PHILOSOPHICAL PSYCHOLOGY, 2017
https://doi.org/10.1080/09515089.2017.1408063

Routledge
Taylor & Francis Group

The elements of emotion

Chad Brockman

Independent Researcher, Englewood, CO, USA

ABSTRACT

I join the growing ranks of theorists who reject the terms of traditional debates about the nature of emotion, debates that have long focused on the question of whether emotions should be understood as either cognitive or somatic kinds of states. Here, I propose and defend a way of incorporating both into a single theory, which I label the "Integrated Representational Theory" of emotion (IRT). In Section 2 I begin to construct the theory, defining and explaining emotions in terms of three pieces of content: representations of (1) the emoter's body, (2) something in the world, and (3) a relevance relation between the objects of these first two pieces of content. I describe four general advantages I think the IRT offers. Finally, in Section 3, I elucidate and defend my account by contrasting it with another, similar proposal: Barlassina and Newen's Impure Somatic Theory. In so doing, I explain two additional advantages of my view: first, it supports a unified explanation of all types of emotional response; and second, it offers the best framework for explaining how the representational contents of an emotion are integrated.

ARTICLE HISTORY
Received 2 March 2017
Accepted 31 August 2017

KEYWORDS
Emotion; integration; intentional; perception; phenomenology; representation

1. Introduction

It is widely acknowledged that emotions involve two kinds of mental states the emoter has: representation of the state of her body and representation of something in the world to which the emotion is (in some sense) directed. However, until recently, debates about the nature of emotion have concerned which of these two kinds of representation is their essential and explanatory element. Cognitive theorists have argued that emotions are best explained solely in terms of beliefs, desires, judgments, and so on, which constitute them, while somatic theorists have argued that they are best explained solely in terms of bodily processes they involve.

I am one of a growing group of philosophers who reject both of the traditional, single-dimension theories of emotion. I believe the problems faced by each of the two traditional theories requires us to try a different approach. Given that

CONTACT Chad Brockman cbrockman71@gmail.com

emotions involve representing both the body and the world, the obvious thing to try is to see if we can combine the core elements of both theories into one explanatory framework. So this is what I undertake here. In Section 2, I propose a theory that explains emotions in terms of three pieces of representational content in the emoter's mind: (1) representation of the state of her body, (2) representation of something in the world, and (3) representation of some relation of relevance between these first two pieces of content. Since the third piece of content integrates the contents of the first two, I label my proposed theory the Integrated Representation Theory (IRT). I end Section 2 by sketching the construction of my proposed theory and outlining some of the general advantages I think it offers.[1]

Finally, I evaluate my proposal by contrasting it with Barlassina and Newen's Impure Somatic Theory of emotion (IST). There is substantial overlap between my analysis and theirs, but I discuss some advantages of my proposed theory over the IST. Specifically, I argue, the IRT offers a unified explanation of all types of emotion, whereas Barlassina and Newen's approach requires different kinds of explanation for biologically basic vs. cognitively complex emotional states. Also, I argue, my view offers a better way of explaining how the first two pieces of representational content get integrated, by positing a third piece of content that integrates them. I conclude by describing these two additional advantages of my proposed theory.

2. A compound theory of emotion

I believe that each of the two traditional Single-Dimension theories of emotion is inadequate. There is a large literature to review, and many issues to consider, before we would be entitled to this conclusion. Here, I will only make a few conclusory statements and set this background issue aside. Roughly, emotions cannot be identified with states consisting only of some combination of beliefs, judgments, or desires, because any or all of these states could be held unemotionally. If the body is not involved in the right way, the state will be better understood as a kind of thought rather than an emotion. On the other hand, emotions also cannot be identified with states consisting only of bodily processes or somatic states, because these processes could fail to have the right kinds of functional connections with the emoter's representation of her world. If cognitive processes are not involved in the right way, the state is at most a certain kind of bodily process, likely accompanied by a certain kind of bodily feeling, but not an emotion.

There have been formidable defenses of the two views I am rejecting against the kinds of claims I have just made about them, and I do not mean to imply that these debates have been resolved. But my review of this literature leads me to conclude that both cognitive and somatic theories suffer for lack of the other's explanatory resources, and that any adequate theory will have to combine their core elements in some way. That is, any adequate theory of the emotions will have to explain them in terms of emoters' responsiveness to both changes in the state of the body

and things or situations in the external world. The present issue is how we might be able to combine these elements in one theory. So I shall now begin to develop such a theory and describe some of the general advantages I take it to have.

2.1. The skeleton of an integrated representational theory of emotion (IRT)

I propose to define an emotion as a mental state in which the emoter represents three things: (1) occurrent changes in the state of her body, (2) something in the world, and (3) some relation of relevance between the objects of the first two representations. It is because the first two pieces of representational content are integrated in the third that I label this the *Integrated Representational Theory* of Emotion, or *IRT*. Of course, the conclusion that we need such an elaborate new theory presents me with a set of challenges that the two traditional, one-dimensional theories do not face: I must justify the inclusion of each of the three pieces of content and show how they add up to a unified explanation. I think these challenges can be met, but also that it is wise for us to be philosophically conservative as we undertake theory construction. Thus, as a matter of methodology I aim to add elements to the theory only as needed, and to resist unnecessary philosophical commitments.

2.1.1. First piece of content: Representation of occurrent changes in the state of the body

Jamesian theorists (1890/1950) have been correct to insist that emotions cannot be explained without reference to bodily changes. Merely believing that something is dangerous does not suffice for being afraid, for example; there must also be a representation of physiological disturbance. At the same time, while emotions necessarily involve representation of bodily changes, they are not reducible to such changes, or to the representation of such changes.

The first philosophical commitment I think we must make here is that the emoter's representations of the state of her body must meet the minimal conditions for intentionality. That is, they cannot be merely causally correlated with, or bear information about, actual changes in the body; it must be possible for the representations to represent non-veridically. Otherwise we would be committed to assuming that having an emotion necessarily involves representations that perfectly track occurrent changes in the state of the body. Assuming that perception is not taken to entail very much in the way of conceptualization or propositional form, I think it would be correct to regard the bodily representations in my proposed theory as perceptual, in roughly the same way that low-level registration by an organism of basic features of its environment are perceptual. And where there is perception there can also be misperception.

One of the most important innovations in Antonio Damasio's development of James's somatic theory (1890/1950) is his addition of mechanisms for representing emotion-related bodily changes *where those changes are not actually*

occurring (See Damasio, 1994, pp. 155–156). That is, in addition to the standard process of sensing bodily changes as they occur, Damasio argues, there must also be systems for generating representations *as if* those changes were occurring. In this way an emoter can undergo and experience similar emotions, whether or not her perception of changes in her body is veridical. Damasio gives several reasons for thinking there must be such "as if" representations of bodily changes. Some of these reasons are based on neuroscientific evidence suggesting that there are "neural devices that help us feel 'as if' we were having an emotional state" by activating body-mapping systems in the brain through causal processes that bypass the body (p. 155). Another reason to think there are these "as if" representations of affective bodily changes, Damasio thinks, is that these systems are needed to explain the high level of efficiency with which we respond to emotional stimuli and learn from them. That is, instead of always having to wait for an emotional response to be fully realized in a body state, and then do the work of processing information about the body's response, we are able to "concoct the fainter image of an 'emotional' body state, without having to reenact it in the body proper" (p. 155). In short, we would be far less able than we really are to be affected by what happens to us, to interpret its emotional significance, and to learn from our emotional experiences, if it were not the case that we could have non-veridical representations of emotion-related changes in our bodies. Thus, Damasio has provided reasons for thinking the contents of these states representing the body are intentional contents in the traditional sense.

A second required philosophical commitment will be to some constraints on the kinds of physiological changes the representation of which could serve as this piece of content. Bodily changes must be represented as part of having an emotion, I contend, but clearly not just any bodily changes will work. One could not be said to be sad about being rejected because one was rejected and happened to have a toothache at the time. Only certain kinds of bodily changes, and corresponding experiences, can constitute sadness, and toothaches are not generally among them. So how do we know which body states can be represented and thereby constitute a certain emotion?

Fortunately it is not necessary to dig too deep into this question here, because the principal competitors of my proposed theory have no advantage in this regard. Cognitive theories do not require emotions to be realized in the body in *any* particular way. So for cognitivists, very little turns on this question about constraints on physical realization. Conversely, showing how to specify constraints on the physiology of emotion is critical for purely somatic accounts like those of James and Prinz.[2] If emotions are identical with states of the body, or explainable solely in terms of body states, then the whole explanation turns on the issue of exactly which body states these are. For in that case it is not just a matter of specifying constraints; there will simply be a fact of the matter, for each emotional state, and each kind of emotion, as to which pattern of physiological changes is involved.

So the least we can say is that the physiological theorist has the heaviest burden of the theories I am considering here.

Finally, as with the other dual-dimension theories, we will need some way of explaining why some somatic representations can figure in emotions while others cannot—why crying can be constitutive of sadness, but toothaches cannot. This is an interesting set of questions, but, so far as I can tell, not important for comparing the theories I discuss here. Any methods available to Jamesians (1890/1950) for connecting emotions with distinct patterns of bodily response will also be available to a dual-dimension theory that subsumes their somatic theory.[3]

Notice an important formal similarity between the two philosophical commitments just discussed: in each case the representations in question are constrained in some respects and flexible in others. First, this part of an emotion's mental content is constrained in that it must represent changes in the body, but it is also flexible in that it can be generated either in response either to actual bodily changes or through processes of thought, imagination, or reinstatement that bypass the body. Second, it is constrained in that it must represent bodily changes that somehow fit the kind of emotion at issue; however, it is also flexible in that the possible patterns of emotion-constituting physiology need not match up one-to-one with the range of actual emotional phenomena. A single body state could figure in more than one emotion. By defining emotions in terms of complex intentional states instead of bodily sensations or perceptions, the IRT can incorporate somatic perceptions without reducing emotions to them.

2.1.2. Second piece of content: Representation of something in the world

Cognitive theorists have been correct in thinking that emotions cannot be explained without reference to the way emoters represent things in the world.[4] Merely being in a body state typical of anger does not suffice for being angry; there must also be a representation of something in the world that is represented as having some relevance to this body state.[5] At the same time, the emoter's representation of the world is not by itself enough to support the whole psychological explanation of an emotion.

Which philosophical commitments does this part of the IRT require? First, here again the representation of something outside the body must be intentional. It must be possible for an emoter to be frightened of an object he only thought he saw moving toward him, or angry about a movement he has misinterpreted as a gesture. We cannot build into the theory an assumption that the emoter correctly represents whatever it is he is emotional about. Many of the things that inspire our emotions are everyday objects of perception, and in those cases our representations must be robust in the usual ways with respect to possible misperception. Further, I assume it is possible for the external object of an emotion to be misconstrued, mereological, mythical, abstract, and so on, and in such cases it can be unclear exactly how to specify conditions of identity or veridicality. And this means we must, if we want to allow for all kinds of objects, leave open the

possibility of either non-veridical or imperfectly distinct representation of the things in the world that an emotion could be about.

The second and only other commitment I think we need here is to the possibility that the representation of the object is one that could possibly exist in the right kind of relevance relation to that of a corresponding body state. In a moment I will elucidate the nature of this relevance relation, and the kinds of content that can occupy it, when I move on to the third piece of content. For now, I claim only that, whichever external object an emoter represents as part of an emotion, in order for that object to figure in an explanation of that emotion it must be possible for it to be represented as having some connection of an appropriate sort with the corresponding representation of a body state. For reasons to be explained momentarily, I believe the burden of this requirement turns out to be light.

2.1.2. Third piece of content: Representation of a relevance relation between the objects of the first two representations

It is the inclusion of a third piece of content that is most distinctive of my proposed theory, because it specifies that the first two pieces must be not only co-present, and not only integrated in the sense of causally interacting, but actually represented in relation to each other within a distinct part of the state's representational content.[6]

I mean for the philosophical commitments required by this last piece of content to be limited in at least three ways. First, this piece of content is the only thing referred to by the term "integrated" in the theory's name, the only kind of integration the theory makes essential for emotion. Of course, we should expect integration at the level of content normally to correspond to neural-level integration, as well as integration within experience. That is, if we thought that interoceptive and exteroceptive information gets integrated at the content level, then presumably we would also expect to find corresponding causal interaction between areas of the brain that process somatic information and areas that process perceptions of external objects. Looking from the other direction, finding such neuroscientific evidence, some of which I will discuss below, would tend to support the IRT. However, I think it does no harm, and might promote my goals of philosophical conservatism and tidiness, to specify that, strictly speaking, our only *commitment* is to integration consisting of this connection between the first two representations in the third piece of content. The relevant kind of integration is content-level integration.

Second, there is little restriction here on the kinds of relations that could be represented as existing between the world and a state of the body—causal, conceptual, historical, rational, narrative, perhaps even stipulated—a huge variety of different types of relations could serve as relevance relations. For in general it is not the nature of this relation between the first two pieces of content that defines the state as an emotion and helps to explain it, according to the IRT, but only the presence of all three pieces and the representation of connections among them.

Emotions are representational states with just this sort of form. The reason we cannot presuppose that the connection between world and bodily response must be a specific kind is that the kinds of relevance relations that can always be forged through the productive processes of associative learning, discursive thought, and conscious action create immense flexibility in the processes by which different mental contents could get connected together. The evidence for this flexibility is that it is possible (albeit obviously unlikely) for a person to develop a fear of sandwiches (which, we can assume, have never harmed her), or to become angry at the number four, or to feel sad about winning the lottery. These examples show that things in the world and bodily responses can get connected in ways that do not reflect real causal or conceptual or rational relations between them. Still, in the emoter's mind, her emotion is in some sense constituted by a relevance relation that she represents between these objects.

Third, we cannot assume that this third representation necessarily involves a conceptual articulation of the nature of the connection between the external object and the body's response. For the theory will need to accommodate relatively primitive emotional responses in which the representation of stimulus and response could not possibly be conceptual. In some such cases, that connection might be little more than the registration and preservation in the brain of their temporal contiguity. (As we will see, evidence from neuroscience strongly suggests that the brain does register and use such cognitively thin neural-level associations between stimuli and bodily responses.) In such cases, the relevance of one represented object to the other might not be anything at the level of conceptualized causal or conceptual relations between them; it might be only a neural registration of their simultaneous presence at a certain temporal point in the brain's processes.

A commitment I think we *do* need to make here is (yet again) to treat this representation as intentional in the sense of possibly involving misrepresentation. The extreme variability I have described regarding ways of determining emotional relevance makes this point about intentionality tricky to explain, but I think we still must allow for the possibility that an emoter could have a genuine emotion, while misrepresenting the relevance of her bodily response to the thing in the world with which she associates it.

To illustrate, recall the example of the person who is afraid of sandwiches. To avoid questions about dispositional vs. occurrent emotions, let us imagine a specific situation in which this person sees a certain cheese sandwich and becomes afraid of it. By my supposition, she must represent the sandwich as having some relevance to her body's fearful response. And I claim that, if she does have all three representations—the fearful response, the sandwich, and the represented relevance relation—then she is indeed afraid of the sandwich. It is easy to dismiss the possibility that this relevance relation she represents could have anything to do with justification, since, for reasons most people find obvious, sandwiches generally aren't dangerous. They are not relevant to our concerns in that way. But fear conditioning is a very open-ended process, and there are many conceivable

ways that this person might have come to regard a sandwich as dangerous. Perhaps she nearly choked on a sandwich long ago. Perhaps she suspects that someone means to use the sandwich to poison her. Perhaps there is an evil Pavlovian in her life who has trained her to fear sandwiches. Or perhaps she did not expect the sandwich to be where she happened to look, its presence surprised her, and the whole experience ended up being a bit spooky. And so on.

Strictly speaking, one could argue that in some of these cases the emoter is misrepresenting the nature of the relevance relation between the sandwich and her body state, since she thinks the sandwich justifies the response when in fact it doesn't. But the connection need not be rational in order to serve the function of representing relevance under the IRT. It only needs to forge a represented connection in the emoter's mind. So however justified or unjustified this representation of relevance might be, if the person represents it then this could be our third piece of content and we would have to say that she is, whether rationally or not, afraid of the sandwich.

Another kind of case is perhaps a bit more troublesome. Several studies have shown it to be possible that, unbeknownst to the person in question, the state of her body might have been brought about by something other than the sandwich. For example, perhaps a terrifying image was flashed to her subliminally just before she saw the sandwich. If the terrifying image caused her fearful body state, but she instead associated it with the sandwich, then we might think this shows that her representation of the sandwich as having relevance to her body state was not veridical. That is, we might conclude that she has misrepresented the sandwich as relevant to her fear response, since in fact something else—the terrifying image—is the thing that actually *is* relevant.[7] There are problems with this line of reasoning, though. One is that multiple things could be relevant, and the fact that one is the most relevant does not prevent the other from being relevant as well. So this kind of argument would not support its conclusion. Another, deeper problem is that we as theorists have to base our judgment about the sandwich's relevance to the fear response on either *our* judgment of its relevance or *hers*. And since she is the one whose representations of both the terrifying image and the sandwich might have constituted her emotion, it will be her representation and not ours that decides the question. She represents her body's response as being about the sandwich, her behavior presumably reflects this fact, her next experience of seeing a sandwich will probably be causally connected with her current state, and we have no good reason to substitute our understanding of the object of her fear for her own. What her fear is about *for her*, what she represents as relevant to it, is what it is about, even if she is wrong about how it was caused.[8]

Still, even if we defer to the emoter's own representation of this relevance relation, there are still important ways in which it is reasonable to think that she has misrepresented things. For she has almost certainly represented the sandwich as causing, or at least inspiring, her fearful response, and if so, she is incorrect about this. She also almost certainly represents the sandwich as justifying her fearful

response, and thus as being relevant to her fear in that way. But this judgment would be based on her false belief that sandwiches are dangerous to her, and thus this basis for her judgment about its relevance would be incorrect as well. But in any case her representation of the sandwich's relevance need not be veridical to play its role in constituting her fear. At least this is what the IRT implies, and it seems to me correct.

2.2. Some general advantages of IRT

A key consequence of my making the same claim for all three elements of the theory is that we have a thoroughly representational account. Now the challenge will be to defend my conclusion that these three pieces of intentional content fit together in a way that will help us explain emotions.

2.2.1. Advantage 1: Homogeneous explanation in terms of intentional content

The theory I propose avoids an obvious worry that either cognitivists or Jamesians (1890/1950) would be likely to have. On the face of it, it seems that we are trying to combine two very different kinds of thing to be explained—feelings of bodily changes on the one hand, and perceptual or propositional representation of objects on the other—and these seem to call for different kinds of explanation. Indeed, I suspect that the quest for theoretical unity has driven many toward one or the other of the two single-dimension theories; seeing no way to combine both bodily feelings and perceptions of objects into one type of state, they make only one of the two explanatory and the other either irrelevant or accidental. One reason I maintain hope that there can be a unified explanation of these two elements is that they can both be explained as types of mental representation. Emoters *represent* things in the world that are the causes or objects of their emotions, and they *represent* the changes in their bodies.[9] At least part of feeling bodily changes involves representing them. So what we have are two kinds of perceptual representation, not perceptions on the one hand and feelings on the other. It remains a challenge to explain exactly how these two kinds of representation, which are likely to differ in important ways, are combined. But from the theorist's standpoint they are liquids in the same solution; we end up with a uniformly representational account.

2.2.2. Advantage 2: Unified explanation of emotions at all levels of cognitive sophistication

Paul Griffiths has long argued that the concept of emotion does not refer to a natural kind, because its reference is divided between different classes of phenomena unified by different "causal homeostatic mechanisms" (2004, p. 36). He argues that some of the things we call "emotions" are "affect programs"—biologically basic, reflex-like responses, such as surprise; others are cognitively complex states, such as being demoralized by the results of a recent election; and still others are socially learned "scripts" for courses of emotion-motivated behavior. I think we

can see how an IRT account cuts across the boundaries between these subtypes of emotion. Roughly, the world-directed content could be little more than low-level sensory registration of something in the environment, or it could be a thoroughly conceptualized or conscious representation. Likewise, the body-directed content could be little more than a low-level registration of changes in the body via an interoceptive system, or it could be a representation of a pattern of changes that is recognized or consciously thought about. Either way, the contents will fit into the IRT model, and thus, the theory will cover the various ways that emotions might be composed of both basic and cognitively complex representations.

In terms of causal analysis, the IRT is also silent on the question of whether emotion processes are causally "driven" primarily by physiological or cognitive processes. A state with the requisite structured content is an emotion, whatever the process by which it was generated. As a result of this flexibility, in the end we should be able to locate all genuine emotions somewhere on the same continua—from basic to cognitively complex, and from physiologically to cognitively "driven." Defined in this way, I think the different emotions might all form a distinct psychological kind. Whether we go on to treat this as a *natural* kind will depend, not on distinctions among the different patterns of emotional response, but instead on whether we are prepared to treat *any* folk psychological categories as natural kinds.

I continue my discussion of this advantage in connection with Barlassina and Newen's "impure somatic theory."

2.2.3. Advantage 3: IRT provides a plausible way of explaining emotions' phenomenology

Another advantage of my proposed theory is speculative, since it will depend on what we conclude about a different set of issues in the philosophy of mind. But consider the widely discussed doctrine of *intentionalism*, which is (roughly) the view that the phenomenal character of a perceptual state is determined, either partially or entirely, by representational content. Now recall that the IRT posits three pieces of representational content: representation (1) of the body, (2) of something in the world, and (3) of some relation of relevance between the two. Also recall that I have specified that all three pieces of content must be straightforwardly intentional. If we also assume this overall picture of how emotions are constructed, then, if we conclude that some version of intentionalism gives the correct explanation of the phenomenology of all this content, we will then have the beginnings of a broad general explanation of emotional feelings. Broad indeed, since it will have to accommodate three pieces of content, the objects of which can vary widely in their determination, their modes of representation, their degrees of intellectualization, and their causal relations to surrounding mental states. Given that its machinery accommodates this immense variability, an intentionalist account like this could go some way toward helping us explain the complex and variable phenomenology of emotional experience.

The most likely challenge here will concern (1), specifically regarding the premise that the phenomenology of bodily experience could admit of an intentionalist explanation. For it was long assumed that bodily processes, when they are perceived, are felt immediately, via a process not mediated by representations or thoughts or information processing. And our experience of having bodily feelings independently of our thoughts, coupled with the old dichotomy between feeling and thought, seems to give intuitive support to this assumption. However, recently Michael Tye and others have argued that even very "bodily" kinds of experiences, such as the experience of having pain in a specific part of one's body, are best understood as intentional states (See Tye, 1995). What pain experience represents, according to Tye, is that there is something wrong in a certain part of the body and this something is painful. What it is to be in pain, then, is to be in a mental state with that kind of representational content. If Tye is correct, pains can be located in different parts of the body, and can come in different "flavors"—stabbing, burning, stinging, and so on, yet all be explained by reference to the same type of intentional content.[10] That is, pain states have intentional content, and it is their having this content that explains how they feel.

Various philosophers have argued for versions of intentionalism about visual and other kinds of perceptual experience as well (See, e.g., Byrne, 2001). Many of those arguments start with a firmer foothold than the one about pains, since it is easier to argue that the phenomenology of, for example, visual experience depends, at least to a large extent, on the content of the experience—that is, on how the objects perceived are represented as being. Cognition is an obvious part of the process also, and cognitive phenomenology, the explanation of the feeling associated with cognitive states in virtue of their specific representational contents, has also been widely discussed.[11] It seems safe to say that, for just about any kind of mental state that has a distinct type of intentional content and a distinct type of feeling, a plausible argument can be made for explaining the feeling in terms of the content.

Suppose intentionalist arguments about the phenomenology of bodily and perceptual experience both succeed. Then under IRT, we would have intentionalist explanations of the phenomenology of the first two elements of emotions. From there it would not be much of a stretch to suppose that the third piece of content, which is the integration of the first two, also gives rise to a kind of experience that is explainable in terms of its representational content. My argument, based in part on the kind of neuroscientific evidence I will mention in the next subsection, and in part on arguments to be developed further in my discussion of Advantage 5 below, will be that emotions really do involve this type of integrated intentional content.

My conclusion at this stage is that if the IRT account gets the intentional structure of emotions right, then we will have the resources for a productive and suitably complex explanation of their phenomenology.

2.2.4. Advantage 4: Consistency with relevant neuroscientific evidence

Again, the first two types of representation specified by the IRT were representations of (1) the state of the body and (2) something in the external world. It is not controversial that there are brain systems devoted to representing these things. But several neuroscientists have also argued that there are probably also distinct systems in the brain that function to integrate these two streams of information, which is to say that there are also distinct brain systems that function to supply the third piece of content for the theory: (3) representation of a relevant connection between (1) and (2). Thus, for example, Damasio writes that "In order for us to feel a certain way about a person or an event, the brain must have a means to represent the causal link between the person or event and the body state …" and that "This sense of precise cause-and-effect may arise from activity in convergence zones that perform a mutual brokerage between body signals and signals about the entity causing the emotion" (1994, pp. 161–162). Gray and colleagues found that "neural activity in a bilateral region in the lateral PFC (Pre Frontal Cortex) depended conjointly and equally on the emotional and stimulus conditions: a crossover interaction with no main effects." Based on these results, they speculate that if the integrated signal from this region has a functional role, "emotion and cognition can conjointly and equally contribute to the control of thought, affect, and behavior" (Gray, Braver, & Raichle, 2002, pp. 4115–4120). Duncan and Feldman-Barrett (2008) also conclude that "widely distributed circuitry" functions to link information about the external environment with somatovisceral information, and that "the function of this circuitry is to link sensory information about a stimulus with a representation of how the stimulus affects the person's internal (somatovisceral) state" (p. 1118).

These converging results from neuroscience obviously merit wider and closer examination than I have given them. But it seems clear that the overall structure of the IRT explanation is at least consistent with current thinking by some neuroscientists about the corresponding overall functional architecture of relevant systems in the brain. For they posit functional systems and processes corresponding to all three pieces of content in the IRT. Most notably, they posit processes of convergence or integration of interoceptive and exteroceptive information at the neural level, the products of which likely underlie cognitive phenomena such as emotional learning and memory, attention-guidance, object-directed feeling, and so on.

3. Comparison with another dual-dimension theory, Barlassina and Newen's impure somatic theory (IST)

Barlassina and Newen (hereafter, B&N) propose a theory according to which emotions "are constituted by the integration of bodily perceptions with representations of external objects, events, or states of affairs" (2013, p. 1). Since their theory is in most ways structurally very similar to the one I propose, I will not give a general

summary of their view, but will instead get straight to the comparison of their proposal with mine.

The substantive advantages I claim for my account over B&N's are (1) that the IRT provides a general framework that supports the same kind of explanation for all types of emotion in terms of contents that can be at any level of cognitive sophistication, thus, avoiding the need to continue splitting the category of emotion into "basic" and "higher cognitive" sub-categories, and (2) that, whereas they say little about the kind of integration they have in mind, the IRT allows us to begin filling in crucial details of that explanation by specifying how the representations of external objects and bodily changes are combined in a distinct, third piece of representational content. I raise and discuss specific arguments by B&N as they become relevant to my own discussion.

3.1. Advantage 5 (supplementing advantage 2): One explanation for both basic and non-basic emotions

Richard Lazarus is a cognitive theorist who explains emotions in terms of the emoter's appraisal of her relationship to her environment. Each type of emotion corresponds to a specific type of emoter-environment relationship, which Lazarus labels an emotion-specific "core relational theme." For example, the core relational theme for the emotion of guilt, according to Lazarus, is "having transgressed a moral imperative" (1991, p. 122). To feel guilty, then, is to appraise your situation in the world as an instance of your having transgressed a moral imperative.

Jesse Prinz adapts Lazarus's notion of core relational themes to his own purely somatic theory. On Prinz's view, emotions are identical with somatic states, but what they *represent* are core relational themes. To explain how the representational content of a somatic state can be a core relational theme, rather than the perceived body state, Prinz appeals to Dretske's (1981) causal and teleological theory of reference, the relevant part of which says (roughly) that a state represents x if that state is reliably caused by x in virtue of having been set up to carry information about x. To continue with the example of guilt, the system that gives rise to guilt has been set up to be reliably triggered by the emoter's appraisal of her situation as an instance of her having transgressed a moral imperative. Thus, Prinz argues, when a person feels the physiological changes that constitute the emotion of guilt, she thereby represents the corresponding core relational theme.

I provide this background explanation because B&N retain key parts of it. In particular, they retain the notion that emotions represent core relational themes. However, Lazarus's account of cognitive appraisal judgments will not work for B&N, since the most evolutionarily basic emotional responses cannot require anything as cognitively sophisticated as an appraisal judgment about a core relational theme. Prinz's account of somatic states representing core relational themes by carrying information about them also will not work for B&N, since, by their own lights, the way a higher cognitive emotion represents a core relational theme is by

having it as the content of a propositional attitude, not as something implied by the emoter's somatic state. Despite these problems, B&N conclude that all emotions do in fact represent core relational themes. However, they argue—and this is the key point here—basic and higher cognitive emotions (hereafter, "non-basic emotions") represent core relational themes in different ways: something like Prinz's account explains how basic emotions represent core relational themes, while something like Lazarus's account explains how higher cognitive emotions represent them.

After summarizing all of this, I now conclude that we can and should dispense with most of it. B&N's view commits us to two philosophical assumptions that I find extravagant and unhelpful. First, the Prinzian premise that a basic emotion necessarily represents a core relational theme, despite there being no appraisal judgment (or belief, etc.) to the effect that the core relational theme has been instantiated, presupposes a naturalistic semantics like Dretske's (1981). For it assumes each basic emotion takes a core relational theme for its content even where the emoter has no cognitive state that could intrinsically represent that core relational theme.[12] The argument was that emotions represent core relational themes in virtue of being states of systems that evolved to be responsive to those core relational themes, and thus, also to have the function of representing them. These were all elements of Prinz's case against cognitivism—his attempt to show that an emotion can represent a complex situation when the only information the emoter has is from her body.

Dretske (1981)—along with the other philosophers who have attempted to naturalize intentionality by appeal to information theory, or teleosemantics, or consumer semantics, or whatever—have long struggled to overcome serious objections. To rehearse any of these debates would take us far afield, but no matter, because I think the relevant point here is much narrower: that the IRT account renders any appeal to naturalized semantics unnecessary. All we need are our three pieces of content, none of which requires a cognitive appraisal of anything (at least not in the most common senses of "cognitive" and "appraisal"). Thus, the IRT avoids the cognitivists' burden of showing how even basic emotions can involve highly intellectual appraisal judgments. And it also avoids the burden of somatic theorists like Prinz, as well as B&N, of showing how relatively primitive and reflex-like emotional responses could possibly involve anything as conceptually articulated or evidently theory-laden as a bona fide representation of a core relational theme would have to be. These hurdles are not on our track; nothing in the IRT requires either a cognitive appraisal or a representation of a core relational theme, for either basic or non-basic emotions. And since the IRT does not require basic emotions to represent core relational themes at all, it does not commit us to a controversial semantic theory like Dretske's (1981).

Also, second, since the IRT does not require representation of core relational themes, it does not require us to distinguish two different ways in which core relational themes are represented in emotions. Anytime a core relational theme is represented as part of an emotion, it will be a conceptually articulated intellectual

state. In such a case, the representation of the core relational theme will fit into the theory *as the third piece of content*, the one that represents a connection between the body-representing and world-representing pieces. Again, this third piece of content need not be intellectual, in which case the emotion would not represent a core relational theme. But it can and usually will be conceptually elaborated and intellectual, as it would have to be if it *did* represent a core relational theme. The important conclusion here is that the fact of a state's being an emotion implies almost nothing about core relational themes or how they are represented; it has only to do with the structure of its intentional content. Core relational themes are only required by emotion theorists, and perhaps sometimes by especially reflective emoters developing a background theoretical understanding of their own emotions. But core relational themes are not required for theorizing emotions, and, more to the present point, their representation does not support a strong distinction between basic and non-basic emotions.

I finally note that, if I am correct that the class of emotions is not divided into basic and non-basic states, then it becomes easier than B&N think to argue that emotion concepts refer to natural kinds. No need for core relational themes; thus, no need for a clever explanation of how basic emotions can represent core relational themes; thus, no need to distinguish basic from non-basic emotions on the basis of their different ways of representing core relational themes. Since emotions do not have to involve core relational themes, the fact that it is possible for emotions to be more or less conscious, intellectual, reflective, and cognitive no longer threatens the unity of emotion as a psychological category in the way Griffiths worries about. All emotions are defined by the structure of their intentional content, and admit of the same kind of explanation regardless of whether they are basic or non-basic, and regardless of whether or not they represent core relational themes. So if we think any complex psychological processes can be natural kinds, emotions might be among them.

3.2. Advantage 6: IRT better explains how emotional content gets integrated

I agree with B&N that emotions are constituted by the integration of representations of both the body and the external world. I also agree that there are many tough questions concerning the specific ways that the contents of these representations could get integrated, and like them I do not claim to have final answers to those questions (see p. 27). However, one assumption about the integration of emotion content seems to run through their entire discussion: that there are (at least) two distinct processes of integration involved. This distinction between different processes of integration again pushes B&N toward dividing the domain of emotion theory into different types of emotion calling for different kinds of explanation. And here again I will push back against that premise.

In the course of their discussion, B&N describe three ways that an emoter's representations of her body and her external world might get integrated. First, they

claim, "emotion could involve *perceptual integration*," in which case the simultaneous representation of an external object and a bodily state "might be considered a *multimodal perceptual state*" (2013, p. 663). Second, "emotion generation could depend on *cognitive integration*;" for example, if my belief that the exam is tomorrow, plus my belief that I am not prepared, together contribute to my state of fear about tomorrow's exam (p. 663). Third, they add, where both the two objects are represented consciously, "we do not simply have a case of informational integration, but also a case of *phenomenal integration*" (p. 672). And states that could be integrated in experience could be either perceptual or cognitive, for example, when one has the experience of bodily anger *and* a perception of an undeserved speeding ticket *and* a belief that one has been treated unjustly.

I am dubious about treating phenomenal integration as a distinct type of content integration, but I will not take up that issue here. Instead I will focus only on the perceptual and cognitive types of representation, since B&N's division of the class of emotions is mainly based on that distinction. Similarly to the way they distinguish basic and non-basic states of emotion, and in some ways connected with that distinction, this time they distinguish emotions that are *perceptually integrated* from others that are *cognitively integrated* (p. 663). Again, while there surely are different processes of integration involved, I do not think this justifies having the structure of our theory reflect this distinction, since the IRT can accommodate these different processes of integration while giving all emotions the same kind of explanation.

I need to do some sorting-out here to make clear exactly what it is that I reject from B&N's account. For they seem to blend together several distinctions that, while related, can themselves be distinguished. First, schematically: in some places they distinguish between perceptually integrated states and cognitively integrated states. Call this the *perceptual/cognitive* distinction. In other places they distinguish between states that integrate only perceptual information about the world and perceptual information about the body, vs. states that integrate both of those plus an additional piece of content, a type-specific propositional attitude (pp. 668–670). Call this the *two-piece/three-piece* distinction. In still other places they refer to a distinction between basic and non-basic emotions, the *basic/non-basic* distinction. So there are three distinctions here: (1) perceptual/cognitive, (2) two-piece/three-piece, and (3) basic/non-basic. These distinctions often get blended together in B&N's discussion. For example, first, at one point they identify cases of multimodal perception with "the simplest cases," and cognitively integrated states with "the most sophisticated cases," thereby seemingly blending the *perceptual/cognitive* distinction with the *basic/non-basic* distinction (p. 638). Second, they sometimes blend the *two-piece/three-piece* distinction with the *basic/non-basic* distinction, for example, when they state that "while basic emotions are constituted by the integration of bodily representations with representations of external objects, higher cognitive emotions require the integration of a further element (i.e., appraisal judgments)" (p. 671). Finally, third, they claim that the propositional attitudes

that give rise to the *two-piece/three-piece* distinction "are nothing but Lazarus's (1991) *appraisal judgments.*" Strictly speaking, appraisal judgments could be simply added to, but not cognitively integrated with, the other two pieces of content. But this would be strange, since appraisal judgments are representations of core relational themes, which are themselves cognitive representations of salient connections between the external object and the emoter, including the emoter's bodily response. So it would take special effort to avoid the conclusion that *three-piece* content must at least involve cognitive integration, and thus, it seems likely that B&N's view also blends the *two-piece/three-piece* distinction with the *perceptual/cognitive* distinction.

The first thing to get clearer about is that B&N do not really mean to commit, as far as I can tell, to two distinct types of integration, perceptual and cognitive. They do not claim that one process of integration is perceptual and the other cognitive. They also do not claim that one process integrates only perceptual content, while the other integrates only cognitive content. Rather, what they claim is that one integrative process includes only perceptual content, and the other includes both perceptual *and* cognitive content. More specifically, what they claim is that one process of integration combines only two pieces of perceptual content, while the other combines two pieces of perceptual content plus one piece of cognitive content. So the relevant distinction regarding perceptual vs. cognitive content is between *states with perceptual contents only*, and *states with both perceptual and cognitive contents*. Their two diagrammatic portrayals of "basic" and "higher cognitive" emotions make clear that basic emotions are two-piece states with perceptual content only, while non-basic or "higher cognitive" emotions are three-piece states with perceptual and cognitive contents (pp. 663, 670). In short, it seems clear that they do not have in mind two different modes of integration, perceptual and cognitive but only two different processes of integration, one of which involves cognitive content while the other does not.

What all of this implies, I think, is that the only distinction doing any explanatory work here is the *two-piece/three-piece* distinction. The *basic/non-basic* distinction matters only insofar as it is stipulated to be identical with the *two-piece/three-piece* distinction (as it seems to be), and the *perceptual/cognitive* distinction is not the one that matters for explanatory purposes. Whether the *basic/non-basic* distinction matters at all within an integrative theory depends entirely on how it is defined. If it is defined as identical with the *two-piece/three-piece* distinction, as B&N define it, then again it is the former distinction that does all the work.

So it is the *two-piece/three-piece* distinction that I will focus on. This is also the distinction that conflicts with the theory I have proposed. For, as I have made clear, the IRT is a uniformly *three-piece* theory of emotion content. My argument is that this unified three-piece theory can and should be used for explaining the entire class of emotions, with no splintering of the domain into subclasses, and that any distinction we want to draw between basic and non-basic emotions can and should be defined in a way that is consistent with this unified account.

The easy part of my job here will be to show how the IRT can accommodate both kinds of emotion, those that involve cognition and those that involve only low-level perception.[13] The more difficult part will be to explain why using the same theory for both is not Procrustean—why this is the best framework for explaining emotions with and without "higher" cognitive content. First, the easy part. As we have seen, the IRT specifies three pieces of representational content, one of which integrates the somatic and external objects by representing some kind of relevance relation between them:

(RC1) Representation of something in the external world.

(RC2) Representation of the state of the body.

(RC3) Representation of a relation between (RC1) and (RC2).

The risk of Procrusteanism stems from the fact that all emotion states could be forced into this form. The question is, should they be? If we treat all emotions as having this form, we get a *uniform* theory that explains all emotions in terms of (RC1)–(RC3). If we do not, we get a *splintered* theory that explains emotions in terms of (RC1) and (RC2) when they involve only low-level perception of the world and the body, but in terms of (RC1)–(RC3) when they also involve a cognitive appraisal.[14] Of course, other things being equal, a uniform explanation is to be preferred. So I think my view starts with some presumption. But are other things equal here? Is there a good reason to break up the set of emotional phenomena and develop two different kinds of explanations of emotion content, a two-piece explanation and a three-piece explanation?

On B&N's view, the cognitive element of a *non-basic* emotion with *three-piece* content would be identical with an "(emotion-)type-specific propositional attitude" (p. 669). In turn, they claim, such type-specific propositional attitudes will turn out to be "nothing but Lazarus's (1991) *appraisal judgments*" (p. 669). I think this analysis ends up fitting the form of my theory just fine. After all, their third piece of content is an appraisal judgment, which is a representation of the external object and the emoter as inhabiting a situation instantiating a core relational theme. In light of the relevance of the core relational theme, the emoter represents the object as having some bearing on her well-being. In other words, the cognitive representations that B&N have in mind always contain representations of the relation between the external object and the emoter—which include the emoter's body, and its responsiveness to the object. So while we do not explain the third piece of content in quite the same way, the IRT account of cognitive emotions seems functionally similar to the IST account. And in any case, there is no disagreement about the number of pieces of content to be integrated when it comes to non-basic or cognitive emotions.

Thus, the only disagreement here concerns the explanation of basic/perceptual/two-piece/non-cognitive emotions: B&N think these should get their own distinct kind of explanation; I disagree. Or, to translate everything into my vocabulary, they believe the explanation of these purely perceptual or non-cognitive emotions

should not appeal to (R3), but I believe it should. One way of putting the question is: Does the absence of any appraisal in an emotion render any reference to (R3) otiose? If both (RC1) and (RC2) are perceptual, and they involve no belief or the like, do we have reason to posit an additional piece of content, (RC3), that represents the integration of (RC1) and (RC2)? Again, B&N say "no," and I say "yes."

B&N clearly think some kind of integration of (RC1) and (RC2) is presupposed if the state we are talking about is to count as an emotion at all. So the question is: Concerning an emotion that involves multimodal perception, but not a cognitive appraisal, is the integration of (RC1) and (RC2) reflected in the state's representational content, or not? If not, we have two pieces of content: a representation of the body and a representation of the external object. But if so, we have three: a representation of the body, a representation of the external object, and a representation of their integration. So again, the *two-piece/three-piece* distinction is the important one here.

I believe there are several reasons to favor using the same form of explanation for all emotions, of which I will now give two. The first reason is just an appeal to empirical intuition: there does not seem to a very bright functional dividing line between emotions that involve cognitive appraisal and those that involve only low-level perception of the body and the external object. Emotional states at various levels of "basicness" seem more or less continuous in relevant ways; for example, emotions do not seem to change their basic functional character all that much when the emoter reflects on her own emotional state or gathers information about the object. An emotion that begins as a basic subcognitive fear response can evolve quite smoothly into a self-conscious and cognitively elaborated state of fear, without the basic nature of the emotional experience changing very much. If fear responses that involve appraisals were starkly different from those that do not, in the way B&N suggest, then this functional continuity is not what we would expect to find.

To explain my second reason for rejecting separate explanations for basic and non-basic emotions within an integrative account, I will focus on a part of B&N's own discussion. At one point they address Prinz's attempts to deal with a problem his theory faces: the problem of explaining how, if emotions are identical with bodily responses, they nevertheless seem to be essentially connected with objects in the world. Prinz's answer, they explain, is that "emotions, i.e., bodily perceptions, are directed toward particular objects because they are *linked* to other mental states that represent those particular objects" (p. 655). Then, they argue, Prinz "oscillates between three different characterizations" of the nature of this link (p. 655): it could be that the two states are linked in that they co-occur, or in that one causes the other, or that they somehow become unified in the emoter's mind or experience. B&N's discussion of these three suggestions sheds light on my current question about the way emotion contents get integrated. For each of Prinz's suggestions describes a way that (RC1) and (RC2) could be unified, and each brings trouble for him. I will address these three suggestions about the way perceptions of

external objects and bodily responses might be linked—*co-occurrence, causation, unification*—in reverse order, so that I can finish with the co-occurrence view, which contains the argument that is relevant to the present discussion.

The problem for Prinz with saying that the contents of (RC1) and (RC2) are unified is that it would abandon his purely somatic theory in favor of an integrative theory like mine or B&N's. It seems clear that this is where B&N expect Prinz to end up (as do I) once he rejects the co-occurrence and causal views. The connection between emotional responses and external objects must be some kind of unification or, better, integration. As regards Prinz, so far, so good.

The problem for Prinz with saying that the referents of (RC1) or (RC2) are linked in that one causes the other is that what the emoter takes as the external object of her emotion might not be its actual cause. B&N develop clever examples to show how this could occur. I'll give a simpler example here: suppose a loud bang frightens you as you are looking at a person who looks like he might be holding a gun. You feel afraid, and the object of your fear is the gun you think he might have. But it turns out that the noise was caused by someone dropping a heavy book on the floor behind you. In this case the object of your fear—the (possible) gun—is not the cause of your body's response. Cases like this suffice to show that causation is not generally what links the emoter's response to an external object. Where causation is part of the story, the way it gets involved in the emoter's mental state is mediated by perceptual or cognitive processing; the mere fact of causation is not enough to forge the representational link. Here again I generally agree with B&N's conclusion.

The argument that is useful here comes in their discussion of the first possible link between the referents of (RC1) and (RC2), *co-occurrence*. Prinz's suggestion here was that an emotion's being directed at a certain object might be a matter of a neural representation of that object existing in the brain at the same time as a neural representation of the body's response (2004, p. 181). But the problem with the co-occurrence view, as B&N point out, is that the representation of the body's response might co-occur with many, many other neural representations, and we need a principled way of specifying which of them is the one that represents the emotion's object. For example, at the same time you heard the loud bang while looking for a gun, you might also have seen a hot air balloon overhead, smelled a skunk, and so on, but the co-occurrence of one of these other perceptions with your fearful response would not make it the object of your fear. Most of the indefinitely many co-occurring things you represent are *not* the object of your emotion. Thus, the co-occurrence explanation, by itself, also fails.

I am arguing that we should treat the emoter's perceptual representation of her body and her world, even for the most basic kinds of emotions, as feeding into a third piece of content that represents their integration. This is where this last argument from B&N comes in: the reason it is not enough that the representation of the body and the external object co-occur is that co-occurrence establishes no special link of about-ness or directedness of the emotional response onto its

specific external object. It must be something more than causation or co-occurrence that forges that link. If the kind of sub-cognitive emotion at issue here is nothing more than a multimodal perception, as B&N claim, it cannot be multimodal only in the sense of being a case of multiple perceptions from different modes merely co-existing in the same brain. To explain the psychological facts *there must be functional integration*.

Finally, the argument I have been setting up: Since there must be functional integration of the first two pieces of perceptual content, we must posit a third piece of representational content in which the two pieces are integrated. For I think there is only one way of resisting this conclusion, and it fails. That is, one could try to argue that an emotion integrates perceptual information about bodily changes and an object in the world, but that this integration does not itself give rise to anything that would count as a representation of anything. This would be to suppose that the integration is only at the neural level, that not all neural processing of perceptual information produces states that have representational content, and that the integration of the two perceptual constituents of a basic emotional state is like this—neurally, but not representationally, integrated. Besides being counterintuitive, and frustrating the goal of psychological explanation, the problem is that these two parcels of information must also be functionally integrated in order to explain the about-ness relation between the object in the world and the body's response. And if these are functionally integrated, then a state that integrates representations of them, and accounts for their functional connection, must surely also *represent* that functional connection.

Generally, the Dretskean (1981) semantics that B&N endorse for their analysis of the contents of basic, sub-cognitive emotions would be of no help in trying to resist my claim that integration always gives rise to a third piece of representational content. For neurally integrated states themselves have regular informational and teleo-functional connections with their causes, which, on those semantic theories, would mean they are representations of whatever they carry information about. And what an integrated bimodal perceptual state would carry information about is: (1) an object as perceived in mode a, (2) an object as perceived in mode b, and (3) a connection of information from both modes. So again we are pushed toward something like the IRT, with three pieces of content, for even the most low-level perceptual states. To avoid this conclusion, B&N would need to explain some other way that a neural state which integrates two other neural states, both of which represent things, could exist without itself having representational content. Failing this, we will have to conclude that, even where both are low-level perceptions, two perceptual states could only be integrated in a further state that has content not identical with that of the two integrated states. So there will always be three pieces of content, with the third integrating the first two.

It is true that what gets represented in emotion is not always represented cognitively, certainly not in the sense of involving belief or appraisal or judgment, and so on. It is also true that most of it *can* be represented in these high-level cognitive

ways. Any part of an emotion's content could come in as part of a propositional attitude, for example. But in distinguishing different levels of "cognitiveness" that the representational components of an emotion could have, we do not commit ourselves to a distinction between different species of emotion calling for different explanations. All can be explained within the IRT framework.

Notes

1. Other theories of emotion that incorporate representation of both the body and the external world include the various forms of "enactive" theories of emotion, such as those of Colombetti and Thompson (2008)and Maiese (2014). According to those theories, emotions are constituted by both kinds of representation, but in such a way that they are in some sense fused together and indistinguishable. Another kind of integrative account is developed by Lewis (2005). Lewis outlines a complex explanation of emotions based on dynamical systems theory, according to which emotions are fundamentally processes of self-organization involving interaction among several different cognitive subsystems including arousal, action tendencies, attention, and feeling. I believe that comparing my integrative account with these others would reveal further advantages of my proposal, but here I only address what I take to be the most formidable alternative to my view, and the one the discussion of which is most illuminating for my purposes here.
2. I refer to these as "purely" somatic theories because James (1890/1950) famously identifies emotions with feeling of bodily changes, and Prinz (2005) argues that bodily changes are both necessary and sufficient for emotion.
3. See, for example, Prinz (2004), especially pp. 72–74, where he argues that different kinds of emotion do in fact have distinct corresponding patterns or "prototypes" of bodily response. I suspect that he infers too much from the philosophical and empirical arguments he gives for this conclusion, and that it will not be possible to match emotions with distinct patterns of physiological change. But this is an argument for a different occasion; my points here are only (1) that this part of Prinz's explanation is consistent with the analysis of perceptions of the body within my account, and (2) that the IRT requires far less specification of bodily processes than a Jamesian theory (1890/1950) does, since it does not purport to define or explain emotions solely in bodily terms.
4. It is important to keep in mind that the emoter's body is part of her world. And clearly it is possible to have emotions about our bodies. So we cannot use the phrase "in the world" to mean the same as "in the external world."
5. The kind of relevance I have in mind will be addressed in the discussion of the third element, representation of the relation between the contents of the emoter's world-facing and body-facing perceptions.
6. These three options—co-presence, causation, and unification—are given in Barlassina and Newen (2013, pp. 656–658), and discussed in some detail in Section 3.
7. Compare with B&N's example of "The Disgusting Can of Coke" (p. 21).
8. A worry one might have at this point is that my emphasis on the person's judgment about the sandwich's relevance, as opposed to her unconscious perception of the terrifying image, is just an expression of cognitivist bias on my part. I have favored the more cognitive emotion over the more basic one in my explanation, it is true. But my reason is not mere bias. Rather, it is my assumption that human emotions nearly always involve cognitive processes, and, when they do, these cognitive

processes subsume most lower level representational processes and determine their psychological impact. Conscious or intellectual emotional processing is not always present, but when present it almost always overpowers more reflex-like processing within the person's psychology. For a human being, a reflex-like emotional response will tend to evolve quickly into something far more psychologically complex.

9. Damasio (1994, esp. p. 151), gives an accessible discussion of the brain's ongoing representation of the state of the body and the contribution of this representation to emotions.
10. As it happens, Tye independently reaches a similar conclusion about the intentional structure of emotions in his article "The Experience of Emotion: An Intentionalist Theory" (2008). He states, "for each emotional experience, there is a perceptual experience (or a thought) typically directed on something external and a bodily sensation or feeling. What the emotional experience does is to bring together the content of the former and the content of the latter" (pp. 39–40).
11. See, for example, Bayne and Montague (2011).
12. The concept of intrinsic representation I have in mind here is something like the one explained by Stufflebeam (2001, esp. pp. 403–405).
13. This distinction between low-level perception and cognition is far from universal, and is sure to invite objections. After all, many cognitive scientists refer to perceptual processes as cognitive, and that distinction is not easy to justify. All I will say about this is that for present purposes, and in the interest of aligning my conceptual background with Barlassina and Newen's, I am thinking of cognitive content generally as belonging to "higher" or intellectual processes—that is, as contents residing in the same neighborhood as beliefs, judgments, thoughts, appraisals, and the like.
14. In the interest of brevity, in what follows I will use the word "appraisal" as shorthand for the family of cognitive concepts including appraisals, beliefs, propositional attitudes, judgments, and the like.

Acknowledgments

I would like to thank Jennifer Brockman and two anonymous referees for comments on previous drafts of this paper.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Chad Brockman http://orcid.org/0000-0001-8623-8025

References

Barlassina, L., & Newen, A. (2013). The role of bodily perception in emotion: In defense of an impure somatic theory. *Philosophy and Phenomenological Research, 89*(3), 637–678.

Bayne, T., & Montague, M. (2011). *Cognitive Phenomenology*. Oxford: Oxford University Press.

Byrne, A. (2001). Intentionalism defended. *Philosophical Review, 110*(2), 199–240.

Colombetti, G., & Thompson, E. (2008). The feeling body: Towards an enactive approach to emotion. In W. F. Overton (Ed.), *Developmental perspectives on embodiment and consciousness* (pp. 45–68). New York, NY: Erlbaum.

Damasio, A. (1994). *Descartes' error*. New York, NY: Harper Collins.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Duncan, S., & Feldman-Barrett, L. (2008). Affect is a form of cognition: A neurobiological analysis. *Cognition and Emotion, 21*(6), 1188.

Gray, J. R., Braver, T. S., & Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Sciences, 99*(6), 4115–4120.

Griffiths, P. (2004). Is emotion a natural kind? In R. C. Solomon (Ed.), *Thinking about feeling: Contemporary philosophers on emotions* (pp. 233–249). New York, NY: Oxford University Press.

James, W. (1890/1950). *The principles of psychology*. New York, NY: Dover.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences, 28*, 169–245.

Maiese, M. (2014). How can emotions be both cognitive and bodily? *Phenomenology and the Cognitive Sciences*, *13*(4), 513–531.

Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. New York, NY: Oxford University Press.

Prinz, J. (2005). Are emotions feelings? *Journal of Consciousness Studies, 12*(8–10), 9–25.

Stufflebeam, R. S. (2001). Brain matters: A case against representations in the brain. In W. Bechtel, P. Mandik, J. Mundale, & R. S. Stufflebem (Eds.), *Philosophy and the neurosciences: A reader* (pp. 395–413). Oxford: Blackwell.

Tye, M. (1995). A representational theory of pains and their phenomenal character. *Philosophical Perspectives, 9*(A1), 223–239.

Tye, M. (2008). The experience of emotion: An intentionalist theory. *Association Revue Internationale de philosophie, 62*(242), 25–47.