

A CENTURY OF MISUNDERSTANDING?

WILLIAM JAMES'S EMOTION THEORY

Jake Spinella

University of Illinois Chicago

jspine2@uic.edu

2019 WILLIAM JAMES SOCIETY
YOUNG SCHOLAR PRIZE WINNER



I argue, *contra* traditional interpretations of William James's emotion theory like that of Antonio Damasio and alternate interpretations like those of Phoebe Ellsworth and Lisa Barrett, that James is best classified as a *functionalist* regarding emotion. In arguing for this point, I will make four textual claims: (1) James was an important precursor to Basic Emotion Theory (BET) and his theory is best identified as a flavor of BET; (2) James's theory individuates emotion categories by their *evolutionary, functional roles*; (3) The only necessary condition on something being an emotion is that it is a bodily feeling; and, (4) Contrary to Barrett and Ellsworth, James was loath to offer a definitive list of basic emotions not because he loathed taxonomy but rather because he thought psychology was not yet a natural science with well-defined theoretical categories. I will then argue that a proper understanding of James's emotion theory defangs some critiques of BET and of Neo-Jamesian theory.



Phoebe Ellsworth, in her 1994 article “William James and Emotion: Is a Century of Fame Worth a Century of Misunderstanding?” wryly observed: “Ask anyone about William James’s theory of emotion and you will almost certainly hear about the bear.”¹ This opening sentence sets the stage for Ellsworth’s critique of the standard interpretation of James’s theory of emotion. The standard interpretation of that theory sees James claiming that emotions like anger, disgust, fear, etc., are discrete categories constituted exclusively by the perception of internal bodily feelings. This article, coupled with the 1994 release of Antonio Damasio’s explicitly “Neo-Jamesian” *Descartes’ Error: Emotion, Reason, and the Human Brain*, led to a significant resurgence of scholarly and scientific interest in James’s emotion theory that persists to this day. Ellsworth’s article has spawned a large secondary literature regarding the proper historical interpretation of James’s emotion theory, with numerous emotion theorists claiming that their theories are the true inheritors of James’s legacy.² I will begin this paper by outlining the traditional interpretation of James as a proponent of Basic Emotion Theory (“BET”) and then profile Ellsworth’s alternate interpretation as a counterpoint. I argue that James’s emotion theory has been largely misunderstood, not just by the aforementioned representative examples but also by Ellsworth herself. In arguing for this historical point, I will forward four claims: (1) James was an important precursor to BET and his theory is most comfortably identified as a flavor of BET or proto-BET; (2) James’s proto-BET individuates individual emotion categories by the *evolutionary, functional roles* of emotions rather than by later BET’s focus on *emotion signatures* in facial expressions, the autonomic nervous system, etc.; (3) The only necessary condition on something being an emotion in James’s theory is that it is a bodily feeling, though appraisals often in fact play important roles in emotion generation; and finally, (4) contrary to both Barrett and Ellsworth, James was loath to offer a definitive list of basic emotions not because he loathed taxonomy but rather because he thought psychology was not yet a natural science with well-defined theoretical categories. After marshalling evidence for

these four claims, I will canvass how many of the most popular and recent readings of James (both friendly and critical) get him wrong, and I will subsequently extract some lessons for the contemporary emotions debate whose argumentative dialectic is (to this author, at least) largely the same as it was when James was writing. In particular, I will argue that a proper understanding of James's emotion theory defangs some traditional critiques of BET and of Neo-Jamesian theory, forcing critics to reformulate their critiques.³

I. THE TRADITIONAL ACCOUNT OF JAMES'S EMOTION THEORY

Perhaps the most famous passage in the last two centuries of emotion theory comes from James's 1884 article in *Mind* entitled "What is an Emotion?". Pre-Jamesian accounts of emotions saw emotions as intrinsically motivating mental events that induced us to action. James, in "What is an Emotion?", sought to turn this common wisdom on its head:

Our natural way of thinking about these standard emotions is that the mental perception of some fact excites the mental affection called the emotion, and that this latter state of mind gives rise to the bodily expression. My thesis on the contrary is that *the bodily changes follow directly the perceptions of the exciting fact, and that our feeling of the same changes as they occur is the emotion*. Common sense says, we lose our fortune, are sorry and weep; we meet a bear, are frightened and run; we are insulted by a rival, are angry and strike. The hypothesis here to be defended says that this order of sequence is incorrect, that the one mental state is not immediately induced by the other, that the bodily manifestations must first be interposed between, and that the more rational statement is that we feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful, as the case may be. Without the bodily states following on the perception, the latter would be purely cognitive in form, pale, colorless, destitute of emotional warmth. We might then see the bear, and judge it best to run, receive the insult and deem it right to strike, but we could not actually *feel* afraid or angry.⁴

Emotions, *per James*, are not mental events that induce action but perceptions of internal bodily feelings that arise when we engage in action. In this regard, emotions are caused by “exciting perceptions” and by actions rather than actions being caused by emotions. This account of emotions as bodily feelings is coupled with a Darwinian story about how this special subset⁵ of bodily feelings came to be:

. . . the nervous system of every living thing is but a bundle of predispositions to react in particular ways upon the contact of particular features of the environment. As surely as the hermit crab’s abdomen presupposes the existence of empty whelk-shells somewhere to be found, so surely do the hound’s olfactories imply the existence, on the one hand, of deer’s or foxes’ feet, and on the other, the tendency to follow up their tracks . . . The labors of Darwin and his successors are only just beginning to reveal the universal parasitism of each special creature upon other special things, and the way in which each creature brings the signature of its special relations stamped on its nervous system with it upon the scene.

Every living creature is in fact a sort of lock, whose wards and springs presuppose special forms of key—which keys however are not born attached to the locks, but are sure to be found in the world near by as life goes on. And the locks are indifferent to any but their own keys.⁶

Emotions, for the James of “What is an Emotion?” are bodily feelings that prepare us to act in certain ways in order to solve certain problems due to both adaptive pressures by natural selection and our individual faculties for associative learning and habit formation:

To discuss thoroughly this objection [that the objects of our emotions are conventional] would carry us deep into the study of purely intellectual *Æsthetics*. A few words must here suffice. We will say nothing of the argument’s failure to distinguish between the idea of an emotion and the emotion itself. We will only recall the well-known evolutionary principle that when a certain power has once been fixed in an animal by virtue of its utility in presence of certain features of the environment, it may turn out to be useful in presence of other features of the environment that had originally

nothing to do with either producing or preserving it. A nervous tendency to discharge being once there, all sorts of unforeseen things may pull the trigger and let loose the effects. That among these things should be conventionalities of man's contriving is a matter of no psychological consequence whatever . . .⁷

This reading of James's emotion theory, though from before the time when BET was known as such, shares numerous commonalities with the later theory forwarded in different guises by Silvan Tomkins,⁸ Paul Ekman,⁹ Carroll Izard,¹⁰ Damasio,¹¹ and others. Canonical BET proposes that emotions are evolutionarily selected-for "modules" or "affect programs" with a distinct neural or behavioral signature for each basic emotion such as anger, fear, sadness, etc.¹² Basic emotions are evolutionarily selected-for, found universally in human culture, and constitute more complicated emotions via their combination.¹³ Given the high degree of similarity between James's theory in "What is an Emotion?" and later examples of self-consciously adopted BET, many BET proponents have retroactively dubbed James a basic emotions theorist. Later on, I will argue, *contra* Ellsworth and Barrett, that this decision to call James a proponent of BET is a legitimate historical move, but not in the way that he has been commonly understood as such. I will now profile Ellsworth's alternate interpretation.

II. ELLSWORTH'S COGNITIVE-APPRAISAL INTERPRETATION OF JAMES

Ellsworth's main textual source for her reinterpretation of James's theory of emotion is his 1894 *Psychological Review* article "The Physical Basis of Emotion." In this article, James responds to various criticisms of his theory that had been developed since the publication of "What is an Emotion?". Ellsworth's main contention is that James is best characterized as a progenitor of the later cognitive-appraisal theory of emotions rather than as a proponent of BET. The cognitive-appraisal theory of emotions, first forwarded by Stanley Schacter and Jerome Singer, claims that emotions are

combinations of undifferentiated states of physical arousal (i.e. emotions do not have distinct neural or behavioral signatures) coupled with a *cognitive appraisal* of a situation.¹⁴ On this account, fear is, roughly, a physiological state of high arousal and negative affect coupled with an appraisal of the eliciting situation, say, encountering a bear, as dangerous. Another way of putting this is that both bodily feelings and cognitive appraisals are necessary conditions for an emotion to occur. Ellsworth argues that the textual basis for this interpretation lies in James's clarifications in "The Physical Basis of Emotion" coupled with a careful analysis of the phrase "perceptions of the exciting fact" found in the canonical quote in "What is an Emotion?"¹⁵ Ellsworth takes "perceptions of the exciting fact" to obviously mean cognitive appraisal—we judge the bear as frightening, and, coupled with our bodily disturbances, are put into an emotion state of fear. Furthermore, Ellsworth argues that the common reading of James, where the temporal sequencing of emotion events is such that behaviors determine emotion rather than *vice versa*, is incorrect. Rather, the proper interpretation of James problematizes the idea of a privileged temporal ordering, because he sees these processes as simultaneous:

Debates about the primacy of cognition, bodily responses, or feeling make little sense when emotions are considered as a stream. The question of the role of peripheral feedback only makes sense when phrased as the question James' hypothesis originally posed: Are bodily sensations necessary for the subjective feeling of emotion? The question of whether what is occurring is an emotion at all becomes a matter of semantics, of different theorists' preferences for different moments in the flow of events when, according to their different definitions, "cognition," or "affect," or "bodily feedback," or "emotion" has been achieved. Over the past century, James' stunning paragraph, describing the sequence of events in large units of perception (see a bear), behavior (tremble, run), and feeling (feel afraid) has drawn our attention away from the recognition that none of these units is elemental, none is stable. They are all in motion, all the time, and there is no reason to believe that one must end before another begins.¹⁶

While this description of the process of emotion generation sounds Jamesian in temperament with its focus on process, flux, and boundary-mixing, it does not cohere well with James's established writing on the temporal sequencing of physiological processes found in his other work. The reason James has the temporal ordering of emotion events that he does is because of his interpretation of contemporary physiological knowledge found in his 1880 article "The Feeling of Effort." In that article, James makes the claim, *contra* Wilhelm Wundt and others, that bodily sensation is an *afferent feeling*. This is to say that the cognitive intention to move does not create the feeling of effort beforehand by "innervating the nerve currents" of the muscles, skeleton, and viscera of the relevant motor region, rather the nervous activity of the muscles, skeleton, and viscera generate the feeling of effort, which we then take notice of:

In opposition to this popular view, I maintain that the feeling of *muscular* energy put forth is a complex *afferent* sensation coming from the tense muscles, the strained ligaments, squeezed joints, fixed chest, closed glottis, contracted brow, clenched jaws, etc., etc. That there is over and above this another feeling of effort involved, I do not deny; but this latter is purely moral and has nothing to do with the motor discharge. We shall study it at the end of this essay, and shall find it to be essentially identical with the effort to remember, with the effort to make a decision, or to attend to a disagreeable task.¹⁷

Bodily feelings, for James, strictly precede cognitive interpretations, and he thought this because of the contemporary physiological understanding of how afferent nerve currents in the motor system worked. Since James launched an extended defense of the idea of a privileged temporal sequencing of physiological processes in "The Feeling of Effort," we can reasonably assume that James cared about establishing a temporal and explanatory ordering of events involving bodily sensations (including emotions), contrary to Ellsworth. Ellsworth also gets the reconstruction of James's

temporal ordering wrong, even as she decries attempts to identify a privileged temporal ordering of emotion generation. She characterizes James's main innovation as changing the common sense understanding of emotion processing of Stimulus → Interpretation → Affect → Bodily Response by switching bodily response and affect. But really, James's view at the time of "What is an Emotion?" seems to be Stimulus → Bodily Response → Interpretation → Affect rather than Stimulus → Interpretation → Bodily Response → Affect.

Note that this textual claim can coexist with the further claim that there is no privileged temporal ordering *tout court*. James may well have been searching for a privileged temporal ordering that is appropriate for the purposes of scientific psychology. This gloss can render much of James's work on the psychology of his day consistent with the self-professed perspectivalism found in the rest of his corpus. Later, I will elaborate on this point in further detail.

Ellsworth's alternative reading of James is called into further question by James's belief that "exciting perceptions" can bypass interpretation, as when "we abruptly see a dark moving form in the woods our heart stops beating, and we catch our breath instantly and before any articulate idea of danger can arise," or when we have a strong association between stimulus and response through classical conditioning.¹⁸ Furthermore, her account of James as a cognitive-appraisal theorist faces significant difficulty in reconciling the idea that emotions are epiphenomenal with appraisal theory generally, since cognitive appraisals are usually taken as action-motivating. While Ellsworth is correct in her main contention that James has been misread in the century since his death, her alternative reading does not pass muster as an exegesis of James's theory of emotions. I will now offer my own interpretation of James's theory of emotions in the following sections.

III. JAMES'S FUNCTIONALIST THEORY OF EMOTIONS

Contrary to both the mainstream interpretation of James as a traditional proponent of BET and Ellsworth's alternate account of James as a proto-appraisal theorist, I will argue that James is best

characterized as a *functionalist* about emotion categories. A functionalist with regard to emotion categories is different from a traditional Basic Emotion theorist, who usually defines emotions as evolutionarily task-specific but whose conditions of membership are predicated on *interindividual similarity* rather than *functional role*. This is another way of saying that traditional BET posits certain ‘signatures,’ whether facial, autonomic, or cognitive, that are the essence of what it is to be in a certain emotion state. In contrast, James’s emphasis is squarely on the role played by each of our emotion categories rather than by any specific similarity in every instance of, say, anger:

Both Dr. Worcester and Mr. Irons are struck by this variability in the symptoms of any given emotion; and holding the emotion itself to be constant, they consider that such inconstant symptoms cannot be its cause . . . People weep from excess of joy; pallor and trembling accompany extremes of hope as well as of fear, etc.

. . . How can any definite emotion, he [Dr. Lehmann] asks, exist under such circumstances, and what is there then left to give unity to such concepts as anger or fear at all? The natural reply is that the bodily variations are within limits, and that the symptoms of the angers and of the fears of different men still preserve enough *functional* resemblance, to say the very least, in the midst of their diversity to lead us to call them by identical names. Surely there *is* no definite affection of ‘anger’ in an ‘entitative’ sense.¹⁹

James explicitly rejects the idea that there are any facial expressions, autonomic signatures, or cognitive appraisals that are fixed and invariant in an emotion category, though he also affirms the reality of certain congenital dispositions that arise because of our physiological and evolutionary organization:

That one set of ideas should compel the vascular, respiratory, and gesticulatory symptoms of shame, another those of anger, a third those of grief, a fourth those of laughter, and a fifth those of sexual excitement, is a most singular fact of our organization, which the labors of a Darwin have hardly even begun to throw light upon. Where such a prearrangement of the nerve centres exists, the way to

awaken the motor symptoms is to awaken first the idea and then to dwell upon it. The thought of our enemy soon brings with it the bodily ebullition, of our loss the tears, of our blunder the blush. We even read of persons who can contract their pupils voluntarily by steadily imagining a brilliant light—that being the sensation to which the pupils normally respond.²⁰

On this reading, fear is united when one predisposes oneself to neutralizing a threat, in extreme cases through fight or flight, but also in more mundane cases like avoiding rain through the purchase of an umbrella. Anger predisposes one to respond aggressively to correctly perceived injustices and inequities committed by others, etc.²¹ The next question this account needs to answer is: what fixes the function of a certain emotion category? The answer is two-fold: the first, and most general, answer is the Darwinian one. We have the predisposition to act in certain ways because of the process of natural selection. A human whose fear response included walking towards a bear with open arms would not survive very long in our ancestral epoch, and so natural selection favored certain kinds of responses to certain kinds of problems. Much of this line of thinking can be found in the previously quoted material from “What is an Emotion?” on page four of this paper. The second way our emotions can be functionally defined is by way of associative learning: we fear getting wet and so avoid it by buying an umbrella. We do this through force of habit rather than evolutionary selection insofar as we had to go through multiple individual experiences of getting wet before we established the association between rain and feelings of unpleasantness from wet clothes, hair, etc.

One tension in this reading of James that must be addressed, however, is the relationship between emotions and motivations for action. Recall James’s assertion in “What is an Emotion?” that emotions do not motivate action but rather arise as a result of action. If emotions are defined by their functional role, they are implicitly defined in terms of predispositions to act in certain ways, given certain apprehensions of bodily feeling, to accomplish certain evolutionarily-prescribed goals. If this is in fact the case, then there is an inconsistency in James claiming that emotions are both

functional predispositions and that they do not motivate action. I think this tension dissolves if the following passage is carefully read, however:

I think that all the force of such objections lies in the slapdash brevity of the language used, of which I admit that my own text set a bad example when it said ‘we are frightened because we run.’ Yet let the word ‘run’ but stand for what it was meant to stand for, namely, for many other movements in us, of which invisible visceral ones seem by far the most essential; discriminate also between the various grades of emotion which we designate by one name, and our theory holds up its head again. ‘Fear’ of getting wet is not the same fear as fear of a bear. It may limit itself to a prevision of the unpleasantness of a wet skin or of spoiled clothes, and this may prompt either to deliberate running or to buying an umbrella with a very minimum of properly emotional excitement being aroused. Whatever the fear may be in such a case it is not constituted by the voluntary act.²²

When attended to, this passage indicates that James eventually abandoned the position he was most famous for: that emotions are not themselves motivating. This commitment falls out of the theory when suitably clarified in “The Physical Basis of Emotion” because internal visceral changes, rather than behaviors, are what cause emotions. This leaves space for emotions to motivate action—perceptions of internal bodily change give rise to emotional mechanisms that *then* predispose one to act in a variety of ways—and it seems as if James has acknowledged this point by calling “we are frightened because we run” a “bad example.”²³ A coherent functionalist account of James’s emotion theory is committed to the claim that James ultimately rejected the epiphenomenality of emotions. Indeed, it seems as if James was committed to the motivating nature of emotions in other writings, and he seems to have not made a similar point regarding epiphenomenality at any point past 1884’s “What is an Emotion?”, which lends plausibility to the interpretive claim that he ultimately revised his idea regarding the motivational status of emotions. So much the better for his theory.

The next aspect of James's theory to be addressed is what constitutes the *process* of emotions in the body and brain. My functionalist reading of James agrees with the traditional interpretation in claiming that emotions are perceptions of bodily feelings and not perceptions of bodily feelings *and* a cognitive appraisal like that of Ellsworth and later, Lisa Barrett.²⁴ James repeatedly stresses the physiological nature of emotions in all his written works on emotions, even going so far as to claim that if we:

try to abstract from our consciousness of [the emotion] all the feelings of its characteristic bodily symptoms, we find we have nothing left behind, no "mind-stuff" out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains.²⁵

The cognitive-appraisal interpretation can only be saved on this point if, as Ellsworth argues, "exciting perceptual facts" are all and only appraisals.²⁶ But along with the passage from the *Principles* quoted in section II regarding how certain emotions can be elicited without the mediation of an appraisal mechanism, James is also committed to the idea of *objectless emotions*. Objectless emotions do not obviously involve appraisals,²⁷ and there are passages where the existence of such objectless emotions leads James to *explicitly reject* appraisals as being necessary conditions for emotion generation:

Both Dr. Worcester and Mr. Irons insist on the fact that consciousness of bodily disturbance, taken by itself, and apart from its combination with the consciousness of an exciting object, is not emotional at all . . . The facts must be admitted; but in none of these cases where an organic change gives rise to a mere local bodily perception is the reproduction of an emotional diffusive wave complete. Visceral factors, hard to localize, are left out; and these seem to be the most essential ones of all. I have said that where they also from any inward cause are added, we *have* the emotion; and that then the subject is seized with objectless or pathological dread, grief, or rage, as the case may be.²⁸

The balance of textual evidence indicates, contrary to alternative readings of James, that emotions are fundamentally perceptions of the feeling of internal bodily change in the viscera. Visceral factors are, for James, those parts of the inner body that are the sources of *involuntary feedback*, such as glands, reflexes, etc., as opposed to the largely (though not entirely) voluntary internal feedback of the muscles.²⁹ This is not to say that visceral factors are the *only* internal bodily feelings relevant to emotions—it is just the case that they are the most important, given the voluntary nature of the muscular system in James’s account. To the extent that there are involuntary musculoskeletal sensations that prime us for action, those play a role in emotion generation as well.³⁰ On this reading, then, the presence of evolutionarily and individually selected-for internal visceral feedback that predisposes the user to action is what differentiates emotions from other affect-states—visceral feedback is a necessary condition on emotion generation.

While James denies that appraisals are a necessary condition for emotion generation, he does not deny the importance of appraisals in most cases of emotions—of course, cases of shame, anger, and fear routinely involve the assessment of something as meriting shame, anger, or fear—but James wants to make the claim that despite appraisal’s importance in paradigm cases, the only thing present in *all* cases of emotions is the perception of felt bodily (especially visceral) change.

My final interpretive claim is that, contrary to Ellsworth and Barrett, James was in fact searching for a privileged set of emotion categories but that he thought there was a privileged schema only *relative to the purposes of scientific psychology*. Ellsworth makes the claim that James was not a Basic Emotion theorist because he would be loath to privilege a certain set of emotional categories as the “right” or “real” one:

James probably would not have condemned the study of the processes by which people or cultures select meaningful events from the infinity of possibilities or the comparison of different emotional

representations across individuals or cultures. What he did condemn was the assumption that some particular selection was the right one or the real one, that by studying and documenting one particular arbitrary set of distinctions in detail we might discover truths about emotion that would hold for all people and all times. He would have rejected the idea that there are 6 or 7 or 10 or 20 basic emotions, and the effort to prove that a particular affective experience should or should not be considered an emotion. Likewise he would have been contemptuous of efforts to establish definitive distinctions among related emotions such as guilt, shame, embarrassment, and humiliation or empathy, sympathy, compassion, and pity unless the effort was designed to apply only to a particular culture at a particular moment.³¹

While it is true that James never explicitly offered a definitive list of basic emotions, he did come very close:

Rapture, love, ambition, indignation, and pride, considered as feelings, are fruits of the same soil with the grossest bodily sensations of pleasure and of pain. But it was said at the outset that this would be affirmed only of what we then agreed to call the “standard” emotions; and that those inward sensibilities that appeared devoid at first sight of bodily results should be left out of our account.³²

Rapture (read as happiness), love, indignation (read as anger), and pride are canonical examples of basic emotions in later BET. Throughout his writing on emotion, James routinely refers to these paradigm cases of emotion as paradigm cases, even going as far as calling them “standard emotions.”³³ While he did not explicitly say that these are privileged emotion categories, it at the very least reduces the evidentiary plausibility of Ellsworth’s claim that James would have rejected a list of basic emotions outright. Ellsworth cites the following passage to further support her claim that James would be loath to privilege a particular categorization as *the* correct categorization:

This is all I have to say about the emotions. If one should seek to name each particular one of them of which the human heart is the seat, it is plain that the limit to their number would lie in the introspective vocabulary of the seeker, each race of men having found names for some shade of feeling which other races have left undiscriminated. If then we should seek to break the emotions, thus enumerated, into groups, according to their affinities, it is again plain that all sorts of groupings would be possible, according as we chose this character or that as a basis, and that all groupings would be equally real and true. The only question would be, does this grouping or that suit our purpose best?³⁴

Ellsworth, strictly speaking, is correct in saying that James would never claim there is *one* privileged emotion category *irrespective of one's purposes*, but he is very clear that there can be privileged emotion categories given the necessities and demands of the inquiry in question. And, in fact, James advocated for a shift to a more naturalist methodology in scientific psychology that would see rationalist introspective psychology replaced by the psychology of “the biologists, nerve-doctors, and psychical researchers.”³⁵ In doing so, James explicitly advocates for a set of privileged categories with respect to psychological science, namely the ongoing identification of mental states with brain states and peripheral nerve currents:

One great reason why Professor Ladd cares so little about setting up psychology as a natural science of the correlations of mental with cerebral events, is that brain states are such desperately inaccessible things. I fully admit that any *exact* account of brain states is at present far beyond our reach; and I am surprised that Professor Ladd should have read into my pages the opinion that psychology as a natural science must aim at an account of brain states exclusively, as the correlates of states of mind. Our mental states are correlated *immediately* with brain states, it is true; but, more remotely, they are correlated with many other physical events, peripheral nerve currents for example, and the physical stimuli which occasion these. Of these latter correlations we have an extensive body of rather orderly knowledge. And, after all, may we not exaggerate the degree

of our ignorance of brain states themselves? We do not know exactly what a nerve current is, it is true; but we know a good deal *about* it . . . Now the provisional value of such knowledge as this, however inexact it be, is still immense. It sketches an entire program of investigation, and defines already one great *kind* of law which will be ascertained. The *order in time* of the nerve currents, namely, is what determines the *order in time*, the coexistences and successions of the states of mind to which they are related. Professor Ladd probably does not doubt the nerve-current theory of motor habits; he probably does not doubt that our ability to learn things ‘by heart’ is due to a capacity in the cerebral cortex for organizing definitely successive systems of paths of discharge. Does he then see any radical reason why the *special time-order* of the ‘ideas’ in any case whatever of ‘association’ may not be analogously explained? And if not, may he not go on to admit that the most characteristic features of our faculty of memory, of our perception of outer things, of our liability to illusion, *etc.*, are most plausibly and naturally explained by acquired organic habitudes, stamped by the order of impressions on the plastic matter of the brain? But if he will admit all this, then the diagrams of association-paths of which he preserves so low an opinion are not absolutely contemptible. They do represent the *sort* of thing which determines the order of our thoughts quite as well as those diagrams which chemists make of organic molecules represent the sort of thing which determines the order of substitution when new compounds are made.³⁶

James brought this naturalistic attitude to the study of emotion as well, and his loathing of contemporary emotion taxonomies is best explained not by Ellsworth’s invocation of his relativism but rather by their unsuitability *for the purposes of scientific psychology*:

Were we to go through the whole list of emotions which have been named by men, and study their organic manifestations, we should but ring the changes on the elements which these three typical cases involve. Rigidity of this muscle, relaxation of that, constriction of arteries here, dilatation there, breathing of this sort or that, pulse slowing or quickening, this gland secreting and that one dry, *etc.*, *etc.* . . . We should find a like variation in the objects

which excite emotion in different persons . . . The internal shadings of emotional feeling, moreover, merge endlessly into each other . . . The result of all this flux is that the merely descriptive literature of the emotions is one of the most tedious parts of psychology. And not only is it tedious, but you feel that its subdivisions are to a great extent either fictitious or unimportant, and that its pretences to accuracy are a sham. But unfortunately there is little psychological writing about the emotions which is not merely descriptive. . . . But as far as “scientific psychology” of the emotions goes, I may have been surfeited by too much reading of classic works on the subject, but I should as lief read verbal descriptions of the shapes of the rocks on a New Hampshire farm as toil through them again. *They give one nowhere a central point of view, or a deductive or generative principle* . . . Is there no way out from this level of individual description in the case of the emotions? I believe there is a way out, but I fear that few will take it. The trouble with the emotions in psychology is that they are regarded too much as absolutely individual things . . . But if we regard them as products of more general causes (as ‘species’ are now regarded as products of heredity and variation), the mere distinguishing and cataloguing becomes of subsidiary importance. Having the goose which lays the golden eggs, the description of each egg already laid is a minor matter. Now the general causes of the emotions are indubitably physiological.³⁷

What James is advocating here is an explicitly *revisionary* program for emotion theory in scientific psychology. The revision consists in rehabilitating our scientific-psychological emotion categories into a more naturalistically respectable idiom by focusing on “a deductive or generative principle.” This generative principle, for James, is the Darwinian functionalist view: emotions are perceptions of bodily change that are evolutionarily derived and predispose us to act in certain ways, and this principle encourages a reorientation away from the highly abstracted and metaphysical categories of introspective psychology and toward a focus on the physical, neural basis of the emotions. So, while James may have never offered a list of basic emotions, it may very well be that if we rehabilitated our psychological categories to his preferred functionalist and

physiological specifications, he would be more than happy to privilege these emotional categories *for the purposes of affective science*. And, crucially, all BET requires (when plausibly characterized) is that there be a privileged set of discrete emotion categories for the purposes of affective science.³⁸ Since James *wants* a privileged set of discrete emotion categories for the purposes of affective science, he is most plausibly described as a Basic Emotions theorist, though very different from classical BET. Now that I have offered a compelling third interpretation of James's emotion theory, I will close by suggesting how this interpretation of James's emotion theory can help inform contemporary debates surrounding BET.

IV. WHAT BASIC EMOTION THEORY IS AND WHAT IT IS NOT

BET has recently come under serious attack from various flavors of constructionism with regard to emotion. The central claim of those critical of BET is that, much like James's critics before them, BET cannot account for the radical heterogeneity of our emotion categories. Recent empirical evidence, they contend, suggests that the only thing that fear, anger, happiness, love, etc. have in common is the fact that we label them as such.³⁹ BET, in its positing of neural or behavioral signatures for each emotion category renders itself empirically inadequate because none of these signatures have been forthcoming in practice. What this interpretation offers the contemporary emotion debate is a way for Basic Emotion theorists to respond to this critique of the research program. Since James explicitly disavowed interindividual neural/behavioral signatures of the kind that later theorists claimed as constitutive of emotion, it allowed him to account for the variability of emotion categories by uniting them under their common function. In fact, many recent proponents of BET have made this move, classifying basic emotions in terms of "action readiness,"⁴⁰ which is a broad enough category, by virtue of its functional nature, to account for much of the variability constructionists see as problematizing BET.

Another advantage of this interpretation of Jamesian theory is that it allows for both the classical Jamesian and Neo-Jamesian accounts to adequately explain how perceptions of bodily feeling are

action-motivating. A traditional critique of Jamesian and Neo-Jamesian theory is that it dispenses with or severely reduces the action-guiding role of emotions. This is a legitimate critique on both the standard reading and Ellsworth's reading of James's emotion theory insofar as it seems glaringly obvious that emotions do motivate action, and yet James strenuously denied their action-motivating character.⁴¹ But once it is made clear (and it hasn't been sufficiently made clear in previous James scholarship) that James eventually discarded this portion of his theory in exchange for emotions acting as *functional predispositions to act*, a major critique of Jamesian and Neo-Jamesian theories must be reformulated (by showing how predispositions to act still aren't motivating *enough* for a plausible theory of emotion) or thrown away entirely. Finally, an understanding of James's aspirations for turning psychology into a properly natural science might motivate us to keep BET, given its fecundity as a psychological research program over the past century, even in the face of difficult critique. We might take the lack of unification in our emotion categories not as evidence of absence of unity but rather as a plea for better research methods, more fine-grained terminology, and more clever experiments.⁴² These observations resonate nicely with some of James's opening and closing words in "A Plea for Psychology as a 'Natural Science'":

Psychology, indeed, is to-day hardly more than what physics was before Galileo, what chemistry was before Lavoisier. It is a mass of phenomenal description, gossip, and myth, including, however, real material enough to justify one in the hope that with judgment and good-will on the part of those interested, its study may be so organized even now as to become worthy of the name of natural science at no very distant day. I hoped that my book would leave on my readers an impression somewhat like this of my own state of mind. I wished, by treating Psychology *like* a natural science, to help her to become one . . .

It seems to me, finally, that a critic of cerebralism in psychology ought to do one of two things. He ought either to reject it in principle and entirely, but then be willing to throw over, for example, such results as the entire modern doctrine of aphasia—a very hard thing

to do; or else he ought to accept it in principle, but then cordially admit that, in spite of present shortcomings, we have here an immense opening upon which a stable phenomenal science must some day appear. We need not pretend that we have the science already; but we can cheer those on who are working for its future, and clear metaphysical entanglements from their path. In short, we can aspire.⁴³

CONCLUSION

I began this paper with a profile of the traditional view of William James as a traditional Basic Emotions theorist, and then I contrasted it with an alternate reading found in a series of papers by Phoebe Ellsworth that characterized James's emotion theory as a precursor to cognitive-appraisal theories of emotion. I argued that both accounts misread James and offered a third account of James as a proponent of a *functionalist, evolutionary* version of BET. In forwarding this third reading, I made four textual claims: (1) That James is best characterized as a proponent of BET; (2) That James thought the only necessary condition for emotion generation was the perception of internal bodily feelings; (3) That James's BET individuates emotion categories by their evolutionary and functional roles; and, (4) that Ellsworth and others are wrong in asserting that James was loath to offer a privileged taxonomy of emotion categories for the purposes of scientific psychology. After discussing each of these claims in detail, I subsequently suggested ways in which this interpretation could be profitably applied to contemporary debates concerning BET and Jamesian theories of emotion by problematizing constructionist critiques of both. I closed with the observation that, whatever its ultimate status, BET has proven to be a useful theoretical program, and that this reading of James might encourage us to treat its deficiencies not as reasons to discard it but, more aspirationally, as reasons to improve it. Here is to hoping.

BIBLIOGRAPHY

- Adolphs, Ralph. "How Should Neuroscience Study Emotions? By Distinguishing Emotion States, Concepts and Experiences." *Social Cognitive and Affective Neuroscience* 12, no. 1 (January 2017): 24–31.
- Barrett, Lisa Feldman. "Are Emotions Natural Kinds?" *Perspectives on Psychological Science* 1, no. 1 (2006): 28–58.
- . *How Emotions Are Made: The Secret Life of the Brain*. Boston: Mariner Books, 2018.
- Battman, Wolfgang, Stephan Dutke and Wolfgang Schönplflug. *Processes of the Molar Regulation of Behavior*. Lengerich, Germany: Pabst Science Publishers, 1996.
- Baumeister, Roy F., Kathleen D. Vohs, C. Nathan DeWall, and Liqing Zhang. "How Emotion Shapes Behavior: Feedback, Anticipation, and Reflection, Rather Than Direct Causation." *Personality and Social Psychology Review* 11, no. 2 (May 2007): 167–203.
- Damasio, Antonio R. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G.P. Putnam, 1994.
- . *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. Boston: Mariner Books, 2000.
- Ellsworth, Phoebe C. "William James and Emotion: Is a Century of Fame Worth a Century of Misunderstanding?" *Psychological Review* 101, no. 2 (1994): 222–29.
- . "Basic Emotions and the Rocks of New Hampshire." *Emotion Review* 6, no. 1 (2014): 21–26.
- Ekman, Paul. "Expression and the Nature of Emotion." In Scherer and Ekman, eds., *Approaches to Emotion*, 319–43.
- Frijda, Nico. *The Laws of Emotion*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, 2007.
- Fodor, Jerry. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.
- Izard, Carroll. "Facial Expressions and the Regulation of Emotions." *Journal of Personality and Social Psychology* 58, no. 3 (1990): 487–98.
- James, William. "A Plea for Psychology as a 'Natural Science'" in *Essays in Psychology*, 270–77.
- . *Essays in Psychology*. Cambridge, MA: Harvard University Press, 1983.
- . "The Feeling of Effort," in *Essays in Psychology*, 83–124.

- . “The Physical Basis of Emotion,” in *Essays in Psychology*, 299-314.
- . “What Is an Emotion?” in *Essays in Psychology*, 168-87.
- . *Principles of Psychology* (3 Vols.). Cambridge, MA: Harvard University Press, 1981.
- Prinz, Jesse. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press, 2006.
- Reisenzein, Rainer. “Emotional Action Generation.” In *Processes*. Edited by Wolfgang Battmann, et al., 151–65.
- Reisenzein, Rainer and Achim Stephan. “More on James and the Physical Basis of Emotion.” *Emotion Review* 6, no. 1 (2013): 35–46.
- Russell, James A. “Core Affect and the Psychological Construction of Emotion.” *Psychological Review* 110, no. 1 (2003): 145–72.
- Scherer, Klaus, and Paul Eckman, eds. *Approaches to Emotion*. Abingdon: Psychology Press, 2017.
- Schachter, Stanley, and Jerome E. Singer. “Cognitive, Social, and Physiological Determinants of Emotional State.” *Psychological Review* 69, no. 5 (1962): 379–99.
- Teroni, Fabrice and Julien A. Deonna. “Getting Bodily Feelings into Emotional Experience in the Right Way.” *Emotion Review* 9, no. 1 (2016): 55–63.
- Tomkins, Silvan S. *Affect, Imagery, Consciousness, Vol. 1: The Positive Affects*. Oxford, England: Springer, 1962.
- Weiner, Bernard. *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*. New York: Guilford, 1995.

NOTES

¹ Ellsworth, “William James and Emotion,” 222.

² For some representative examples, see Barrett, *How Emotions Are Made*; Damasio, *Descartes’ Error*; Prinz, *Gut Reactions*; and, Ekman, “Expression and the Nature of Emotion.”

³ There is some tension between James’s views about interpretation, argument, and predication and the historical-exegetical enterprise of reconstructing his view in a rigorous way. James was always skeptical of definite pronouncements on what *x* phenomenon *really* is or is not. That said, his enormous influence in emotion theory and the number of different theorists who jockey both critiquing and lauding his influence invites a

careful analysis regarding what he said, to the best of our ability. While a rigorous reconstruction may leave something out, I think James himself would understand the usefulness of this particular endeavor for the ultimate purpose of clarifying his historical influence on emotion theory.

⁴ James, "What Is an Emotion?," 170.

⁵ It is important to note that James's theory never states that *all* bodily feelings are emotions; rather, it is the case that emotions form a special subset of bodily feelings that are defined by their functional, adaptive value.

⁶ James, 170–1.

⁷ James, 175.

⁸ See Tomkins, *Affect, Imagery, Consciousness*.

⁹ See Ekman, "Expression and the Nature of Emotion."

¹⁰ See Izard, "Facial Expressions and the Regulation of Emotions."

¹¹ See Damasio, *Descartes' Error*.

¹² Modules are a term in cognitive science and evolutionary psychology for domain-specific, task-specific, independently organized processing architectures instantiated in the brain. See Fodor, *The Modularity of Mind*, for the canonical expression of the modularity hypothesis, though most discussion of modularity in evolutionary theory since Fodor's publishing of *The Modularity of Mind* have significantly weaker conditions for modularity than Fodor demands.

¹³ Ekman, perhaps the most prominent contemporary Basic Emotion theorist, posits six universal emotion expressions and therefore six basic emotions: anger, happiness, surprise, sadness, disgust, and fear. Any and all other emotion states are constituted by combinations of these six basic emotions. Others identify more or less, depending on their commitments, usually ranging from between six (Ekman) and twelve (Izard).

¹⁴ See Schachter and Singer, "Cognitive, Social, and Physiological Determinants of Emotional State."

¹⁵ See pages 2-3, above.

¹⁶ Ellsworth, "William James and Emotion," 228.

¹⁷ James, "The Feeling of Effort," 85.

¹⁸ James, *The Principles*, 1072.

¹⁹ James, "The Physical Basis of Emotion," 303–4.

²⁰ James, "The Feeling of Effort," 105–6.

²¹ See again James, "What is an Emotion?," 170–71.

²² James, “The Physical Basis of Emotion,” 302.

²³ James, 302.

²⁴ And the more recent interpretation of Reisenzein and Stephan, “Emotional Action Generation,” who I take to be the most sensitive interpreters of James’s emotion theory currently working.

²⁵ James, “What is an Emotion?,” 173.

²⁶ Ellsworth, “William James and Emotion,” 225.

²⁷ Perhaps objectless emotions involve a general appraisal of the state of the world at large. But the function of appraisals as motivators to action in appraisal theories seems to lose its force when applied to objectless emotions, which generally are not nearly as motivationally strong as intentional emotions. The claim that objectless emotions do involve appraisals seems, at least to this reader, ill-motivated and ad hoc.

²⁸ James, “The Physical Basis of Emotion,” 305–6.

²⁹ Ellsworth, “William James and Emotion,” 223.

³⁰ This point is not well appreciated, and Damasio, in *Descartes’ Error* and *The Feeling of What Happens*, cites it as one of his main points of departure from traditional Jamesian emotion theory. But, in fact, they are not disagreeing at all, except perhaps on the relative importance of musculoskeletal feedback.

³¹ Ellsworth, “Basic Emotions,” 24–25.

³² James, “What is an Emotion?,” 170.

³³ James, “What is an Emotion?,” 170.

³⁴ James, *The Principles*, 1097.

³⁵ James, “A Plea for Psychology,” 277.

³⁶ James, 275–77.

³⁷ James, *The Principles*, 1063–65. Emphasis mine.

³⁸ Proponents of BET have not emphasized this point enough, in my opinion.

³⁹ See Barrett, “Are Emotions Natural Kinds?” and *How Emotions Are Made*, and Russell, “Core Affect,” for some now very influential critiques.

⁴⁰ See Frijda, *The Laws of Emotion*; Teroni and Deonna, “Getting Bodily Feelings”; and Adolphs, “How Should Neuroscience Study Emotions?” for representative examples of functionalist takes on BET.

⁴¹ Most recent empirical evidence for the proposition also seems to corroborate our strong intuition. See Baumeister, et. al, “How Emotion

Shapes Behavior”; Reizenstein, “Emotional Action Generation”; and Weiner, *Judgments of Responsibility*, for the relevant data.

⁴² Indeed, recently utilized statistical analysis techniques, like Multivariate Pattern Analysis, have found relatively robust statistical correlations to the level of significance, though the use of these methods in emotion theory remains controversial.

⁴³ James, “A Plea for Psychology,” 270, 276–77.