theories here, often called “prototype theories” of emotion, have a longer history in the psychological literature (Russell, 1999, 2003). The aim of these theories is to incorporate theories of biological affect and socially learned emotion in order to provide a more complete theory of our emotional concepts and emotional experience.

A prototype theory of emotion is poised to explain the data on facial feedback without being strictly non-cognitive. Prototype theories lie somewhere between cognitive and non-cognitive theories. They make a distinction between occurrent emotional episodes and the emotional prototypes we use to understand those episodes. While our emotional prototypes often include propositional attitudes as component parts, our occurrent emotional episodes need not include them. Our script for anger, for example, may include the belief that I have been wronged, and hence include propositional attitudes (beliefs) as parts. A particular episode of anger, however, could be entirely non-cognitive so long as the other components of the episode (behavioral dispositions, phenomenology, and so on) more closely match the script for anger than other rival scripts. In this sense prototype theories of emotion suggest that both cognitivists and non-cognitivists get something right about emotion, but at different levels of analysis. It’s interesting that Prinz ignores this option since, as with his appeal to the independent plausibility of Dretske’s theory, prototype theories of emotion are motivated by an independently plausible theory of mental representation. At the very least, Prinz’s objections leave untouched theories of emotion that construe emotions as having a prototypical character.

Prototype theories of emotion also make an important distinction between affect and emotions proper. Affect is understood as “core affect,” and is definable as a two-dimensional space that delineates an agent’s level of arousal and valence. Affect is understood as a biologically primitive non-cognitive state. All of our experiences can be mapped onto some point in the space represented by core affect. Figure 1 represents one standard way of representing this “core affect” space. Affect represents the “feel” of an emotional state, but prototype theorists insist that core affect is only one component of an emotional experience. Affect should not be understood as an

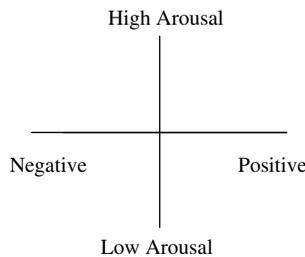


Figure 1 A Circumplex Depicting Core Affective States. Modified from Russell (2004).

emotion. Other components of an emotional experience would incorporate the elements Prinz includes within an emotion's calibration file: our history with experiencing different affective states, the action tendencies they've been associated with, characteristic thoughts that accompany the emotion, objects that have tended to elicit specific emotional responses, etc.

A prototype theory would embrace the SSP: it would accept that there seem to be more emotions than there are distinctive physiological states. The SSP is only a problem if one thinks that emotions are identifiable with particular physiological states. The prototype theorist can explain this problem away, however, by arguing that two particular emotional experiences can and often do share the same point in core affective space (i.e., two type-distinct emotions may have the same level of arousal and valence) but that emotions are differentiated by other components that make up the prototype of each emotional concept. In cases of somatic similarity, a transformation function that ties the particular affective state being experienced with the level of similarity that it bears to the emotional concepts we have at hand would be enough to differentiate one emotional state from another. Anger and contempt, for example, may share similar points in affective space, but can be differentiated by the fact that we tend to make a distinction between the common eliciting conditions, action tendencies, characteristic thoughts, etc., that are associated with each emotional experience.

Note that for prototype theories of emotion, one may be said to experience a *genuine* instance of an emotional episode such as anger even if some of the characteristic features of a paradigm experience of anger are missing. So long as the experience falls closest to the space carved out by our conception of anger (relative to other emotional concepts), then that experience is a genuine instance of anger. Most relevant for my purposes here is that, according to prototype theories of emotion, judgments themselves may be lacking in some instances of anger and present in others; but this fact doesn't serve to make one an instance of "genuine" anger and the other some ersatz form of anger ("schmanger"). This claim is empirically testable. Subjects can and do judge that narratives of occurrent cognitive *and* non-cognitive anger represent genuine instances of anger, rather than some another emotion (see Russell & Fehr, 1994, for one such study). Prototype theories of emotion therefore can account for the cognitive features of emotional experience without positing them as *necessary* conditions, and without relegating the seemingly cognitive features to mere emotional antecedents (as Prinz does by locating them within an emotion's calibration file).

This isn't the place for a full defense or articulation of prototype theories of emotion, but there is much to be said in their favor. I mention them here because I want to offer them as rival theories of what emotions are that can account for the features Prinz seems most interested in capturing about emotion without the need for calibration files. Furthermore, prototype theories are fully compatible with both subjectivism and relativism, and hence might be better suited to Prinz's own meta-ethical purposes. In this sense, then, prototype theories of emotion do not fall prey to the arguments Prinz amasses against cognitive theories of emotion, and are poised to

explain our experience with emotions at least as well as Prinz's non-cognitive embodied appraisals can. Because prototype theories of emotion promise to explain the apparent cognitive and non-cognitive elements of emotion as resulting from prototypical as opposed to occurrent emotional experiences, a prototype theory can accommodate Prinz's non-cognitivism as a component of a theory of occurrent emotional episodes. Both Prinz's preferred somatic theory of emotion and the prototype theory I have offered here would lead to similar relativistic conclusions about the nature of moral language. I close by considering Prinz's arguments on the consequences of embracing this kind of moral relativism.

4. Closing Thoughts on Moral Relativism

Prinz lays the relativistic implications of his theory out for all to see, and defends the position vigorously. As Prinz articulates his ultimate position, metaethical relativism can be defined by the following schema:

Metaethical Relativism (extended definition): The truth conditions of a moral judgment depend on the context in which that judgment is formed, such that:

A judgment that X *ought* to \emptyset is true if and only if it is wrong not to \emptyset on the value systems of both the speaker and X.

A judgment that \emptyset -ing is wrong is true if and only if \emptyset -ing is the object of a sentiment of disapprobation among the contextually salient individual(s) (usually the speaker). (pp. 174–180)

Here Prinz notes an important and underappreciated difference between kinds of moral judgments. For the metaethical relativist, what we might call “ought” judgments only make sense within a moral community that shares many overlapping values. If I demand that you ought to keep your promises, this demand only makes sense if we both already agree that promise keeping is something to be valued. On the other hand, Prinz argues that judgments about the rightness or wrongness of an action or state of affairs merely *express* the moral values of the speaker. Because these judgments do not place demands on the addressee, these kinds of judgments apply beyond the scope of the speaker's moral community such that I could plausibly say that many applications of sharia law are wrong as an expression of my disapprobation towards that practice. Only ought judgments, what Prinz calls “oughtittudes,” are meant to play a role in giving reasons to the *addressee* of the moral judgment.

Prinz's view is the result of commitments he believes are implied by his theory of emotion, and commitments that result from his theory of what moral emotions are. Prinz argues that moral emotions are those emotions that evoke contempt, anger, or disgust. Any norm violation that triggers these emotions is a moral violation. Having argued that emotions are best understood as somatic representations of concerns, and that *moral* emotions are those emotions that correspond to specific kinds of violations, he also argues that culture can dramatically influence the specific nature of

our values by shaping our grounding norms. Grounding norms are “rock bottom values” (p. 125). They represent values for which we cannot give justification. We may be able to defend a progressive taxation scheme, for example, by appealing to notions of fairness. If asked to justify our concern for fairness, however, we may be at a loss. Fairness, in this example, is a grounding norm. Although we cannot help but care about fairness we may not be able to non-circularly justify that concern. Grounding norms are the end of the line when it comes to rational discourse. If two parties do not share grounding norms, then rational moral disagreement is impossible for them. “If two people have different grounding norms, they must resort to other means of persuasion” (p. 125) beyond reasoning. The result of committing oneself to the claims that emotions are somatic representations of concern, that moral emotions result from specific emotion-inducing violations, and that grounding norms are culturally variable is meta-ethical relativism. What Prinz goes on to do in his final chapter is defend this form of relativism, and in that spirit I find that Prinz is more or less successful.⁴

Prinz thinks that the principle challenge to relativism is that it is considered to be an “insidious doctrine.” It is thought to be insidious because it has the following implications:

1. If morality is relative then no morality is any better than any other.
2. Our morality is not privileged.
3. Moral disputes are spurious with those who don’t share our morality.
4. Discovering the truth of (2) will weaken our confidence in our own moral values. (pp. 205–206)

As an example, Prinz has us consider political ideologies. He argues that the difference between liberals and conservatives might terminate in a difference of grounding norms, and thus that political debates between these two groups are spurious. Not only are the debates intractable, but liberals and conservatives may be, literally, talking past one another. They fail to actually disagree with one another.

Grounding norms seem to come from the way we are “constituted,” but here, where we really need an account of how grounding norms are constituted, we aren’t given much. Grounding norms themselves don’t serve as reasons in moral discourse (they are “rock bottom,” recall); rather they seem to delimit the moral population with which we can hope to engage in moral discourse at all. If this weren’t true, then “debates between [liberals and conservatives] would be much easier to resolve, and the differences would diminish as the facts came in” (p. 193).

If Prinz is right about this, then he thinks that worry (3) might be true of political debates. They would be spurious because adherents of different political ideologies don’t share the same grounding norms. They would be talking past one another instead of engaging in a substantive debate. Furthermore, discovery of (2) would make it such that worry (1) also crops up quite quickly. “An intellectually honest liberal would have to admit that there is no position from which conservatism is objectively worse” (p. 206) because there is no neutral standpoint from which to make this kind of judgment. This will conclude with the truth of (4) and the

weakening of the intellectually honest liberal's own commitment to her values. Prinz finds this kind of worry to be "the most pressing challenge facing the relativist" (p. 207).

Some of Prinz's arguments appeal to the positive consequences that would result from realizing the truth of relativism. Relativism should function to promote *tolerance*. Discovering the truth of claims (1) and (2) should therefore be seen as having the positive consequence of making us less likely to harshly judge those with different values, because we realize that our moral reasons may not apply or even be comprehensible to them.

Of course this only seems like a positive consequence from a certain standpoint. Why think anyone, or even many people, would support tolerance if they discovered that their moral judgments were relative? Why wouldn't it foment intolerance instead by painting those with different grounding norms as distinctly "other," "not one of us," and "to be beaten into submission since rational debate is hopeless" (p. 208)? It isn't clear why tolerance follows from relativism as easily as Prinz seems to think it would. Prinz seems aware of this worry, but thinks that intolerance would be hard to sustain psychologically once the truth of relativism is known. But this seems to be a placeholder for an argument instead of an argument itself.

Furthermore, Prinz argues that relativism doesn't undermine our moral convictions. He argues that this worry is doubly exaggerated since he thinks that the harm of losing some of our moral convictions is overstated, along with the degree to which relativism implies the weakening of those convictions. If we come to realize that our opposition to differing cultural practices is based on holding to an untenable—and here I assume he means empirically false—form of moral absolutism, then Prinz thinks it wouldn't be a bad consequence if we were to weaken our opposition to those practices.

Prinz ends his book by arguing that relativism doesn't imply that all moral systems are equal. There isn't a *moral* standpoint from which we can claim that one moral system is better than another because of the truth of relativism. There may however be a *non-moral* standpoint from which one could conclude that one moral system is better than another. Prinz appeals to non-moral values like consistency, the effects on well-being that moral systems have, the universalizability of a system's moral claims, the degree to which false beliefs are required to sustain a moral system, the degree of social stability a moral system brings about, the ease of implementation of different moral systems, and consistency with "pre-moral norms," as tools with which to assess one moral system or to compare moral systems against one another (pp. 290–297).

Of course we have to value these non-moral values in order for them to serve as arbiters between moral systems, and Prinz readily admits this point. Prinz is committed to arguing that no standpoint exists from which we can claim that one moral system is *morally* better than another, but that we can make other comparative statements about moral systems that can function both to make better/worse claims about moral systems and also account for a sense of moral progress (though this

sense relies, again, on it being the case that parties share the non-moral values being used to make the comparative statements).

I conclude my analysis of Prinz's book by saying that the approach that Prinz is developing here is perhaps the best argument that a relativist has at her disposal to stave off the claims that relativism is insidious. If Prinz is right about the scope of our moral claims, especially oughtitudes, then he can successfully defend an important sense of moral progress by appealing to just the kinds of non-moral values he points us to. Some of these non-moral values are apt to run deep. That is, we would be hard pressed to give up our commitments to consistency and well-being, and values like these are the right ones to appeal to in order to both make corrections from within (that is to improve our own moral practices) and to make claims about competing moral systems.

There is much to like about Prinz's arguments throughout the book, and they need to be taken seriously by subjectivists and objectivists alike. While I fought with him on many particulars, Prinz's arguments are compelling, ingenious, and hard to refute. Sentimentalists would be richly rewarded by taking his arguments into account.

Notes

- [1] This is a large group that crosses disciplinary boundaries between philosophy and psychology.
- [2] Prinz needs to show that whatever it is that emotions are, they must be in some way independently identifiable and separable from cognitions in general. Representations here are simply standing in for one particular kind of cognition that seems, *prima facie*, to belong to our folk concept of emotion.
- [3] Interestingly, Prinz suggests that when we don't know which mechanism put us into the somatic state, we may need outside help, psychotherapy, to figure out which emotion we are actually experiencing (p. 66). Here again the question of conceptualization seems to rear its head doubly so on this occasion. The therapist must conceptualize the emotion the patient is experiencing based on his or her reports and the patient, as a result of this, then conceptualizes her own experience as a result of this process.
- [4] Prinz takes these consequences quite seriously: my grounding norms would not be seen as reasons for someone who did not share them. Philosophers who want to show that moral norms provide robust reasons for action are embarking on a fool's errand (pp. 125)

References

- Buck, R. (1980). Nonverbal behavior and the theory of emotion: The facial feedback hypothesis. *Journal of Personality and Social Psychology*, 38, 811–824.
- Dretske, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.
- Laird, J.D. (1974). Self-attribution of emotion: The effects of expressive behavior on the quality of emotional experience. *Journal of Personality and Social Psychology*, 29, 475–486.
- Levenson, R.W., Ekman, P., & Friesen W.V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27, 363–384.
- Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. New York: Oxford University Press.
- Rourangeau, R., & Ellsworth, P.C. (1979). The role of facial response in the experience of emotion. *Journal of Personality and Social Psychology*, 37, 1519–1531.

- Russell, J. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76, 805–819.
- Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145–172.
- Russell, J., & Fehr, B. (1994). Fuzzy concepts in a fuzzy hierarchy: Varieties of anger. *Journal of Personality and Social Psychology*, 67, 186–205.
- Solomon, R. (2004). *Thinking about feeling*. Oxford: Oxford University Press.

Copyright of *Philosophical Psychology* is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.