

WHAT IS MISSING IN THE “BASIC EMOTION VS. CONSTRUCTIONIST” DEBATE?

PRAGMATIST INSIGHTS INTO THE RADICAL TRANSLATION FROM THE EMOTIONAL BRAIN

Fausto Caruana

University of Parma

fausto.caruana@unipr.it

ABSTRACT: The theoretical bases of affective neurosciences are going through a turbulent period. In answering the apparently naïve question “what is an emotion?” scholars supporting a basic emotion theory (BET) argue that emotions are discrete mental entities localized in the brain in the form of affect programs. Recently, psychological constructionists (PC) have attacked the stronghold of BET by defending an anti-essentialist position inspired by James’s theory of emotion, arguing that emotions are cognitive-based and culturally constructed phenomena, and not natural kinds. Who is right? To answer this question, I will highlight some theoretical pitfalls related to the shared assumption that brain-imaging results are sufficient to delimit the boundaries of the playing field. In contrast, I will capitalize upon electrical stimulation studies, which demonstrate the existence of discrete emotions but, most importantly, emphasize the interpenetration of emotional experiences and expressions. The latter observation links emotions to action tendencies, in line with a longstanding tradition that goes back to the criticisms made by Dewey and Mead of James’s theory of emotion, and survives today in contemporary embodied and enactive theories. This approach proves to be the best equipped to account for the fullness of available empirical data. Who is right, then? Neither of the two. The time has come for a theoretically sound, and empirically grounded, alternative to both BET and PC.

1. The deconstruction of emotions: from basic emotions to constructionism

Philosophy of science teaches that a successful strategy in the early stages of any biological research is to identify an ordinary phenomenon, and localize it in a unique mechanism (Bechtel 2008; Bechtel and Richardson 2010). In the field of neuroscience, direct localization has led to the view that language processing is encoded by a single brain area, Broca’s area in the left inferior frontal gyrus, or that vision depends on a single neural mechanism located in the occipital lobe. Albeit direct localization rarely turns out to be correct – in the two cases mentioned, for example, it was not – this strategy has considerable heuristic value in opening up new lines of research. Notably, the naïve attempts mentioned

above made it possible to discover to what extent language and vision depend on complex mosaics of structurally and functionally different mechanisms. In recent years, something similar is taking place in emotions research.

Direct localization came in affective neuroscience between the 30s and the 50s. In that period, scientists believed they had identified the seat of a unique neural mechanism underlying all emotions, the “limbic system.” A paradigmatic example is the well-known Papez circuit, a cortico-subcortical circuit that – according to its creator – constituted the neural basis of our emotional life.¹ Starting from the 70s, however, the idea that a single mechanism could be responsible for all emotions was gradually replaced by the view that different emotions are supported by different mechanisms. Inspired by Charles Darwin’s work on the expression of emotions, it was also suggested that such mechanisms are identifiable in a limited number of basic emotions,² typically denoted by terms borrowed from everyday language, and on which high-level emotions are built. This view has been referred to as basic emotion theory (BET). Following BET, basic emotions are universal, or substantially unchanged among different cultures, and triggered by corresponding affective programs, neural mechanisms regulating complex, coordinated and automatic responses (Griffiths 1997; Ekman 1999). Affect programs are complex, as they orchestrate multiple variables including facial expression, body posture, and autonomic and neuroendocrine responses. They are coordinated, as they organize specific activation sequences. Finally, they are automatic, as they trigger reflex-like responses.³ The concept of affect program made it possible to tackle the issue of the basis of emotions at the computational and neural levels, and

¹ For an historical overview of these studies, see LeDoux (1999).

² It must be noted, however, that Darwin never spoke about “basic” emotions.

³ In this sense, affect programs are similar to the modules described by Fodor (1983). It should be noted, however, that both Ekman and Griffiths consider emotional responses to be more complex than reflexes and homeostatic drives (Griffiths 2004).

numerous studies have been undertaken to identify the cortical or, more often, subcortical regions housing these programs.

At the end of the nineties, in a work that soon became a classic, Paul Griffiths (1997) went even further. Griffiths made a case for eliminating the concept of “emotion” from the psychological and neuroscientific vocabulary, claiming that this category does not denote a natural kind, as discoveries concerning specific emotions cannot be extended, by induction, to other emotions. Knowledge concerning the mechanisms of, say, fear is not helpful to understand the mechanisms underlying other emotions, such as happiness or surprise. In addition, what we know about the mechanisms of fear is not even transferable to all instances of “fear”: think for example how different the fear is that is evoked by someone who suddenly jumps out of the dark, compared to fear of flying, or fear of a stock market crash. Thinking that these commonsense concepts are scientifically useful only because it is possible to provide a precise conceptual definition is an error comparable to that made by Aristotelian physicists, who mixed all objects above the orbit of the moon in a single category, that of superlunary objects. Although it is true that there are objects above the orbit of the moon, i.e. superlunary objects, nothing follows from the fact that an object is superlunary. In the same way, nothing follows from the fact that a psychological state is an emotion. Griffiths’ conclusion is that, since there is not a unique psychological mechanism underlying different emotions, this concept should be eliminated from our scientific vocabulary and replaced by three different independent concepts: affect programs, high-level cognitive emotions, and socially constructed emotions.

As might be expected, therefore, concepts borrowed from ordinary language on emotions have gone through a gradual deconstruction. Over time, also, Griffiths and others have moved further away from the classical perspective, arguing that the concept of basic emotion

can be saved only by sacrificing a number of assumptions. These assumptions include *essentialism* (if some properties are reliably and systematically grouped, then the objects exhibiting these properties constitute a natural kind; there is no need to postulate a common “essence”), and the view that emotional categories can be identified by concepts taken from our ordinary language (Scarantino and Griffiths 2011; see also Scarantino 2015)⁴. In this view, emotion concepts can still explain regularities in emotional behavior, without committing reference failure (they refer to something – in contrast to concepts like “phlogiston”), or projectability failure (it is possible to project information from some members to other members of the same category – in contrast to concepts like “superlunary object”). A key point of BET, however, survives: basic emotions are different from high-level emotions (say, shame) – either because complex emotions are constructed by assembling basic emotions or because, following Griffiths, complex emotions are completely different from basic ones.

The last brick in the deconstruction of emotions has been removed by psychological constructionists (PC). Disciplines such as quantum mechanics, relativity, plate tectonics or evolution by natural selection has taught us how the truths of science can go far beyond prescientific commonsense, the role of which ran out early on in research. According to PC, the view that there are regularities related to basic emotions would be one of the prescientific insights to be abandoned (Russell 2015). In this view, what is encoded by our putative emotional brain regions are not specific emotions but, in contrast, primitive psychological or physiological components that make them up: disgust, or fear, would be nothing but a particular combination of these primitive components, held together by binning processes influenced, *inter alia*,

⁴ Further reflections have emphasized that the concept of “basicness” can be independently understood at the conceptual, psychological or biological levels (Ortony and Turner 1990; Griffiths and Scarantino 2009).

by language and culture. Most importantly, each of the primitive components that make up our emotions is not inherently emotional, as it also contributes to other, non-emotional, cognitive functions. Last but not least, the view that some emotions are more basic than others, and that high-level emotions are constructed by assembling basic emotions, loses all its meaning. According to Lisa Feldman Barrett, albeit it is an incontrovertible fact that we often (though not always) experience emotions as discrete and separate events, or that we categorize other expressions in terms of basic emotions, there is no scientific support for the hypothesis that these behaviors depend on anatomically and functionally distinct brain circuits. The various occurrences of, say, “fear” do not share any phenomenological, behavioral or neurophysiological substrate. Consequently, a theory of emotions should deal with the following facts: (a) there is no one-to-one correspondence between putative basic emotions and physiological, expressive, phenomenological, or behavioral traits; (b) variability and context-dependence are the norm, not the exception.

2. The “BET vs. PC” Debate

While supporters of both BET and PC now agree that emotions, as they appear in our ordinary language, do not mirror our brain ontology, these two parties disagree on whether emotions are “natural kinds” or not. To put it simply, the debate concerns the question of whether the concept of emotional “basicness” still has any scientific value or not (Griffiths 2004; Barrett 2006; Barrett et al. 2007; Izard 2007; Scarantino and Griffiths 2011). Endorsing one or the other positions has concrete consequences. As an example, if emotions are natural kinds, the same emotions could be more or less present in different cultures and, most importantly, in different animals. As a consequence, we should be allowed to study emotional circuits in animals using them as reliable models for testing psychiatrically effective drugs.

Constructionist claims are typically supported by functional magnetic resonance (fMRI) data. More specifically, many of these studies are based on meta-analyses (Kober et al. 2008; Lindquist et al. 2012b, 2016; Brooks et al. 2016), i.e. studies in which a large number of experimental data, collected from different published works, performed by different groups, is re-analyzed and pooled together. Meta-analyses well fit the aims of PC. By meta-analyses, in fact, it has been possible to demonstrate that studies originally interpreted as supporting a one-to-one mapping between basic emotions and brain areas are also in line with PC predictions. Pooling many studies together, indeed, it appears that, very often, the same areas contribute to different emotions, and individual emotions activate different areas. Accordingly, the discovery that brain regions typically associated with, say, disgust, are also activated by other emotions, say fear, could be employed to argue that such brain regions encode emotion-unspecific “core affects” (Lindquist et al., 2012; see below for a replication at this point). Some supporters of BET have countered by resorting to most sophisticated fMRI data analysis techniques, such as multivoxel pattern analysis (MVPA), showing that emotion-specific neural patterns are not distinguishable at the level of the individual brain area, but rather at the network level (Kragel and LaBar 2016; Saarimäki et al. 2016). Constructionist responses, however, were not long in coming, setting the stage for a potentially endless debate (Clark-Polner et al. 2016).

Here I question the shared, albeit tacit, assumption that we can entrust a big chunk of our understanding of emotions to correlation studies, such as fMRI studies, without consequences. These studies are of crucial importance, of course, but they also have a number of basic limitations that cannot be overlooked. I fear that the price to pay is potentially high. How can we evaluate the hypothesis that emotions are useful movements (Dewey 1894), actions (Döring 2014), action tendencies (Frijda 1987), predictions of action tendencies (Lowe and

Ziemke 2011), affiliative communicative displays (Proust 2016), action-oriented embodied representations (Hufendiek 2016), skillful engagements with the world (Griffiths and Scarantino 2009), affordances (Griffiths and Scarantino 2009; Hufendiek 2016), or activations of action systems (Panksepp 2005), if tests on emotions are entirely conducted isolating the subjects within fMRI scanners, where actual emotional circumstances cannot be realistically simulated, and where true emotional reactions cannot be instantiated? I will turn to this important point in the second part of the paper. Before focusing on this technical weakness of correlative data, I want to focus your attention on an epistemological one.

3. The radical translation from the brain

Functional MRI studies are *correlation* studies. Correlation studies investigate how a part of the brain reacts to stimuli presented by an experimenter. Notably, these studies constitute a greater part of contemporary cognitive neuroscience including – beside fMRI – EEG, MEG and single neuron recordings. Complementary to correlation studies are *interference* studies, i.e. studies in which the experimenter evaluates how interfering with the normal brain functions impacts on the overt behavior of a subject. Interference studies include electrical, chemical, and magnetic stimulations, as well as real and virtual brain lesions. As I will show hereafter, the interpretation of correlation studies, including fMRI studies, suffers from epistemological problems of the kind described by Quine (1960) in the famous experiment on the “radical translation.” In contrast, these problems do not affect interference studies, to which I will turn my attention in the second part of the paper. Since Quine’s line of thought is very well known, a very brief reminder should suffice.

Quine describes the case of a linguist having the task to translate the language of a tribe that has never had contact with English speakers. The linguist is assisted by a native, who allows him to observe his verbal reactions

to different situations. Observing the scurrying of a rabbit, the native produces a statement, which the linguist transcribes as “gavagai.” However, this hides a problem: the statement “gavagai” can be translated with “rabbit”, but also with a number of alternative possibilities, such as “un-detached rabbit part”, “rabbithood”, and so on. All these alternatives are perfectly plausible. Which one, if any, is the *true* translation? Unfortunately, increasing the amount of empirical evidence cannot be of help in solving the ambiguity, all the hypotheses of translation being equally appropriate to explain further occurrences of “gavagai.” Starting from these alternatives, manuals for translating one language into another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with one another (Quine 1960, 26). To put it simply, one can never be *truer* than the other. Strikingly, however, every translation affects in a very peculiar way our interpretation of the society and culture of the tribes. And so far this is what Quine says.

There are, I think, many similarities between the example described by Quine and the neuroscientific inquiry. The standard procedure in place during correlation studies, including fMRI studies on emotions, mirrors the case of a linguist presenting the native with rabbits, and recording his verbal responses. Similar to the case of the linguist, the neuroscientist involved in correlation studies tries to translate the native language by ostension, i.e. presenting external stimuli (e.g. emotional situations) to the native and recording his spontaneous responses: the brain’s “responses.” In this neuroscientific version of Quine’s story, the neuroscientist plays the part of the linguist while the neuron is the native speaker.⁵ The occasional sentence “gavagai” is expressed by a “significant increase in neural activity”, and the stimulus meaning “rabbit” can be

⁵ Actually, the spatial resolution of fMRI is not that good, so “voxel” would be more appropriate than “neuron.” But let us postpone this issue for the moment.

replaced by any experimental stimulus. How can the neuroscientist be sure about what the neuron “perceives” during stimulus presentation, if not by making very strong (arbitrary) assumptions about what we might call the ontology of the neuron?⁶ In accord with the original account, the reference is inscrutable.

Compared to the Quinean linguist, however, the neuroscientist faces an additional problem: while some normative principles are available to the linguist interpreting the native’s sentence, the same is not true for the neuroscientist. For instance, the Principle of Charity constrains the interpreter to maximize the rationality of the subject: “[...] the more absurd or exotic the beliefs imputed to a people, the more suspicious we are entitled to be of the translations; the myth of the prelogical people marks only the extreme” (Quine 1960, 68). As has recently been noticed, the Principle of Charity depends heavily on empathy, an innate shared sense of similarity between speakers, which remains the guiding strategy of the linguist to interpret the native language (Baghramian 2016). This empathic step, which allows for learning by ostension, introduces, in contrast, a crucial problem for the neuroscientist.

Learning by ostension has two requirements: first, the native (or a teacher) and the linguist (or a child) should receive the same stimulation; second, the two subjects must categorize the perceived reality using similar conceptual frameworks. In other words, it requires a common ontology. During correlation studies on emotion, e.g. based on the presentation of emotional situations or faces depicting basic emotions, both the neuroscientist and his peculiar native – a nerve cell, or a brain area – receive the same stimulation, satisfying the first postulate. However, the implicit standard of similarity, which plays such a crucial role in the case of the linguist or a child, is clearly unavailable to the neuroscientist. Following Quine, “people have to be in

substantial agreement, however unconscious, as to what counts as similar if they are to succeed in learning, one person from another, when next to assent to a given observation sentence. Here, then, is an irreducible kernel of relativism: all sensory evidence as reflected in observation sentences is relative to the neural organization that determines what different triggerings of nerve endings will favor the same response. Subjects radically at odds in this neural way could never learn observation sentences or anything else from one another. Our training even of a dog, horse, bear, seal, or elephant hinges on a conformity of his inarticulate similarity standards to our own” (Quine 1984, 293-294). In our case, however, the subject is not a person, or an animal, but a nerve cell. Too bad. In fact, if I am correct, this analysis leads to the following conclusion: in correlation studies, each assignment of meaning to neural activity is a kind of interpretation and, following Quine, faced with multiple interpretations, any interpretation is *true*r than the other. There are no deeper truths to look for.

A “proof of concept” could be of some help to convince the reader that this is a very concrete problem. Imagine a macaque presented with pictures of animals, depicted on a monitor screen. During the presentation of the stimuli, the electrical activity of a neuron is recorded, and aligned to the stimulus onset. During the presentation of a tiger, the neuron increases its firing rate, in line with the view that, the tiger being a predator, the macaque’s visual system has been equipped with special tiger-detectors by evolution. The increase in firing rate acts as an affirmative statement uttered by our native, the nerve cell and, accordingly, the neuroscientist indicates the cell as a tiger-selective neuron (similarly, we can imagine Quine’s linguist presenting the native with different animals and listening to the native saying “gavagai” only when a tiger is presented). Subsequently in our story, the neuroscientist adds some experimental controls, deconstructing the original image of the tiger into

⁶ I always felt embarrassed talking about the ontology of the neuron, until I found out that Dennett spoke about the ontology of the elevator (Dennett 2013).

subcomponents and discovering that the same neuron is also active during the presentation of a black and white image of a tiger. In addition, the scientist presents a composition of white rectangles, replacing the tiger's orange stripes, and black rectangles, replacing the black ones. The neuron fires again. In contrast, we can imagine that the neuron stops firing when the entire picture is replaced by separate black or white rectangles. From a scientific point of view, this is a well-controlled experiment, demonstrating that the “ontology” of this specific neuron does not include “colored tigers’ heads.” I did not invent anything: this is indeed the result of a classic study by Tanaka (1993). However, what can we say about the “ontology” of this neuron? Similar to the case described by Quine, this neuroscientific result can be interpreted in very different ways: the constructionist will say that “the analysis of visual stimuli effected by inferotemporal neurons is not complex enough to specify a particular biological object on the basis of a single cell discharge” (Jacob and Jeannerod 2003, 59), while the evolutionist could still argue that the neuron is selective to tiger-like patterns, also including false positives (black-white rectangles superimpositions) for evolutionary reasons. Further studies will not be of any help in discriminating the two interpretations.

This, in my view, is what is happening in the debate between BET vs. PC: the “essentialist” BET neuroscientist will publish an article on the “cortical representation of rabbithood”, while the constructionist will reply with a commentary on the “un-detached rabbit part.”⁷

While correlation studies require the interpretation of the response of a nerve cell (or brain region), electrical stimulation (and other interference) studies are in a privileged position. Let us consider electrical

stimulation: the delivery of small quantities of electricity in a specific part of the brain elicits specific movements or modification of the overt behavior in the experimental subject. The task of the experimenter is to interpret this behavior, a natural task that we perform all the time. In addition, when performed on humans, electrical stimulation allows the scientist to take advantage of the first-person experience provided by the experimental subject, who can verbally report on the sensations elicited by the stimulation – allowing for the Heterophenomenological approach described by Dennett (1991, 2003). Suddenly, the epistemological limitations mentioned above are somewhat defused, leading translation troubles back to a more common situation: undetermined but, nevertheless, assisted by some normative principles – as in the case of the Quinean linguist.

4. Beyond correlation: what stimulation tell us about emotions

Stimulation studies have important advantages over correlation studies, including access to the behavioral responses elicited by the stimulation. The relationship between subjective experiences and actions is of particular interest when discussing emotions. How did a bear met in the woods differ from one watched in the zoo? Asking this question, Dewey (1894, 1895) noticed that the main difference between a bear met in the woods (i.e. something to be run from) and one watched in the zoo is that only the former is something to be avoided, and consequently explained “by reference to movements having some use.” In contrast, the experience of the bear in the zoo involves no racing of the heart or increase in respiration, because the individual does not contemplate escape (see also Ward and Throop 1989; Backe 2001). However, whether actions are the stuff the emotions are made of, as Dewey and others suggest (Mead 1895; Frijda 1987; Panksepp 2005; Lowe and Ziemke 2011; Caruana and Gallese 2012; Döring 2014; Gallese and Caruana 2016;

⁷ Notably, Dennett and Millikan consider the indeterminacy of radical translation really negligible in practice: in their view, it is extremely unlikely that there could be two different pathways that produce two interpretations of the same value, globally indeterminate and radically different (Dennett 2013, Ch.30). As should be clear, I disagree with them on this.

Hufendiek 2016), is something that fMRI studies would hardly be able to clarify. Beside the epistemological problems mentioned above, in fact, there are also very important technical limitations.

When we are locked in an fMRI scanner, we can talk about emotions, recognize them in other individuals, categorize them or, at most, remember our own past emotional states – but we can hardly have vivid emotions. Think of mirthful laughter. Correlative (fMRI, EEG, MEG or similar) studies cannot easily tackle two “hard problems” posed by this emotional expression: (1) “Laughter is a social behavior that virtually disappears in isolated people being scrutinized in a laboratory setting” (Provine 2000). (2) Laughter involves facial grimaces, vocalizations and postural movements, which render it impossible to study it within an fMRI scanner. Not surprisingly, the majority of correlative data on laughter are based on the visual or auditory perception of laughter produced by others. The same is true for many other emotions: whatever leads you to laugh out loud, or makes you scream with fear, or disgusts you until you vomit, will remain outside the laboratory. The number of imaging studies on emotion in which nobody feels any emotion is astonishing. Panksepp is probably right: “correlative approaches, such as brain imaging in humans or psychophysiology, are not strong enough to judge what is ‘basic’ in the basic emotions. [...] We must be selective in choosing which techniques are appropriate for addressing affective questions. For instance, one should do their best to get people into strong affective states” (Panksepp 2007, 282-283).⁸ Hereafter, I will review some examples.

⁸ Not to mention the fact that, for some time now, this technique has been questioned because of its alarming propensity to false positives (Logothetis 2008; Bennett et al. 2009; Eklund et al. 2016). And to say nothing of the uncomfortable reality expressed by the following observation of Panksepp, who gives voice to a distrust shared by many physiologists: “With regard to modern brain imaging, simply consider the fact that the pseudocolor statistical maps of neural activity changes (rarely more than a few percentage points different than baseline levels) hide vast oceans of neurophysiological and neurochemical activities, with multiple functional systems interpenetrating [...] and with overlapping,

4.1. Disgust: from facial grimaces to disgust-behavior

According to constructionists’ meta-analyses, the insula is active during the experience of disgust, but also during other emotions – in particular, sadness and fear. Moreover, fMRI studies indicate that disgust processing also recruits a variety of other areas (Phan et al. 2002). According to PC, therefore, studies showing a correlation between insula activity and disgust processing cannot say the last word on whether disgust is a discrete emotion anchored to the insula’s circuitry. After all, they continue, disgust may not be a discrete emotion, and discrete emotions may not exist at all. Stimulation data, however, tell a different story.

A few years ago we found that, in the primate, the electrical stimulation of a specific sector of the anterior insula evokes facial expressions that mimic the natural expression of disgust (Caruana et al. 2011; Jezzini et al. 2012). This motor response was accompanied by a decrease in heart rate – consistent with the view that disgust experience is linked to vagal activity and, consequently, to an increase in bradycardic tone. In addition, if the stimulation was delivered during spontaneous feeding behavior, the elicited response included – besides the production of disgusted facial expressions and heart rate modulation – complex disgust-behavior consisting in the refusal of food: throwing away food already on hand, or spitting out the food in the mouth. The stimulation of other regions of the insula, only a few millimeters away, evoked very different responses, suggesting that the insula is a mosaic of anatomo-functional fields orchestrating different behaviors (Jezzini et al. 2012).

interacting circuits generating affective mentality. Each imaged pseudocolor region of the human brain hides enormous complexities and individual variabilities that need to be considered. What we can surely say of the typical final products of most averaged results from human brain imaging studies, consisting of isolated islands of arousal, is that ‘the brain does not work that way’ (Panksepp, 2007, 282).

These data, which show a systematic connection between the activation of a given brain structure and some instances of disgust, are in line with a lot of evidence from human interference studies: the stimulation of the insula in surgical patients evokes disgust and nausea, as also verbally reported by patients, and selectively alters the ability to recognize facial expressions of disgust (Krolak-Salmon et al. 2003; Isnard et al. 2004; Papagno et al. 2016). Moreover, it has been observed that, in temporal lobe epilepsy, vomiting and nausea only occur when the seizure invades the insular territories (Catenoux et al. 2008). Altogether, interference studies show systematic involvement of the anterior insula in a discrete emotion, disgust, in such a reliable way that disgust and nausea are commonly employed by neurologists and surgeons to identify this region. Most importantly, albeit it happens that the electrical stimulation of this region may not trigger any behavioral response or subjective experience, nevertheless disgust is the only emotion that could be directly elicited here.

There is evidence that, from the behavioral point of view, oral disgust is closely linked to moral disgust, suggesting that moral transgressions depend on an expansion of the role of disgust over evolutionary time (Chapman et al. 2009; Chapman and Anderson 2012). It would be crucial in the future to understand if, at least in principle, all the several instances of disgust could be modulated by targeting this insular region.

4.2. Laughter and smiling

The hypothesis of a direct link between a sense of merriment, typically associated with laughter, and the pregenual sector of the anterior cingulate cortex (pACC), has been suggested by some fMRI studies supporting some versions of BET, but subsequently questioned by constructionist meta-analyses. As discussed above, it is likely that these studies will not be able to overcome the limitations of the radical translation from the emotional

brain, as I called it, leading to an endless controversy. We investigated the hypothesis of an involvement of the pACC in mirth and laughter production by resorting once again to electrical stimulation, in drug-resistant epileptic patients. In about ten patients, stimulation of the pACC evoked laughter and smiling. In half of these subjects, the stimulation also elicited a sense of merriment and mirth. The other half did not report any specific emotions associated with the expression evoked, although patients often reported interoceptive sensations involving the stomach or the whole body (Caruana et al. 2015). Curiously, new evidence showed that the same region is also selectively active during the observation of others' laughter, thus suggesting that this region could also play a key role in laughter contagion (Caruana et al. 2016a). As in the case of the insula, discussed before, it occurs that electrical stimulation of this region may not trigger any response or experience, but of great importance is the fact that, to date, stimulation of this region has not evoked any *other* emotional phenomenon but laughter. In this it is unlike the subgenual sector of the anterior cingulate (sACC), immediately ventral to it, which theories of basic emotions associate with sadness and which, accordingly, is now a target region for the treatment of depression by neurosurgeons (Mayberg et al. 2005). Curiously, laughter can also be evoked by stimulation of other cortical sites, including some frontal and temporal regions, and subcortical centers such as the hypothalamus (typically compromised in patients with gelastic seizures, i.e. epilepsies characterized by pathological laughter) and the nucleus accumbens (a nucleus whose stimulation evokes smiling and merriment in humans; Gibson et al., 2016) – these are centers whose stimulation evokes laughter also in rats! (Burgdorf et al. 2001; Panksepp and Burgdorf 2003; see Caruana et al., 2015, 2016b for an overview). Although the link between laughter and joy is incredibly complex and controversial – most of our daily laughter is not due to humor but to communicative reasons (Provine 2000) – the systematic link between specific anatomical structures and these expressions is a solid scientific result that cannot be overlooked.

4.3. Primate affiliative responses

In answering those who stressed the role of stimulation data in the basic emotion debate (Panksepp 2007; Caruana and Gallese 2012), supporters of PC object that stimulation of specific anatomical sites does not systematically evoke emotional responses: in some cases, for example, stimulation may be totally ineffective (Barrett et al. 2007; Lindquist et al. 2012a). This statement is partially true, but naïve. There are several reasons why this could happen – ranging from technical reasons involving the stimulation parameters to the state of wakefulness of the subject, and others. Most importantly, in line with the view that the context matters, in some cases stimulation may require some additional contextual factors. Let us consider the insula once again. The stimulation of a ventral region of the monkey insula, caudal to the region eliciting disgust, evokes lip-smacking, an affiliative social expression (Caruana et al. 2011). In our original study, this expression was evocable only if the stimulation was released while the subject was establishing eye contact with the experimenter, and was totally ineffective otherwise – thus suggesting that this region is involved in controlling affiliative behavior provided that specific social factors, modulated elsewhere, were present. Hence, the right way to interpret these results is that the production of affiliative expressions is anatomically *distributed*, but still anatomically *constrained*. One could argue, similarly, that many other triggering factors, yet to be discovered, are critical in generating certain responses to stimulation, in other areas.

4.4. Panksepp’s basic emotional systems

All the studies mentioned above only represent a selection, based on my own personal experience, among the amount of data describing emotional responses elicited from specific brain regions. A number of evidences, derived from electrical or chemical

stimulation studies in other animal models, should be added to the list, including different types of affective behaviors elicited in rodents by the stimulation of subcortical centers. Using this technique, Jaak Panksepp described seven basic emotional systems, remarkably similar in different species of mammals: SEARCH, FEAR, RAGE, SEXUAL DESIRE, CARE, SUFFERING and GAME. It is noteworthy that Panksepp identified basic emotional systems using capital letters, in order to avoid semantic misunderstandings, and that such affective systems are not identified by commonsense emotional concepts (as both BET and PC suggest). In Panksepp’s hands, these systems embody primary processes, solutions common to different animal species – beyond cultural or species-specific modulations. According to Panksepp, in fact, our daily emotions result from mixing processes of primary (affective), secondary (learning and thinking) and tertiary (thoughts of thoughts) level (Panksepp 2007; Panksepp and Biven 2012), an account also compatible with many assumptions of PC.

4.5. Stimulation data and the BET vs. PC debate

Altogether, stimulation data raise three observations. First, full-fledged emotional experiences can be directly elicited by stimulating specific brain regions. Needless to say, this evidence is not sufficient to demonstrate the existence of a one-to-one correspondence between brain regions and emotions. After all, as I mentioned, laughter could be elicited by stimulating many different regions, albeit all of them are part of a single anatomical network. By the way, there is abundant evidence that a one-to-one correspondence between structure and function is empirically untenable. Today, this model is by and large replaced with that of “pluripotentiality” characterizing the relationship between structure and function, leading to the view that every brain region plays different functional roles in different functional networks and, accordingly, that its contribution to a given function depends on its anatomical constraints, as

well as on its interaction with other structures. This phenomenon has been referred to as neural exploitation, neural re-use, or neural recycling (Gallese and Lakoff 2005; Dehaene and Cohen 2007; Gallese 2008; Pessoa 2008; Anderson 2010; Scarantino 2012). At the same time, however, stimulation data demonstrate that “anatomy matters” – some responses, and subjective experiences, are constrained by anatomy – and that the link between structure and emotion is not an arbitrary one: some specific emotions have preferential response-initiating areas.

Second, and in agreement with the previous point, these studies highlight the role that the social and natural environment plays in modulating the evoked response. In the case of insula stimulation, social variables such as establishing eye contact with another individual, or the presence of food, are needed to drive and modulate the affiliative and disgust-related responses, respectively. Similar modulatory effects, such as the presence of offspring, or individuals occupying specific hierarchical positions, have been described in past stimulation experiments on rodents (see Frijda 1987, Ch. 7.2). Altogether, these data suggest that, during electrical stimulation, external cues act as environmental or social affordances, modulating the elicited response. Stimulation data do not say the last word on the existence of “basic” emotions, nor do they encourage this hypothesis. However, they support the view that there are “discrete” emotions, coherent patterns distinguishable on the basis of neural, physiological, behavioral and expressive features (Colombetti 2009, 2014). Most importantly, they also leave space for the high variability and context dependence of emotions.⁹

⁹ Compared to the view that there are “discrete” emotions, BET implies stronger assumptions, such as the hypothesis that some emotions are more basic than others, and that complex emotions result from the integration of basic emotions. The latter view is very controversial (for overview see Colombetti 2009 and Hufendiek 2016, Ch.2.5.), and difficult defend with

Finally, stimulation studies highlight the role of action in emotions, and the interpenetration between emotional experiences and overt emotional expressions, showing that (a) stimulation of putative sensory areas, such as the insula, triggers emotional expressions, i.e. action patterns including disgust grimaces, complex disgust behavior, heart rate modulations, or affiliative responses, and (b) stimulation of putative motor centers, such as the pACC, triggers, in addition to emotional action patterns, subjective emotional experiences as well. These data vindicate a longstanding theoretical hypothesis linking emotions to action tendencies, adding a new element in the debate between BET and PC, which deserves to be deepened.

5. Stimulation data, through the lens of John Dewey

We have come to electrical stimulation as a strategy to get around the problem of the radical translation from the brain, and discovered that – besides supporting some kinds of discrete emotion theory – stimulation studies also show that the link between the overt behavioral response and the subjective experience is closer than expected. While the existence of discrete emotions is potentially compatible with many key assumptions of constructionism, the interplay between expression (action patterns) and experience is arguably the foremost discontinuity with current constructionists’ accounts. Notably, the relationship between emotions and action tendencies is one of the first things that thrill in the eye of those who study emotions using techniques that can highlight their expressive side. This point is nicely expressed by Jaak Panksepp: “Emotional feelings may largely reflect the brain activities that control spontaneous emotional action tendencies. Although the motor system has typically been conceptualized as an unconscious output system of the brain [...] there is a great deal of data to suggest that those systems do have a raw-feeling

stimulation data.

aspect when they are aroused. [...] This view [...] places a motor-action homunculus at the center of emotional life rather than any sensory homunculus. This shift of emphasis [...] helps solve a variety of troublesome philosophical problems” (Panksepp 2005, 65). In contrast, constructionists – studying people performing “cold” tasks while lying in an fMRI scanner – stress how language, culture and cognition assemble basic physiological processes: “emotions emerge when people make meaning out of sensory input from the body and from the world using knowledge of prior experiences” (Lindquist et al., 2012, 123). Even more explicit is Lisa Feldman Barrett: “discrete emotions emerge from a conceptual analysis of core affect. Specifically, the experience of feeling an emotion, or the experience of seeing emotion in another person, occurs when conceptual knowledge about emotion is brought to bear to categorize a momentary state of core affect. The conceptual knowledge that is called forth to categorize affect would be tailored to the immediate situation, represented in sensorimotor cortex, acquired from prior experience and supported by language. Categorizing the ebb and flow of core affect into a discrete experience of emotion corresponds to the colloquial idea of ‘having an emotion’” (Barrett 2006, 49). The motor side of emotions, their possible identification with action patterns or action tendencies, has a very small place in this framework.

5.1. William James and PC

In different respects, PC shares several assumptions with the theory of emotion formulated by William James in 1884. These similarities have also been explicitly recognized by constructionists on several occasions (Barrett et al. 2009; Barrett and Russell 2015). Since James’s theory of emotion was subsequently criticized for having discarded the role of action in emotion, i.e. the same criticism I am making of PC, it is worth looking into the similarities and the common limitations of the two theories.

The physiological concerns underlying James’s theory of emotion are evident from the early lines of the 1884 essay “What is an emotion?.” The paper opens by asking which of the two hypotheses about the brain basis of emotions are true: “either separate and special centres, affected to them alone, are their brain-seat, or else they correspond to processes occurring in the motor and sensory centres” (James 1884, 188). He continues: “If the latter be the case, we must ask whether the emotional «process» in the sensory or motor centre be an altogether peculiar one, or whether it resembles the ordinary perceptive processes of which those centres are already recognised to be the seat.” The fact that part of the problem starts from a purely physiological consideration is also supported by the reference to the pioneering electrical stimulation work by David Ferrier, a leading neuroscientist of his time. Indeed, one purpose of James is to criticize the view that emotions constitute a physiological category of their own, and the fact that David Ferrier failed to found purely emotional brain centers plays in favor of this view. Even when, in the *Principles of Psychology*, he deals with the issue of laughter, his position seems to be mostly devoted to criticizing the quest for a sharp border between emotional essences and sensorimotor response: if you ask someone “to imagine away every feeling of laughter and of tendency to laugh from their consciousness of the ludicrousness of an object, and then to tell you what the feeling of its ludicrousness would be like, whether it be anything more than the perception that the object belongs to the class ‘funny’, they persist in replying that the thing proposed is a physical impossibility, and that they always must laugh if they see a funny object” (James 1890, 451-452). Hence, we can reasonably argue that William James and contemporary constructionist theories of emotion agree that emotion categories are merely descriptive rather than biological kinds (Barrett et al. 2009). In other words, James and the constructionists share an anti-essentialist attitude (contra BET). The similarities do not end there.

James and the constructionists seem to agree that psychology and neuroscience should avoid postulating the existence of discrete mental entities that, in the field of affective neuroscience, correspond to discrete affect programs. In the words of James, traditional psychology “talks like one who should say a river consists of nothing but pailsful, spoonsful, quartpotsful, barrelsful, and other moulded forms of water. Even were the pails and the pots all actually standing in the stream, still between them the free water would continue to flow” (James 1890, 255). In emotion research, in fact, basic emotions *à la* Ekman – strongly criticized by constructionists – are typically depicted as static and stereotyped expressions, generated by the reification of some fragments of our continuous and uninterrupted production of communicative facial expressions: pailsful, spoonsful, quartpotsful, and barrelsful. Variability and context dependence shape our emotions, which are not the stereotyped output of an affect program, as in contrast BET suggests.

A third contact point between James and contemporary constructionism – indeed, their weak point, as it should be clear at this point – is that both of them pay little attention to the enactive dimension of emotions.¹⁰

5.2. Dewey’s and Mead’s “enactive” theory of emotion

Historically, the lack of emphasis on the enactive dimension in James’s theory of emotions, which relegates emotion to a passive dimension, was stressed by John Dewey and George Mead’s theory of emotion.

¹⁰ It must be recognized that, at the beginning of his 1884 paper, James mentions the role of actions and motor centers in emotions. Similarly, in the previous quotation from the *Principles*, he includes the “tendency to laugh” as a driving element in the definition of what is “funny.” Albeit James is recognized as the father of the “somatic theory”, and the James-Lange theory emphasizes the role of the visceral *feed-back*, and not that of the output, one could argue that his position is a little more complex than the way it is commonly popularized.

Dewey formulated his own theory to overcome the passiveness of emotions implied by James’s theory, finally suggesting that expression, affect and cognition are phases of a single functional coordination, and that agency is a core feature of emotional experience (Dewey 1894, 1895). In other words, Dewey argues that emotional experience is a parallel aspect of the processes involved in goal-oriented acts, emerging from a dynamic interaction between the agent and the social and natural environment: “the expressions of emotion are to be accounted for not by reference to emotion, but by reference to movements having some use, either as direct survivals or as disturbances of teleological coordination.” (Dewey 1895, 1). Dewey criticizes the view that emotional expressions are expressions of independent, pre-existing, internal states – an assumption that, he says, is induced by the fact that we rate certain movements as expressive when looking at them from the standpoint of the observer (Dewey 1894, 555. See also Backe 2001; Garrison 2003 for recent reconstructions). The focus on the enactive dimension of emotions was further clarified, and extended, one year later in his famous paper on the reflex arc concept in psychology, where he argued that sensory stimulus, the central connections and the motor responses shall be viewed “not as separate and complete entities in themselves, but as divisions of labor, function factors, within the single concrete whole” (Dewey 1896, 358).

But what kind of actions are emotions supposed to be? Typically, actions are classified as instrumental or communicative (Gergely and Jacob 2012). While some emotions could probably be identified as instrumental actions, such as those aimed at defending ourselves, the emotional expressions elicited by electrical stimulation in the studies mentioned above (smiling and laughter, affiliative displays, disgust) are often explained in terms of communicative gestures. This fact fits well with the interpretation of emotional gestures carried out by George Herbert Mead (1895; 1934; 2001) who, capitalizing upon Dewey’s theory, considered affiliative

emotional displays as prelinguistic forms of communication. In Mead’s words, “such beginnings of acts, and organic preparations for action, which have been called expressions of emotion are just the cues which have been selected and preserved as the means of mediating social conduct. Before conscious communication by symbols arises in gestures, signs, and articulate sounds there exists in these earliest stages of acts and their physiological fringes, the means of coordinating social conduct, the means of unconscious communication [...] They had been already naturally selected and preserved as signs in unreflective social conduct before they were specialized as symbols.” (Mead 2001, 3).

5.3. Dewey’s and Mead’s legacy today

Today, enactivists stress the enactive and action-oriented nature of emotions in a very Deweyan fashion: “we should refuse to conceive of the felt quality of anger or fear as something over and above, and thus logically distinct from, organismic ways of responding [...] the phenomenal character of an experience is not identified with, or determined by, extra ingredients over and above the dynamic responses of organisms. Instead they are identified with specific, concrete activities of sentient beings – extended ways of responding, reacting, acting and interacting” (Hutto 2012). Accordingly, an enactivist reply to the constructionist theory would mirror the original criticism moved by Dewey to James, in accord with the interplay between experience and expression highlighted by the stimulation data reviewed above.

Furthermore, contemporary echoes of Mead’s considerations concerning the primary communicative role of emotional acts survive in some theoretical account of emotions. Following Paul Dumouchel (1999), emotions are mechanisms of social coordination among conspecifics, their primary role being to signal the probability an individual will act towards other agents in

one way rather than another. For this reason, Dumouchel agrees in considering emotions as forms of actions. Hutto seems to support something not very different, criticizing the popular conjecture that the functions of emotional expressions to produce effects on other organisms are necessarily “secondary adaptations.” In contrast, he suggests, emotional expressions could be something for which social animals have selectively been directly calibrated for. Indeed, there is no reason to deny that the function of at least some of our capacities for emotional response are primarily social, not the expression of some internal state (Hutto 2006; p.30). Before him, Frijda (1987) also recommended that the function of emotional expressions is not communicating our mental states, or promoting understanding. Rather, it is communicating requests and intentions, and influencing others’ behavior. In a similar fashion, Joëlle Proust (2016) recently theorized that impulsive and habitual signals, including emotional expressions and affiliative displays, convey information to others about what a situation affords, and thereby motivate a response in the receiver. In this view, they are, first and foremost, expressive acts prompting a reactive action.

Thus, if I am correct, we can envision a direct connection between stimulation data and a pragmatist account of emotions, which originates in Dewey’s and Mead’s theory of emotion, which survive today in contemporary enactive theories, and prove to be the best equipped to account for the fullness of available empirical data – a theoretically sound, and empirically grounded, alternative to both BET and PC.

6. “Im Anfang war die Tat”: the pragmatist legacy as an alternative account of emotions

Recapitulating, in the present paper I have discussed a pragmatist inspired, and scientifically informed, approach to the neural basis of emotions, as an alternative to BET and PC. From a scientific point of view,

I capitalized upon the heuristic value of electrical stimulation studies, for two reasons. First, these studies are less prone to the epistemological pitfalls of what I called the radical translation from the brain – a neuroscientific version of Quine’s well-known thought experiment. The indeterminacy of the radical translation from the brain, I argued, underdetermines the results of correlative studies, including fMRI studies – on which the debate between BET vs. PC is based. Second, electrical stimulation can elicit strong emotional reactions, in contrast to standard correlative techniques, thus proving to be the most appropriate strategy for addressing affective questions.

The analysis of stimulation data leads to two important considerations: (a) these data demonstrate that discrete emotions can be elicited by stimulating specific brain regions, contrary to constructionists’ predictions. They also account for the fact that discrete emotions are context-dependent, as the context modulates the elicited response; (b) stimulation data emphasize the interplay between emotional experiences and expressions. This last observation links emotions to action tendencies, in line with a longstanding tradition that goes back to Dewey’s and Mead’s corrections to James’s theory of emotion. Notably, constructionists do not seem to consider the emotional expression as a constitutive aspect of emotion, sharing with James the limitations criticized by Dewey and Mead.

Thus, in the hands of James and his heirs emotional experience is *embodied*, but not *enacted*. Considering emotions as teleological habits aimed at facing environmental challenges and regulating communication, in contrast, sheds new lights on the close relationships between emotional expressions and subjective emotional experiences, described by stimulation studies. If, as I believe, such enactive view can hardly be forced within the BET-PC categories, this pragmatist alternative must be conceived as a third, independent, account of emotions. On one hand, this view redeems a theoretical tradition that associates

emotions with useful movements (Dewey 1894), actions (Döring 2014), action tendencies (Frijda 1987), predictions of action tendencies (Lowe and Ziemke 2011), action-oriented embodied representations (Hufendiek 2016), affiliative communicative displays (Proust 2016), skillful engagements with the world (Griffiths and Scarantino 2009), affordances (Griffiths and Scarantino 2009; Hufendiek 2016), or activations of action systems (Panksepp 2005). On the other hand, it offers a useful framework to account for a huge bunch of contemporary stimulation studies that, albeit forgotten in the BET vs. PC debate, proved to be the most promising ones in the task of radically translating from the emotional brain. Which is what a theory is supposed to do.

Acknowledgments

I want to thank Marco Viola and Giovanna Colombetti for valuable comments to a previous version of the article.

References

- Anderson, Michael L. 2010. “Neural Reuse: A Fundamental Organizational Principle of the Brain.” *Behavioral and Brain Sciences* 33 (4): 245–66.
- Backe, A. 2001. “John Dewey and Early Chicago Functionalism.” *History of Psychology* 4 (4): 323–40.
- Baghramian, M. 2016. “Quine, Naturalised Meaning and Empathy.” *Argumenta* 2 (3). *Italian Society for Analytic Philosophy*: 25–42.
- Barrett, Lisa Feldman. 2006. “Are Emotions Natural Kinds?” *Perspectives on Psychological Science* 1 (1). SAGE Publications: 28–58.
- Barrett, Lisa Feldman., and James A. Russell. 2015. *The Psychological Construction of Emotion*. Edited by Lisa Feldman. Barrett and James A. Russell. Guilford University Press, New York.
- Barrett, Lisa Feldman, Maria Gendron, and Yang-Ming Huang. 2009. “Do Discrete Emotions Exist?” *Philosophical Psychology* 22 (4). Routledge: 427–37.

- Barrett, Lisa Feldman, Kristen A Lindquist, Eliza Bliss-Moreau, Seth Duncan, Maria Gendron, Jennifer Mize, and Lauren Brennan. 2007. “Of Mice and Men: Natural Kinds of Emotions in the Mammalian Brain? A Response to Panksepp and Izard.” *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 2 (3). NIH Public Access: 297–312.
- Bechtel, William. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. Routledge.
- Bechtel, William, and Robert C. Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. MIT Press.
- Bennett, Craig M, Abigail A Baird, Michael B Miller, and George L Wolford. 2009. “Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument for Multiple Comparisons Correction.” *Neuroimage* 47 ((Suppl 1), S125).
- Brooks, Jeffrey A, Holly Shablack, Maria Gendron, Ajay B Satpute, Michael H Parrish, and Kristen A Lindquist. 2016. “The Role of Language in the Experience and Perception of Emotion: A Neuroimaging Meta-Analysis.” *Social Cognitive and Affective Neuroscience*, August.
- Burgdorf, J, B Knutson, J Panksepp, and S Ikemoto. 2001. “Nucleus Accumbens Amphetamine Microinjections Unconditionally Elicit 50-kHz Ultrasonic Vocalizations in Rats.” *Behavioral Neuroscience* 115 (4): 940–44.
- Caruana, F., F. Gozzo, V. Pelliccia, M. Cossu, and P. Avanzini. 2016. “Smile and Laughter Elicited by Electrical Stimulation of the Frontal Operculum.” *Neuropsychologia* 89: 364–70.
- Caruana, Fausto, Pietro Avanzini, Francesca Gozzo, Stefano Francione, Francesco Cardinale, and Giacomo Rizzolatti. 2015. “Mirth and Laughter Elicited by Electrical Stimulation of the Human Anterior Cingulate Cortex.” *Cortex* 71: 323–31.
- Caruana, Fausto, Pietro Avanzini, Francesca Gozzo, Veronica Pelliccia, Giuseppe Casaceli, and Giacomo. Rizzolatti. 2016. “A Mirror Mechanism for Smiling in the Anterior Cingulate Cortex.” *Emotion* 17 (2): 187–90.
- Caruana, Fausto, and Vittorio Gallese. 2012. “Overcoming the Emotion Experience/expression Dichotomy.” *The Behavioral and Brain Sciences* 35 (3): 145–46.
- Caruana, Fausto, Ahmad Jezzini, Beatrice Sbriscia-Fioretti, Giacomo Rizzolatti, and Vittorio Gallese. 2011. “Emotional and Social Behaviors Elicited by Electrical Stimulation of the Insula in the Macaque Monkey.” *Curr Biol* 21 (3). Elsevier Ltd: 195–99.
- Catenoix, H el ene, Jean Isnard, Marc Gu enot, J erome Petit, Claude Remy, and Fran ois Maugu iere. 2008. “The Role of the Anterior Insular Cortex in Ictal Vomiting: A Stereotactic Electroencephalography Study.” *Epilepsy & Behavior* : E&B 13 (3): 560–63.
- Chapman, H. A., D. A. Kim, J. M. Susskind, and A. K. Anderson. 2009. “In Bad Taste: Evidence for the Oral Origins of Moral Disgust.” *Science* 323 (5918): 1222–26.
- Chapman, Hanah A., and Adam K. Anderson. 2012. “Understanding Disgust.” *Annals of the New York Academy of Sciences* 1251 (1): 62–76.
- Clark-Polner, Elizabeth, Timothy D Johnson, and Lisa Feldman Barrett. 2016. “Multivoxel Pattern Analysis Does Not Provide Evidence to Support the Existence of Basic Emotions.” *Cerebral Cortex*.
- Colombetti, Giovanna. 2009. “From Affect Programs to Dynamical Discrete Emotions.” *Philosophical Psychology* 22 (4). Routledge : 407–25.
- Colombetti, Giovanna. 2014. *The Feeling Body: Affective Science Meets the Enactive Mind*. MIT Press.
- Dehaene, Stanislas, and Laurent Cohen. 2007. “Cultural Recycling of Cortical Maps.” *Neuron* 56 (2): 384–98.
- Dennett, Daniel. 2003. “Who’s On First? Heterophenomenology Explained.” *Journal of Consciousness Studies* 10 (12): 19–30.
- Dennett, Daniel. 1991. *Consciousness Explained*. Boston: Little Brown.
- . 2013. *Intuition Pumps and Other Tools for Thinking*. W. W. Norton & Company.
- Dewey, John. 1894. “The Theory of Emotion I. Emotional Attitudes.” *Psychological Review* 1 (6): 553–69.
- . 1895. “The Theory of Emotions II: The Significance of Emotions.” *Psychological Review* 2: 13–32.
- . 1896. “The Reflex Arc Concept in Psychology.” *Psychological Review* 3: 357–70.
- D oring, Sabine. 2014. “What Is an Emotion? Musil’s Adverbial Theory.” *Monist* 97 (1): 47–65.
- Eklund, Anders, Thomas E Nichols, and Hans Knutsson. 2016. “Cluster Failure: Why fMRI Inferences for Spatial Extent Have Inflated False-Positive Rates.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (28): 7900–7905.
- Ekman, P. 1999. “Basic Emotions.” In *Handbook of Cognition and Emotion*, edited by T. Dalgleish and M. Power, 45–60. Chichester, UK: John Wiley and Sons.
- Fodor, Jerry A. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. MIT Press.
- Frijda, N. H. 1987. *The Emotions*. Cambridge University Press, Cambridge.
- Gallese, Vittorio. 2008. “Mirror Neurons and the Social Nature of Language: The Neural Exploitation Hypothesis.” *Social Neuroscience* 3 (3–4): 317–33.

- Gallese, Vittorio, and Fausto Caruana. 2016. "Embodied Simulation: Beyond the Expression/Experience Dualism of Emotions." *Trends in Cognitive Sciences* 20 (6): 397–98.
- Gallese, Vittorio, and George Lakoff. 2005. "The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge." *Cognitive Neuropsychology* 22 (3): 455–79.
- Garrison, Jim. 2003. "Dewey's Theory of Emotions: The Unity of Thought and Emotion in Naturalistic Functional 'Co-Ordination' of Behavior." *Transactions of the Charles S. Peirce Society* 39 (3): 405–43.
- Gergely, György, and Pierre Jacob. 2012. "Reasoning about Instrumental and Communicative Agency in Human Infancy." *Advances in Child Development and Behavior* 43: 59–94.
- Gibson, William S, Shinho Cho, Osama A Abulseoud, Krzysztof R Gorny, Joel P Felmlee, Kirk M Welker, Bryan T Klassen, Hoon-Ki Min, and Kendall H Lee. 2016. "The Impact of Mirth-Inducing Ventral Striatal Deep Brain Stimulation on Functional and Effective Connectivity." *Cerebral Cortex*.
- Griffiths, Paul. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. University of Chicago Press.
- . 2004. "Is Emotion a Natural Kind?" In *Thinking About Feeling: Contemporary Philosophers on Emotions*, edited by Robert Solomon. Oxford University Press.
- Griffiths, Paul and Andrea Scarantino. 2009. "Emotions in the Wild: The Situated Perspective on Emotion." In *The Cambridge Handbook of Situated Cognition.*, edited by P Robbins and M Aydede. Cambridge University Press.
- Hufendiek, Rebekka. 2016. *Embodied Emotions. A Naturalist Approach to a Normative Phenomenon*. Taylor and Francis.
- Hutto, Daniel D. 2006. "Unprincipled Engagements: Emotional Experience, Expression and Response." In *Radical Enactivism. Intentionality, Phenomenology and Narrative. Focus on the Philosophy of Daniel D. Hutto*, edited by Richard Menary, 13–38. John Benjamins Publishing Company.
- . 2012. "Truly Enactive Emotion." *Emotion Review* 4 (2). SAGE Publications: 176–81.
- Isnard, Jean, Marc Guénot, Marc Sindou, and François Mauguière. 2004. "Clinical Manifestations of Insular Lobe Seizures: A Stereo-Electroencephalographic Study." *Epilepsia* 45 (9): 1079–90.
- Izard, Carroll E. 2007. "Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm." *Perspectives on Psychological Science : A Journal of the Association for Psychological Science* 2 (3): 260–80.
- Jacob, Pierre, and Marc Jeannerod. 2003. *Ways of Seeing – The Scope and Limits of Visual Cognition*. Oxford University Press.
- James, William. 1884. "What Is an Emotion?" *Mind* 9: 188–205.
- . 1890. *The Principles of Psychology*. New York: Holt.
- Jezzini, Ahmad, Fausto Caruana, Ivilin Stoianov, Vittorio Gallese, and Giacomo Rizzolatti. 2012. "Functional Organization of the Insula and Inner Perisylvian Regions." *PNAS* 109 (25): 10077–82.
- Kober, Hedy, Lisa Feldman Barrett, Josh Joseph, Eliza Bliss-Moreau, Kristen Lindquist, and Tor D Wager. 2008. "Functional Grouping and Cortical-Subcortical Interactions in Emotion: A Meta-Analysis of Neuroimaging Studies." *NeuroImage* 42 (2): 998–1031.
- Kragel, Philip A., and Kevin S. LaBar. 2016. "Decoding the Nature of Emotion in the Brain." *Trends in Cognitive Sciences* 20 (6). Elsevier: 444–55.
- Krolak-Salmon, Pierre, Marie-Anna Hénaff, Jean Isnard, Catherine Tallon-Baudry, Marc Guénot, Alain Vighetto, Olivier Bertrand, and François Mauguière. 2003. "An Attention Modulated Response to Disgust in Human Ventral Anterior Insula." *Annals of Neurology* 53 (4): 446–53.
- LeDoux, Joseph E. 1999. *The Emotional Brain : The Mysterious Underpinnings of Emotional Life*. Phoenix.
- Lindquist, Kristen A, Ajay B Satpute, Tor D Wager, Jochen Weber, and Lisa Feldman Barrett. 2016. "The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature." *Cerebral Cortex* 26 (5): 1910–22.
- Lindquist, Kristen a, Tor D Wager, Hedy Kober, Eliza Bliss-Moreau, and Lisa Feldman Barrett. 2012. "The Brain Basis of Emotion: A Meta-Analytic Review." *Behavioral and Brain Sciences* 35 (3): 121–43.
- Lindquist, Kristen A, Tor D Wager, Hedy Kober, Eliza Bliss-Moreau, and Lisa Feldman Barrett. 2012. "The Brain Basis of Emotion: A Meta-Analytic Review." *The Behavioral and Brain Sciences* 35 (3): 121–43.
- Logothetis, Nikos K. 2008. "What We Can Do and What We Cannot Do with fMRI." *Nature* 453 (7197): 869–78.
- Lowe, Robert, and Tom Ziemke. 2011. "The Feeling of Action Tendencies: On the Emotional Regulation of Goal-Directed Behavior." *Frontiers in Psychology* 2. Frontiers: 346.
- Mayberg, Helen S, Andres M Lozano, Valerie Voon, Heather E McNeely, David Seminowicz, Clement Hamani, Jason M Schwalb, and Sidney H Kennedy. 2005. "Deep Brain Stimulation for Treatment-Resistant Depression." *Neuron* 45 (5): 651–60.
- Mead, George H. 1895. "A Theory of Emotions from the Physiological Standpoint." *Psychol. Rev.* 2: 162–64.
- . 1934. *Mind, Self, and Society*. Edited by Charles W. Morris. University of Chicago Press.

- . 2001. “The Social Character of Instinct.” In *Essays in Social Psychology*, edited by M.J. Degaan. Transaction Publishers, New Brunswick NJ.
- Ortony, A, and T J Turner. 1990. “What’s Basic about Basic Emotions?” *Psychological Review* 97 (3): 315–31.
- Panksepp, Jaak. 2005. “Affective Consciousness: Core Emotional Feelings in Animals and Humans.” *Consciousness and Cognition* 14 (1): 30–80.
- . 2007. “Neurologizing the Psychology of Affects: How Appraisal-Based Constructivism and Basic Emotion Theory Can Coexist.” *Perspectives on Psychological Science : A Journal of the Association for Psychological Science* 2 (3): 281–96.
- Panksepp, Jaak, and Lucy. Biven. 2012. *The Archaeology of Mind : Neuroevolutionary Origins of Human Emotions*. W.W Norton.
- Panksepp, Jaak, and Jeff Burgdorf. 2003. “‘Laughing’ rats and the Evolutionary Antecedents of Human Joy?” *Physiology & Behavior* 79 (3): 533–47.
- Papagno, Costanza, Alberto Pisoni, Giulia Mattavelli, Alessandra Casarotti, Alessandro Comi, Francesca Fumagalli, Mirta Vernice, Enrica Fava, Marco Riva, and Lorenzo Bello. 2016. “Specific Disgust Processing in the Left Insula: New Evidence from Direct Electrical Stimulation.” *Neuropsychologia* 84 (April): 29–35.
- Pessoa, Luiz. 2008. “On the Relationship between Emotion and Cognition.” *Nature Reviews. Neuroscience* 9 (2): 148–58.
- Phan, K Luan, Tor Wager, Stephan F Taylor, and Israel Liberzon. 2002. “Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI.” *NeuroImage* 16 (2): 331–48.
- Proust, Joëlle. 2016. “The Evolution of Primate Communication and Metacommunication.” *Mind & Language* 31 (2). Wiley-Blackwell: 177–203.
- Provine, Robert R. 2000. *Laughter: A Scientific Investigation*. Viking.
- Quine, W. V. O. 1960. *Word and Object*. MIT Press.
- . 1984. “Relativism and Absolutism.” *Monist* 67: 293–96.
- Russell, J.A. 2015. “My Psychological Constructionist Perspective, with a Focus on Conscious Affective Experience.” In *The Psychological Construction of Emotion*, edited by Lisa Feldman Barrett and James Russell, 183–208. The Guilford Press.
- Saarimäki, Heini, Athanasios Gotsopoulos, Iiro P Jääskeläinen, Jouko Lampinen, Patrik Vuilleumier, Riitta Hari, Mikko Sams, and Lauri Nummenmaa. 2016. “Discrete Neural Signatures of Basic Emotions.” *Cerebral Cortex* 26 (6): 2563–73.
- Scarantino, Andrea. 2015. “Basic Emotions, Psychological Construction, and the Problem of Variability.” In *The Psychological Construction of Emotion*, edited by Lisa Feldman Barrett and J.A. Russell, 334–76. The Guilford Press.
- Scarantino, Andrea, and Griffiths, Paul. 2011. “Don’t Give Up on Basic Emotions.” *Emotion Review* 3 (4). SAGE Publications: 444–54.
- Scarantino, Andrea. 2012. “Functional Specialization Does Not Require a One-to-One Mapping between Brain Regions and Emotions.” *Behavioral and Brain Sciences* 35 (3): 161–62.
- Tanaka, K. 1993. “Neuronal Mechanisms of Object Recognition.” *Science (New York, N.Y.)* 262 (5134): 685–88.
- Ward, Lloyd Gordon, and Robert Throop. 1989. “The Dewey-Mead Analysis of Emotions.” *The Social Science Journal* 26 (4): 465–79.