

An Enactivist Theory of Emotional Content

Daniel Shargel & Jesse Prinz

Most of the prominent theories of emotion identify them with some kind of representational state: judgments, perceptions, desires, sui generis mental attitudes, or compounds of these states. These models emphasize the idea that emotions have meaning or content; they present the world as being a certain way. Embodied theories of emotion complicate this model. Facial expressions and accelerated heart rates do not, at least on first inspection, resemble judgments or perceptions. We believe embodied theories are right, though we won't review the evidence here (see Prinz, 2004). Instead, we want to consider how embodied theories should cope with the question of content. We will review some of the dominant strategies that defenders of embodiment have taken, and we will argue that each faces problems. We will then consider two more recent approaches: one that emphasizes affordances, and another that defines emotions as "enactive." Building on the former approach, we will develop a version of the latter: an enactivist theory of emotional content. Our theory makes a strong claim about the ontology of emotions: that emotions do not represent pre-given features of the world, but rather bring new properties into existence. Thus, emotions do not represent objective features of the world, nor do they represent response-dependent features. Indeed, we don't think it is accurate to say emotions represent at all. Rather, they bring a special class of properties into existence.

1. Embodiment and Emotional Content

Embodied theories of emotion claim that bodily changes, or perceptions thereof, are essential components of emotional states. They have a long history, debating back to the earliest days of theorizing about emotions, since it has long been observed that the body undergoes changes when we experience emotions. Aristotle (1987) related anger to warm blood around the heart, and he defends a view according which emotions involve physical states playing characteristic functional roles. Descartes (1989, pp. 39-40) advanced the view that emotions are judgments that follow upon movements in the animal spirits. Francis Hutcheson (2003) defines emotions as confused sensations accompanied by bodily motions. Two centuries later, Charles Darwin wrote a systematic study of emotional expressions, suggesting that bodily states can be used to differentiate emotions. Such accounts reveal that the body has long figured into theories of the emotions. But embodiment got a major boost in the late 19th century, when William James and Carl Lange simultaneously advanced more thoroughgoing embodied theories. For James (1884), emotions are perceptions of bodily changes, and Lange (1885) rails against the view that bodily conditions are mere contingent effects of emotions.

The James-Lange approach continued to find supporters in the 20th century and among contemporary authors (see Damasio, 1994; Prinz, 2004; Shargel, 2014). Research has shown that emotions correlate with patterned bodily changes, that induced bodily changes can induce emotional experiences, and that disruptions in bodily perception of

change can have an impact on emotional intensity. The embodied approach has not earned universal support, however. One key reason for this is that it is not immediately clear how changes in the body, or perceptions thereof, can have meaning or content. As Aristotle (1987) and Descartes (1989) argued, emotions present the world as being a certain way. In the 20th century, this conjecture took precedence and a range of “cognitive appraisal” theories emerged, which identify emotions with judgments or other cognitive states (Kenny 1963, p. 14). On appraisal theories, each emotion corresponds to some kind of mental representation, corresponding to a “core relational theme”—a relationship between organism and environment that bears on well-being (Lazarus, 1991). Thus, sadness represents loss, anger represents offense, and fear represents danger. When we experience these emotions, we are representing such relations. This helps to explain the distinct role of each emotion, as well as the criteria for assessing their appropriateness, or rationality, on any given occasion. To fear a benign snake, for example, might be considered irrational because it would be mistaken to represent a benign animal as dangerous.

Under pressure from cognitive appraisal theories, defenders of embodied approaches have faced a challenge. How can one insist that emotions are embodied, while also accommodating the intuition that they have meaning? If fear is constituted by bodily changes such as blood flow to the extremities and a racing heart, how can it signify danger? Such abstract content is thought to exclusively belong to mental representations, not bodily transformations.

To face this challenge, proponents of embodied theories have adopted several strategies. We want to review these now, and also indicate why each of the leading strategies might be regarded as problematic. One strategy is to bite the bullet and simply deny that emotions have content. As we read them, both James and Lange essentially ignore the issue of content, so they might have been partial to this approach. One of us has also defended a version of the bullet-biting view, claiming that emotions may lack precise intentional content (Shargel, 2015). There are benefits of bullet biting, since emotions differ in important ways from typical representational states. When we are angry, for example, the anger influences behavior directed at many things in our surrounding, not just an individual object. This is a point we will revisit at the end. But bullet biting also has a considerable cost. Content helps to explain why emotions play different roles. It explains why anger is related to aggression by supposing that anger represents offense. It would be good to have some account that explains this.

A second strategy is to endorse a hybrid model, on which emotions have both a cognitive component and a bodily component. The challenge here is justifying the unification of these components, as well as the identification of that state with the emotion. What makes the two elements part of a single state? They are demonstrably dissociable (we can represent the presence of danger without entering into a physiological state of fear, and conversely). Moreover, such approaches implicitly concede that bodily states cannot have meaning. This forecloses the interesting possibility that embodied states make substantive contributions to content.

A third strategy, once favored by one of us (Prinz, 2004), is to argue that physiological states form part of an embodied representational system. Physiological states carry information about significant situations, so by perceiving these states we

indirectly perceive those situations. On this model, emotions use interoceptive modalities to achieve a form of exteroceptive perception. Prinz tries to pull this off by adopting a popular naturalistic approach to representation: informational semantics (Dretske, 1981). According to informational semantics, a mental state gets its meaning in virtue of the information it carries. Information, in turn, is characterized in terms of some kind of covariation relation. If state M reliably occurs in the presence of stimulus S, then M indicates the presence of S. To a first approximation, then, a racing heart and blood flow to the extremities can represent danger because such states tend to arise when and only when we are in danger. This is called the embodied appraisal theory, since bodily state, or perceptions thereof, come to have the kind of content that is traditionally associated with appraisal judgments.

This approach faces some serious hurdles. Apart from worries about informational semantics in general,¹ standard versions of informational semantics are difficult to adapt to the case of emotions, since they presuppose that the representational content of a mental state can exist independently of that state. Representation is said to involve a reliable co-occurrence between mental states and items in the external world. But some experiences seem to represent things that lack this kind of external existence. For example, if one experiences gustatory pleasure after eating a bite of peanut butter, there is nothing in the peanut butter that corresponds to its pleasantness; others find it unpleasant. Some emotions seem to fit this model as well. Prinz (2004) tries to argue that emotions do not represent secondary qualities (things like deliciousness, which depend on us), but this is a very controversial position with respect to emotional content. Finally, the approach faces another objection, which threatens hybrid theories as well. Informational semantics claims that content is determined by covariation between mental states and the world. But, on that approach, there is nothing about a given mental state that determines what it refers to – mental states are no more than arbitrary symbols (Fodor, 1990). Therefore it is irrelevant that emotional components are realized by bodily states and bodily perceptions (Shargel, 2014). One could easily represent the same content without the bodily change, since the theory appeals to their information-processing characteristics and not their bodily character. Moreover, situations such as danger can result in different bodily reactions, ranging from flight to steadfast confrontation. If content were determined by covariation, the theory might render such differences semantically irrelevant; this is an embarrassing outcome for a theory that emphasizes embodiment, since such differences seem to correspond to distinct embodied construals of the precipitating events. The embodied appraisal theory tries to make bodily states the bearers of content, but, in the same gesture, it makes them irrelevant to content determination.

2. Emotional Content and Affordances

¹ For example, informational approaches have been criticized on the grounds that a given mental state (or bodily state) can have arbitrarily many causes, so information alone cannot settle which of these fix the meaning of the state, and which are false positives (Fodor 1990).

We want to turn now to another account of emotional content, which departs from these earlier proposals. Hybrid theories and the embodied appraisal theory both follow cognitive appraisal theories when it comes to the content of emotions. While hybrid theories simply build in cognitive appraisals, the embodied appraisal theory assigns standard appraisal contents to embodied states. Both approaches give the impression that the bodily changes that take place are irrelevant when it comes to determining content. There is a recent alternative that promises to get around this problem. Paul Griffiths and Andrea Scarantino (2009) have advanced the interesting proposal that emotions represent affordances. We will briefly describe the account and some of its advantages and disadvantages.

The notion of an affordance owes to psychologist of perception J.J. Gibson. Gibson (1979) argues that visual states represent more than shapes and locations of objects. Vision also informs us about possibilities for action. Thus, if we see a hammer, we not only see the length and orientation of its handle; we also see that it is graspable. Graspability is an interesting feature because it is relational. It is not intrinsic to the hammer. Rather it depends on aspects of the observer's body: hand aperture, the ability to close ones fingers, the presence of a prehensile thumb, and so on. It is a feature that depends on how the observer is embodied. To see an object as graspable may also recruit motor responses. One might spontaneously prepare to grasp a hammer because its shape is associated through learning with grasping (Grèzes, 2003).

Griffiths and Scarantino notice that certain aspects of affordances lend themselves to a theory of emotional content. Emotions have long been recognized to have associated actions. When the body changes during an emotion, these changes are not arbitrary, but rather underlie what Nico Frijda (1986) calls an "action tendency." However, Frijda and the many others who make this observation have not suggested that the action tendency is part of what an emotion represents. This is where Griffiths and Scarantino come in. Griffiths (2004), for example, proposes that anger represents the fact that aggression is afforded, and shame represents that the acceptance of another person's dominance is afforded. Griffiths and Scarantino (2009) elaborate on the view. Instead of aggression, they propose that anger represents a retribution affordance, and they note that retribution can take many different forms—sulking, insulting, storming out, etc.—depending on the context.²

The affordance theory has clear advantages over some of the proposals surveyed above. For one thing, affordances are a kind of object that depends on having a particular embodiment, and the perception of affordances characteristically involves bodily changes. Aggressing, sulking, and storming out are all physical activities. The bodily changes that take place when a person is angry are not arbitrarily related to these, but rather prepare the body to act. If emotions represent affordances, emotional embodiment bears a direct relationship to emotional content.

² The affordance theory is part of a broader theoretical model in which the authors emphasize the communicative impact of emotions. Sometimes we show anger, for example, in order to achieve an end. We will touch on this aspect of their account in the final section. Here we want to focus on the claim that emotions represent affordances.

The affordance theory also provides theorists with the opportunity to develop an account of emotional content that avoids informational semantics. Affordances depend on physical abilities, and this invites an embodied approach to content determination. Affordances also bypass the need to posit a cognitive accompaniment, as in the case of hybrid theories, since they are recognized through embodied perception. We may perceive affordances by perceiving the physical abilities that determine them.

We think the affordance theory is going in the right direction, but it still faces some problems. For one thing, Gibson (1979) regards affordances as relational, but in some sense objective; they are invariant features of the environment out there to be perceived. Just like shape and location, affordances exist even when they go unnoticed. A hammer affords grasping even when it hasn't been seen. Embodied responses to a hammer play a role in our "picking up" on the affordance, but the affordance was already there. Changes in the body can alter affordances, on a Gibsonian account, but only when these are long-lasting changes that lead to new invariant relations. For example, after doing a lot of weight training, certain heavy objects might become liftable. But, in general, liftability is a function of one's physical traits, not one's occurrent states. For Gibson, a heavy object can have the property of liftability for a weightlifter even when she is occupied by some other task. A hammer is heftable even when we have no incentive to lift it. These are standing features of the world. Scarantino (2004) notices this aspect of affordances, but he also insists, perhaps departing from Gibson, that affordances can change with context. A tree, he notes, does not afford climbing for a squirrel if there is a barrier in the way. This kind of context sensitivity is fairly minimal, however. The example implies that any accessible tree within the squirrel's immediate environment affords climbing, even if the squirrel is currently foraging on the ground. Climability can be instantiated independent of the animal's current activities.

With emotions things seem different. It would be absurd to say, for example, that there is an affordance to aggress every time one passes another person on the street. We might just as well say that there is an affordance to compliment every stranger about the shape of their ear lobes. We posit emotional affordances to explain emotional motivation and behavior, so we must avoid adopting a conception of affordances that saddles us with these trivial entities. In fact, it doesn't even seem plausible to suppose that there is an aggression affordance when one is insulted. In the right mood, one might be totally unbothered by biting words. In such cases, one's body might remain docile in ways that are neither conducive to aggressing nor motivated to aggress. The emotional affordance – they entity with real explanatory value – seems to arise only when an actual transformation in the body takes place. The transformation enables the body to act, and it brings with it a motivation pull. In contrast, Gibsonian affordances are dispositional properties which need not have any command over us. Griffiths and Scarantino implicitly recognize this. In their fully developed theory, they note that emotional affordances can change dynamically in the course of a social interaction. But this feature of their account marks a contrast with Gibson's idea that affordances are environmental (or relational) invariants. Without further explanation, it is unclear whether affordances can play the role required by their theory.

There is also a further worry, which concerns an assumption that Griffiths and Scarantino share with most other approaches to emotional content. They think emotions

are representational; emotions represent affordances. As we have noted, embodied emotional features facilitate actualizing these affordances, so there is a non-arbitrary connection between emotion and embodied responses. However, on their theory, that connection is still contingent. It is possible to represent the relevant affordances without any bodily change, and they do not explain why alterations in bodily features should change our representations of those affordances. In these ways, their affordance theory resembles other members of the appraisal theory family.

In summary, the affordance theory strikes us as a promising approach to the content of emotions, because it makes features of the body relevant to content determination. But the theory also faces objections about the way that it characterizes affordances, and about the contingent connection between representations and embodied features.

3. From Affordances to Enaction

We have raised concerns about some of the prevailing approaches to the content of emotions. Here is a brief summary. No-content views capture the intuitions that embodied responses are unlike conventional representations (in ways we will elaborate on below), but they fail to fully explain the role that emotions play in modulating motivation and behavior. Hybrid views have difficulty justifying the union of cerebral and embodied elements, and do not make much use of the later. The embodied appraisal theory puts meaning into the embodied response, but it doesn't adequately explain the bodily aspects of meaning and relies on controversial semantic theories. Affordances address these last worries nicely, but they ultimately feature some disadvantages of hybrid view. Since they still consider emotions to be representational states, they do not fully capture the non-representational contributions of emotional embodiment. It would be good to find another approach.

One idea that has been emerging in the recent literature may fit the bill. It is the enactive approach to cognition (Colombetti & Thompson, 2008; Colombetti, 2014; Di Paolo & Jaegher, 2010). Here we want to sketch a version of the view, along with some of its implications. The details of our approach will differ a bit from the existing enactive views. For example, we will not be endorsing extended cognition, emergence, or phenomenological methodology. In these respects, our account is a limited form of enactivism. In addition, we reject the thesis that enactivism is the best approach to cognition more generally; we think it is especially suitable for the emotions.

The enactive approach that we want to recommend takes inspiration from Francisco Varela and his colleagues, since we think he provides an attractive way to move beyond the affordance theory. Our focus will be on Varela's account of how minds make meaning (Varela, 1991; Varela et al., 1991). Here the term "enaction" gives a good indication of the view. Two things need to be italicized. First, to "enact" is, for Varela, "to bring forth." Enaction brings things into existence. In this sense, enaction is ontological; it creates new things. Second, enaction involves action. We bring things into existence by doing something with our bodies. It is an embodied theory of cognition because mental activity is grounded in physical activity.

The notion of “bringing forth” is meant to contrast with standard accounts of representation. A key feature of standard representational approaches is that they assume that the world is pre-given, that there are mind-independent entities, category boundaries, properties, and relations waiting to be re-presented by symbols or models inside the mind. Enaction, in contrast, implies a kind of constructive process wherein an organism adds to its environment things that weren’t there before mental processes kicked in. Theories of embodiment help us develop an account of how this construction takes place. How can we alter what exists in our world? According to Varela, we can do so through embodied activity.

We think these ideas lend themselves to explaining the emotions. First of all, we think emotions are embodied, in that they essentially involve bodily states, as well as inner processes that regulate these states. We also think the bodily states in question are related to action. Each emotional episode involves a pattern of bodily change that is conducive to specific kinds of action: fleeing, aggressing, withdrawing, and so on (Prinz, 2014). We also think that emotional content is difficult to characterize in terms of a pre-given world. Against the idea that minds bring forth reality, critics argue that some features of the world exist independently of particular cognizers, and, against embodiment, critics offer examples of mental processes that can be carried out without specific changes in the body. Neither line of criticism is especially compelling in the case of emotions, however, which is our focus here. We think evidence supports the view that emotions are embodied (though, for reasons of space, we cannot discuss the evidence here), and we have raised doubts about the idea that emotions represent pre-given features of the world.

Putting these two features of enaction together—embodiment and bringing-forth—we can now say, in broad strokes, why we think an enactive theory of emotion may outperform rivaling embodied theories. Like no-content views, the enactive approach rejects traditional theories of representation, but, unlike them, it asserts that embodied activities are, in some sense yet to be specified, meaning-giving. Unlike hybrid views, enaction does not posit cool cognitive states to explain emotional meaning, but rather locates meaning in embodied activity. For reasons that will become clearer shortly, the enactive view does so by making reference to actual bodily changes, and is thus more cohesive than the embodied appraisal view, which calls on covariation with things in the environment.

The enactive view most resembles an affordance view. As the term is often used, affordances are things that we perceive in virtue of aspects of our bodies. We see the graspability of a hammer because its handle has proportions that could be easily grasped by hands like ours. On most affordance theories, the perception of an affordance is occasioned by an actual motor plan: when a hammer is presented, we prepare to grasp. This looks like a case where meaning is created by physical activity. As we suggested, however, affordances are also often characterized as standing features of objects, which they have even when the affordances have not been noticed. Hammers are graspable even when unseen. For this reason, it is misleading to say affordances are created by our emotions, at least if we use the visual examples as an analogy. The affordance theory does much to improve on its

rivals, by relating meaning more intimately to embodiment, but it faces this technical hurdle.

The enactive view promises to overcome that hurdle. Enactive content is said to be brought forth by activity. So it is not the kind of content that is in place prior to the emotion. But what might that content be, if it is not representational? This is a question that is difficult to answer without discussing some specific examples. We will do that in the next section. But, to get started, we want to offer a schematic suggestion of how emotions might bring forth meaning. Affordances are possibilities for action but they are possibilities that are already in place. Emotions create new possibilities for action. The embodiment of an emotion makes new actions possible because it places the body into a configuration where it can perform certain actions more easily than it could have before. One might think this is just a quantitative difference: emotions make action easier. We think it is more than that. Prior to an emotion, there is a sense in which any action compatible with our physical constitution is possible. So that claim that a given action could take place is, to that degree, trivial. When an action is potentiated by an actual bodily change, certain courses of action go from being mere possibilities to being something akin to dynamic attractors. Epistemically, a potentiated action becomes a salient option, and, motivationally, a potentiated action has a kind of command over it. When the body prepares to act, we are *eo ipso* pulled toward certain lines of behavior. Motivation is immanent: it is not merely an injunction to act, through its motivational role, but a nascent kind of acting, as a constitutively embodied state—our bodies push and pull us.

All this suggests an important contrast between emotional content and Gibsonian affordances. In perception, affordances are properties that stand in place regardless of whether they are noticed. Affordances are, thus, dispositional, and, consequently, they are not motivating. Emotional content differs on both of these dimensions. This is the difference between possibility and potentiation: Gibsonian affordances merely require possible action, while emotional content requires nascent action. Emotional content can still be expressed using the language of affordances provided this difference with the perceptual case is made explicit: the affordances associated with emotions are state-dependent (they typically arise only once the emotion has been initiated), and imperatival (they motivate action). In contrast, the affordances in Gibson's scheme are state-independent and inert. Thus, we embrace Griffiths and Scarantino's appeal to affordances, but find it useful to develop the view by making these differences explicit. The enactive approach gives us the tools for doing precisely this.

State-dependent, imperatival affordances have the attractive feature of being bound up with embodiment. These affordances arise only when the body undergoes perturbations, and they motivate because the perturbations push and pull us. This offers a more intimate link between content and bodily change than can be found in traditional embodied theories. It also implies that emotional meaning is not given in the world, but rather constructed, which is a satisfying result.

With this schematic characterization in place, we will now consider a range of different emotions, both social and non-social, to illustrate how the enactive view works in greater detail. We will then consider some implications.

4. Putting the Enactive Approach to Work

Emotions are often put into two classes, social and nonsocial. Here we want to give illustrations of both. Griffiths (1997) has argued that social and nonsocial emotions do not form a natural kind, and the affordance theory presented by Griffiths and Scarantino takes social emotions as their primary focus. We think that both social and nonsocial emotions are amenable to a unified analysis on the enactive view, at least with respect to content. We will not attempt to show that all emotions can be explained on our account, but by giving examples in both categories, we hope to illustrate its breath.

We will begin with an example that illustrates the state-dependent nature of emotion enaction in a nonsocial context. A swimmer far from shore begins to struggle with a strong undertow, and realizes that it will take every bit of strength she has to make it back. In virtue of entering a state of fear, her body mobilizes more forcefully to support strenuous action, helping her return safely. This bodily change has important effects. Prior to the fear, the ocean seemed like a place to enjoy a swim, not a site of escape. The bodily response makes escape salient by actually compelling a flight response. Moreover, the physical changes actually enable the swimmer to traverse waters that would have been prohibitively strenuous before her fear set in. So both what is possible for her and what she is motivated to do have changed. In that respect it seems like her fear altered affordances for her. However, the ordinary concept of affordance does not permit this sort of state-dependent change. Gibsonian affordances are stable possibilities of action. Only through long-term training could she make an enduring change in what waters can be traversed, and, even if our swimmer has undergone such training, the ocean would afford recreational swimming to her, not escape. Fear increases her powers and changes the action-type that presents itself to her. It is a temporary solution to a temporary problem, which changes her relationship to the environment.

In our second example, a dancer is thrilled when the DJ plays one of his favorite songs. Like fear, joy mobilizes bodily resources, so he can dance harder and longer. Other types of action are available to him: he could get a drink, take a break, or even go home and get some sleep. Instead, his joyful bodily activation has an imperative force that makes him dance to that song. It gives him extra stamina, compels him to carry on, and thereby makes dancing a compelling choice. Without joy, dancing would not seem feasible any longer.

Fear mobilizes resources to make dangers more escapable, and joy mobilizes resources to make opportunities more achievable. Sometimes, however, there is not much to be done, either because nothing is wanting in the case of relaxed pleasure, or, in the case of sadness, because one has suffered a loss that cannot be repaired. The adaptive response in the latter case is to avoid spending bodily resources in a futile attempt to achieve what is (temporarily at least) out of reach. When sad we reserve bodily resources in a way that changes affordances. Our bodies shift into a mode that makes any

action more effortful, and that functions as an imperative to disengage, withdraw, or remain lethargic.

The enactive view can also be applied to social emotions. Here we take inspiration from Griffiths and Scarantino (2009), though we also want to emphasize the imperatival, state-dependent nature of emotional affordances in these cases. We can start with the example of joy in athletes. Joy often arises in nonsocial contexts, but, as Griffiths and Scarantino point out, there is also a social form of joy—a joy that is dependent of doing something that is valued in the eyes of others. We think it useful to mark this distinction by using the word “pride” for the social case. Griffiths and Scarantino also point to research showing that victorious athletes smile only when observed (see Fernandez-Dols and Ruiz-Belda 1997). An athlete’s smile makes it clear to her audience that she is happy with her accomplishments. By signaling this, she is mobilizing resources that can accomplish significant social benefits. Her smile transforms neutral observers into potential allies, since smiles are contagious, and we all want to collaborate with people who are on a path to success. Social emotions are not mere representations of social relations; they bring forth new social affordances, both for the one having the emotion and those with whom she interacts. The athlete’s emotion creates a new relation with her audience that allows bond-building. Forging social connections becomes an imperative, and the bodily changes, including the expressive component, contribute to that possibility.

Displays of anger follow a similar pattern. They signal preparation to aggressively further interests or values, which rallies support from those whose interests or values align, and discourages opposition from others. Social displays create new social realities, and they do so by credibly publicizing the features that are familiar from the non-social cases: the creation of state-dependent, imperative affordances through bodily mobilization. Anger mobilizes the body for aggressive action. This can transform affordances regarding the ones who provoked a state of anger. It would be absurd to suppose that all people at all times afford being aggressed against. Though physically possible, we rarely feel motivated to attack others. It doesn’t even make sense to say that people afford aggression when they act offensively; some react to offenses with aggression, but others with avoidance, submission, dependency, concern, disarming affection, bemused indifference, and so on. Anger does not pick up on the action that was demanded by the offense, but rather creates the demand; it temporarily transforms those who provoked anger, and potentially innocent others whom one happens to encounter while angry, into potential targets of aggression. Expression signals this, and action potentiation motivates it, thus creating new social relations that were not in place before the emotion.

These examples are just rough sketches to illustrate how the idea of state-dependent imperatival affordances can apply to nonsocial and social emotions. Our goal with these cases is not to provide precise analyses, but rather to give an idea of how enactive content works. Instead of representing core relational themes such as danger, loss, and offensiveness, emotions bring in new possibilities of action, and introduce a kind of to-be-doneness that temporarily alters affordance space.

5. Implications of Emotional Enactivism

Having presented some examples, we now want to consider some implications of the enactive approach. We will begin by indicating how coming to this view relates to our own earlier commitments in emotion theory. In so doing, we hope to illustrate what's at stake for those who are already attracted to embodied approaches to emotions. Then we will consider some implications that bear on widely shared commitments in the literature. We think an enactive approach to emotional content may require the reconsideration of some deeply held views. Finally, we take up the issue of emotions and ontology. This is a central feature of our account, which we want to bring out more explicitly.

In previous work, both of us have been committed to embodied approaches to emotions. As noted, one of us has defended an embodied appraisal view, which treats emotions as perceptions of bodily states that carry information about core relational themes (Prinz, 2004). This account shares features with enactive approaches (see Colombetti & Thompson, 2008), but also faces objections (see above and Colombetti, 2014). If our enactivist approach is on the right tracks, the embodied appraisal theory must be revised. First, it is wrong about the content of emotions: rather than representing core relational themes, such as danger and loss, emotions represent state-dependent, imperatival affordances. In principle, one could define core relational themes broadly enough to accommodate this revision. A core relational theme is an organism-environment relation that bears of well-being. State-dependent, imperatival affordances satisfy this definition. Still, loss and danger seem to be response-independent properties; one can be in danger, for example, without realizing it. One of us (Prinz 2004) did embrace this point, and denied that emotional content is response-dependent. We now think this is a mistake, since emotional content depends on our responses. In fact, we think this dependency goes even deeper than the dependency found in standard secondary qualities, such as colors. If Locke (1979) is right, colors are powers to cause sensations; and such powers exist in things even when they are not having their effects. Emotional content is not dependent on our responses in this dispositional sense; it depends on an organism actually being in an emotional state (state-dependence is stronger than response-dependence). Moreover, the content is imperatival. It has a kind of to-be-doneness, which standard secondary qualities lack. One can define terms such as danger and loss to reflect these two features. One can stipulate, for example, that something counts as a loss only when it is perceived affectively in a way that motivates certain kind of coping strategies. So the revised account can still be characterized as an embodied appraisal view, but these differences in content are significant.

Another departure from Prinz (2004) concerns the correlates of emotional states. In that earlier work, emotions were said to supervene on brain states in somatic input systems. The peripheral nervous system and body were excluded. Enactivists like to argue for a stronger form of embodiment, according to which "mental states" can include actual changes in the body. We are inclined to agree with this. Drawing a principled distinction between brain and body is difficult and tends to promote a kind of dualistic thinking. Also, we explained the motivational aspect of

emotions by saying that the body is engaged in activities when we are in emotional states. Emotions are a kind of doing. This idea is easiest to capture if we treat the body as part of the emotion. That also allows us to say, in social cases, that others can directly perceive the emotions we are in, which is one way in which emotions literally change extant relations in the world. One could stipulate that emotions are internal, but we think that would be both contrived and unhelpful. A theory of how emotions work, and how they come to have their content, is easiest to state if emotions include both internal and external bodily changes.

This brings us to the prior work of Shargel. Shargel (2014) picks up on the ideas in William James's contemporary, Carl Lange. James and Lange are said to have innovated the embodied approach to emotions independently around the same time (James, 1884; Lange, 1885), hence the popular label, the James-Lange theory. But Shargel observes a subtle difference; James says emotions are perceptions of bodily changes and Lange seems to suggest that emotions encompass actual changes in the body. The enactive approach defended here fits nicely with that Langan thesis, which Shargel also endorses. In following Lange, Shargel (2015) also raises doubts about the idea that emotions are representational states. Below, we will describe a way in which our enactive account may qualify as non-representational. At the same time, the enactive account does commit to the view that emotions have content. This, and the specific content proposed, require extending the views expressed in Lange and endorsed by Shargel.

Before we address the difference between content and representation, we want to draw attention to a few features of enactive content that might require revision in many emotion theories, whether embodied or otherwise. First, the enactive view allows a more parsimonious view of what Griffiths (2004) calls Machiavellian emotions. Theorists in the transactional tradition have emphasized cases where emotions serve as strategic moves in an ongoing negotiation. Russell & Fernández-Dols's (1997) athlete example is relevant here, or Hinde's (1985) example of birds that display anger but quickly flee. Griffiths and Scarantino (2009) argue that we cannot understand these cases without locating each emotional response in the context of the ongoing, open-ended transaction. However, these cases complicate representational accounts of emotional content, even for their own affordance view. The agent enters the emotion that is the best next move for advancing that agent's interests, not the emotion that best represents some features of reality. As Griffiths (2004) acknowledges, this makes the emotion seem insincere. After all, if the emotion is sincere but still misrepresents reality, then the agent's strategic move to enter that emotional state requires self-delusion. Transactional emotions, for the representationalist, are therefore either Machiavellian or Orwellian.

The enactive view allows us to avoid this dilemma. In the enactive view, the content of an emotion is simply a function of the affordances it enacts. Emotions produce the same affordances as usual in the transactional cases, plus additional social affordances that are determined by the specific circumstances of the transaction. Hinde's angry bird really is mobilized to fight – there is nothing false about the mobilization. The goal behind the mobilization is to intimidate the rival,

affording greater access to valuable resources. However, there is a big difference between mobilizing to fight in order to alter social affordances and actually initiating a fight with a stronger rival. If the attempt to intimidate fails, it is only sensible to enter a state of fear, redeploying resources to afford more effective flight. On our view, it is sensible to alter emotional states as a negotiation develops, because different stages of the negotiation require different affordances. We can explain this progression without implying insincerity or self-delusion.

With our revised view of emotional content, it is necessary to reconsider associated properties such as intensity. Within a representational framework, there are two natural ways to conceive of the strength of an emotion: as the degree of credence in the representation, or as the magnitude of the features perceived. The former seems more appropriate because it corresponds more closely to the strength of other states, such as belief, with a mind-to-world direction of fit. A strong belief is a belief we hold with greater credence. However, strong credence in a minor danger does not explain the occurrence of a strong fear. The enactive theory of emotions has a very simple explanation of intensity. Emotions modify state-dependent, imperatival affordances, so a strong emotion modifies affordances more strongly. Strong fear mobilizes the body more thoroughly to enhance escape, which leads to a greater imperative to escape.

It is commonplace to say that emotions have two kinds of objects: formal and particular (Kenny 1963, de Sousa 1987). The formal object is, roughly, the type of self-world relationship that any token of a given emotion type would represent (for example, anger might concern offensiveness), and particular object is the specific thing towards which a given emotion is directed (for example, anger about a new government policy towards immigrants). On the enactive view, this distinction may break down to some degree. Once the body is engaged, it is liable to create new relations to a wide range of things that an organism happens to encounter. For example, when something makes a person angry, the anger tends to leak, and the person may lash out at those who did not cause the anger in the first place. Plus, in expressing anger, a person creates a palpable impression of hostility that can be picked up by others, temporarily altering social relations.

Such diffusion could be described as a kind of mistake, and in some cases it does harm—a person might regret an aggressive response to someone who did not cause the initial state of anger. But it is also a pervasive aspect of how emotions work, and the broad relational changes that take place are perfectly real, and demand explanation. Shargel (2015) uses this observation to challenge the claim that emotions have intentional content. The enactive approach suggests another way to put the point: emotions usually can't be confined to having precisely circumscribed particular objects. They have a kind of content that spreads, because the affordances themselves spread. The content doesn't impact everything (anger will not be extended to things that are too distant in space and time), but it impacts more than the objects that elicit the emotion. In fact, there are good reasons for this content to spread. When anger alters our bodily mobilization the effects on the efficacy of our actions aren't limited to any particular object. Imperatival content should have the same broad scope.

Just as we vary emotional states in sync with changes in social dynamics, as in our discussion of transactional cases above, we manage the scope of our emotional responses in accordance with social norms. Although our enactive theory does not feature emotions with particular intentional objects, our social norms sometimes do. In order to comply with these norms we often attribute an intentional object to our emotions. When our anger threatens to violate social norms by impacting someone who was not responsible for making us angry, we respond by selectively suppressing the imperatival character of the emotion so that it does not impact that person. This attribution does not reflect representational content, but it has instrumental value in such cases. We think that emotion theory may need to abandon or revise the notion of particular objects to reflect these facts about how emotions operate (cf. Shargel, 2015).

This brings us to another issue, which is the rationality of emotions. On many theories, emotions have a conventional kind of content, which can be assessed for something like truth or fit. Emotions aim to represent relations with the world, and they may do so accurately or inaccurately. Emotions can be criticized as unwarranted or irrational, when they arise in situations that they fail to fit. The enactive account is consistent with the possibility that emotions may go wrong, but there is a way in which it also invites rethinking about these issues. Consider fear. Fear, we have said, brings an imperatival affordance into being—something that might be glossed in terms of escapability together with an impulse to flight. Along with standard theories, we grant that such a state might arise under conditions that, objectively speaking, do not warrant flight. The enactive account can borrow from other approaches when trying to articulate grounds for saying that certain emotional episodes are inapt. At the same time, there is a way in which enactive content is self-fulfilling. The property of escapability is brought forth by the emotion. It is really there. If one has a fear response to something benign, such as a harmless snake, the emotion will still cause escapability to be instantiated in the situation. In this sense, the emotion is not misrepresenting anything. It can't misrepresent anything. It is bringing an action-related property into existence. One can be criticized for false beliefs about the dangerousness of the snake, or for acting in ways that are unnecessary, but the emotion itself does not misrepresent. This differs from standard theories, which suppose that fear represents something like danger. In addition, the enactive theory may introduce the possibility of another kind of irrationality, which relates more directly to action. In some cases emotions potentiate behaviors that we cannot carry out. Fear when constrained would be an example. In such a case, standard theories would say the fear is perfectly rational (chances are a person is danger if constrained). But constraints prevent fleability, so an emotion that potentiates flight cannot fully succeed in bringing that property into being. This, for us, would be a kind of misfire—a failure to enact a property. We think such misfires deserve consideration in discussions of how emotions can go wrong. We will not develop the point here, but flag it as a direction for future work.

Now we come to the issue of representation. The originators of the enactive approach to cognition make the provocative claim that mental states are not representational (Varela et al., 1991). Most cognitive scientists embrace a

representational theory of the mind, according to which behavior is underwritten by mental representations—inner states that aim to represent things, features, events, or states of affairs in the world. The wholesale rejection of mental representation is a radical view, and it is not one that we will endorse. But the view does have plausibility when applied to emotions. If the enactive account of emotional content is right, then emotions are not in the business of informing organisms about external features of the world. They do not tell us about something that is already there, prior to the emotion itself. Rather, they bring something into being. This bringing forth can be regarded as a kind of content: that which is brought forth is a kind of embodied interpretation. When a body prepares for flight in response to a predator, it presents the predator as something that demands escape, and had it prepared, instead, for aggression, it would represent the predator as demanding confrontation. Like other kinds of mental content, enactive content also drives action and decision-making, and can, thus, play a role in explaining behavior. The term ‘content’ here means something like significance for the organism. The term representation, in contrast, often carries the further assumption that there is something out there in the world that can be re-presented by the organism. On this use, the reference relation is like pointing or labeling; it implies a fixed and independent object. Enactive content is not like this. It is more like making or doing. It creates something that wasn’t there to begin with. This can be expressed by saying that emotions are not representations (or, more conservatively, they differ from how mental representations are standardly understood).

Much more can be said about the topic of representation. For instance, it raises questions about whether emotions have truth conditions and questions about the content of judgments (e.g., moral judgments or aesthetic judgments) that may have emotions as component parts. We will not take up these issues here, but we do want to end by drawing attention to a related issue. The contrast between representing and making underscores the point that emotions bring properties into being. This has been a repeated refrain in our discussion, but our emphasis so far has been on the semantic implications. The notion of enacting or bringing-forth also bears on questions of ontology. There are two ontological implications that we would like to emphasize. First, whereas representational mental states are said to register properties that already exist, enactive content is brought into being. Emotions are, thus, adding to the furniture of the world. Without emotions, the properties that they enact would not exist. This is an interesting case where minds construct reality. Second, the kinds of properties that emotions bring into being are quite unusual. As noted, they are neither primary qualities nor secondary qualities, as these terms are usually understood. Nor are they mere affordances, as that term has been used in accounts of perceptual content. Their imperatival character means that they are properties that have a built-in to-be-doneness. This can be thought of as a kind of normativity. It may be that other mental states also enact such properties (pains, desire, and evaluative attitudes, for example). It may also be these other states enact such properties only because they constitutively involve emotions. We need not settle that here. The crucial point in that enaction allows us to articulate the existence of a kind of property with interesting characteristics.

These properties play an important role for many organisms, and they warrant philosophical reflection. The enactive theory, unlike other accounts of emotional content, draws attention to this aspect of ontology, and that may be its greatest source of interest. On the enactive theory emotions furnish the world with normative properties. This ontological implication is easily missed on other theories of emotional content.

We leave the elaboration of these implications to future occasions. Our goal here has been to motivate and articulate an enactive approach to emotion content. We think this approach has advantages over rival theories, and some surprising implications—psychological, semantic, and ontological. Our enactive theory leaves out aspects of enactivism that have been emphasized by some other authors, and we stop short of the view that all cognition is enactive. But our theory builds on this tradition, which is sometimes dismissed as excessively radical. We have tried to show that aspects of enactivism lend themselves to emotions. Taking enactivism seriously may deepen our understanding of what emotions are and why they matter.

Acknowledgements

We would like to thank Fabrice Teroni and Hichem Naar for extremely detailed and helpful comments.

References

- Aristotle (1987). *De Anima*, trans. H. Lawson-Tancred. London: Penguin Classics.
- Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press.
- Colombetti, G., & Thompson, E. (2008). The feeling body: Towards an enactive approach to emotion. In W. F. Overton, U. Müller, and J. L. Newman (eds.), *Developmental Perspectives on Embodiment and Consciousness* (pp. 45–68). : Erlbaum.
- Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.
- Descartes, R. (1989). *The Passions of the Soul*, trans. S. H. Vonn. Cambridge: Hackett Publishing.
- de Sousa, R. (1987). *The Rationality of Emotion*. Cambridge, MA: MIT Press.
- Di Paolo, E.A., M. Rohde, & De Jaegher, H. (2010). Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne, & E.D. Di Paolo (eds.), *Enaction: Toward a New Paradigm for Cognitive Science* (pp. 33–87). Cambridge, MA: MIT Press.
- Dretske, F. (1981). *Knowledge and the Flow of Information*, Cambridge: MIT Press.
- Fernández-Dols, J.M., & Ruiz-Belda, M.A. (1997). Spontaneous facial behavior during intense emotional episodes: Artistic truth and optical truth. In J.A. Russell and J.M. Fernández-Dols (eds.), *The Psychology of Facial Expression* (pp. 255-274). Cambridge: Cambridge University Press, 1997.

- Fodor, J.A. (1990). *A theory of content and other essays*. Cambridge, MA: MIT Press.☐
- Frijda, N.H. (1986). *The Emotions*. Cambridge: Cambridge University Press.
- Hutcheson, F. (2003). *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*. Indianapolis, IN: Liberty Fund, Inc.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Grèzes J., Tucker M., Armony J., Ellis R., Passingham R. E. (2003). Objects automatically potentiate action: an fMRI study of implicit processing. *European Journal of Neuroscience*, 17, 2735–2740
- Griffiths, P. E. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Griffiths, P. E. (2004). Towards a Machiavellian theory of emotional appraisal. In P. Cruse & D. Evans (eds.), *Emotion, Evolution and Rationality* (pp. 89-105). Oxford: Oxford University Press.
- Griffiths, P., Scarantino, A. (2009). Emotions in the wild: The situated perspective on emotion. In P. Robbins & M. Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 437-453). Cambridge: Cambridge University Press.
- Hinde, R. A. (1985). Expression and Negotiation. In G. Zivin (ed.), *The Development of Expressive Behavior* (pp. 103-116). New York: Academic Press.
- Kenny, A. (1963). *Action, emotion and will*. London: Routledge & Kegan Paul.
- Lazarus, R. S. (1991). *Emotion and Adaptation*. New York: Oxford University Press.
- Locke, J. (1979). *An Essay Concerning Human Understanding*, ed. P. H. Niddich. Oxford: Oxford University Press.
- Prinz, J.J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- Scarantino, A. (2004). Affordances Explained. *Philosophy of Science*, 71, 949-961.
- Shargel, D. (2014). *Constitutively embodied emotions* (Doctoral dissertation). Retrieved from http://academicworks.cuny.edu/gc_etds/472
- Shargel, D. (2015). Emotions without objects. *Biology and Philosophy*, 30(6), 831-844.
- Varela, F.J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Varela, F.J. (1991). Whence the origin of perception? A cartography of current ideas. In F. Varela & J.P.Dupuy (eds.), *Understanding Origins: Contemporary Ideas on the Origin of Life, Mind and Society* (pp.235-265). Boston: Kluwer.