

LSU PHIL 4941, Philosophy of Mind
Spring 2018
www.protevi.com/john

Notes on Lisa Feldman Barrett, *How Emotions Are Made*.

OVERVIEW OF BOOK

LFB is opposed to the linear classical view in which 1) sense stimuli trigger or turn on 2) latent or turned-off specific brain regions (a "footprint") which then 3) trigger motor output (facial expressions, HR changes). Hence, we have a battle in our souls / minds / brains between brute emotions and reason (xi; see also 280)

Barrett rejects this: there are no footprints (xii). Different brain areas (across individuals or even w/in individuals at different times) can support the "same" emotion (degeneracy or many-to-one), while the "same" brain region can support many different emotions (one-to-many).

Barrett holds to constructivism (xiii). Emotional experience emerges from active brain which combines input from ongoing (and constantly modulating) body processes, socially constructed emotion concepts which predict sensory input as well preparing for body action (e.g., crying), and environmental cues. This is all put together so that racing heart, chest tightness, and crying is made sense of as "sadness appropriate to the awful news" (xiii). But the "same" body actions could be made sense of as anger or fear or joy in other circumstances (xiii).

Key point is that our cultural script of classical view of emotions enables us to make sense of our experience of emotions, which due seem to arrive suddenly and involuntarily in consciousness as primitive / irrational / triggered passions that take over our minds and bodies (xiii). So much of the book is structured by this split between the active brain and the passive mind or subject of experience.

Barrett will sometimes say "you are actively making sense of a situation" and other times that "you don't know what your brain is doing." So, on p. 31, "you're not a passive receiver of sensory information, but an active constructor of your emotions." Here the "you" means "your brain." So here the brain is a power of spontaneity, a power or "faculty" that performs actions ("learns," "understands," "makes sense," etc.). However, on 26, "the entire process of construction is invisible to you." Here the "you" is the subject of experiences, which seem to simply arrive in consciousness. Hence, to use cognitive science terms, the brain mechanisms are "subpersonal," and the emotion experiences are "personal."

From an enactivist perspective, I would add that it is the organism that is underlying this active brain / passive mind split. The brain is not the "base" of the mind; rather, as Thompson puts it in his *Psychology Today* blog post, "the brain facilitates the mind, which is an event mediating brain, body, and world such that an organism acts in its environment."

Now we have this split due to NS as we need a really fast predictive brain to keep us alive in complex bio-social niches which have allowed consciousness over evolutionary time. And certain niches perhaps allow our persons to flourish ("person" being that which we are in social interactive reality). In terms of my Wexler paper, our selves are "dynamic patterns of social and somatic interactions." It's our political challenge to construct niches (institutions, ways of thinking, speaking, and being, and so on) that allow the widest possible spread of occasions for flourishing.

BARRETT AND PREDICTIVE PROCESSING

Barrett's basic outlook is that the brain is a prediction engine that prepares for sensation and for body action, with perception being error reduction (chapter 4). This is a strategy for information reduction – the brain need only deal with sensation that doesn't match prediction, rather than tackling the entire massive stream of incoming information (85).

A "concept" (= mental representation of a "category" or set of objects [87]) is unpacked in a context in which fast arriving sensation modulates ongoing brain activity (58), which is already predicting what is going to happen and what should be done about it.

For instance, you're in the fruit aisle of the supermarket so you're predicting apples. When "shiny red" sensations arrive they then fit right into the prediction. So, sensation is not really providing new information needing to be processed. But suppose some joker has slipped in a plastic apple. Picking it up, you get touch sensations that don't match the touch predictions, so your brain is going to shift prediction to "fake apple" which has the effect of allowing sensation to match the new category, and thus reducing processing as there is now a match of input to prediction.

Prediction also goes in parallel to motor preparation. So, when you're predicting "apple" your saliva is going, as you love apples and can't wait to bite into it (or vice-versa, your stomach is queasy as you hate apples). The plastic apple then can be a disappointment or a relief to your anticipated tasting.

Barrett uses the same model for emotions. We have learned emotion categories from our cultures. They are unpacked and used for prediction in appropriate context, and then input from external and internal sources ("interoception" – p. 56 and 66) will correct errors in that prediction – for instance, your friend is not really mad at you; she's just stubbed her toe and is in pain.

NOTES ON INDIVIDUAL CHAPTERS

CHAPTER 1: search for "footprints." Here Barrett rejects the "footprint" idea and introduces 2 key concepts: degeneracy, or "many-to-one" (many neural patterns to produce an emotion)

and "core systems," which are "one-to-many" (the same brain system can produce many different emotions) (19).

CHAPTER 2: emotions are constructed. This chapter introduces prediction: brain systems simulate models in advance of arriving sensory information which is then used as error correction (26-27). Concepts are formed from past experience and are "primary tool for brain to guess meaning of incoming sensory inputs" (28). As this goes on subpersonally, senses seem to deliver the world as is (29).

An important twist: the brain uses simulations to make sense of incoming body signals (29-30). That is to say, an emotion concept is unpacked and an instance of that concept is used to make sense of body input in a worldly context (30): my heart is beating fast and a dog is snarling, I must be experiencing "fear."

Variation or "population thinking": concepts are records of different experiences with loose similarities of body sensation and worldly context (35). This gets us into the philosophical question of "nominalism," which denies the existence of universals. Universals are mere names for similarity profiles of what really exists, concrete instances:

Degeneracy: different instances of fear are constructed by different combinations of the core systems throughout the brain. We can describe the instances of fear together by a pattern of brain activity, but this pattern is a statistical summary and need not describe any actual instance of fear. (36)

CHAPTER 3: failed search for universal emotions.

CHAPTER 4: origin of feeling: affect as general sense of well-being; valence as positive and negative affect (72). It depends on interoception, but why we have feeling is not yet known ("Hard Problem") (73). Affective realism: "you see what your brain believes" (77-78): that is, prediction prepares a way to interpret sensation to produce perception. Rationalist human exceptionalism and limbic system theories (81).

CHAPTER 5: concepts, goals, and words. classical concepts are definitions or sets of necessary and sufficient conditions for membership in a category (88). Prototypes are central profiles of family resemblances; they pick out the typical or most common instance of a kind, so you can have variation from the center but not too much (89). But with neural profiles of emotions, you hardly ever get the prototype: it's all variation and no center (89). So brains construct prototypes on the spot, as needed (89): it imposes similarities across instances, depending on your goal at the time (90). So your brain is constructing concepts as needed (92).

Statistical learning or pattern formation enables predictions (94-95), including about info contained in other minds (96) about goals, intentions, preferences (98). Emotion words are the key then in developing ability to construct emotion concepts as we grow up (102). So we have cultural variation in emotion carried by different emotional vocabularies (103-4).

Passage on page 110 gives "theory of constructed emotions in a nutshell."

CHAPTER 6: how brain makes emotions. Concepts and predictions are the same. You develop concepts as compressions at high neural level, that is, "multisensory summaries" of past experience (that is, match of exteroception and interoception – how external events have changed your well-being and gotten an "affective profile tag" [my term]), and you "unpack" them in a "cascade" that moves from general to specific (it decompresses the concept's compressed state) that predicts and interprets incoming sensory information as matching or departing from the prediction (118-120).

"Control system" (really just an "optimizing function") allows brain to choose a "winner" among competing concept cascades (123); these winners allow body-budgeting and action preparation; the major hubs in control and interoceptive systems handle most brain info and are possibly enablers of consciousness.

CHAPTER 7. Emotions as social reality. Collective intentionality is necessary but not sufficient for social reality, as some non-human animals exhibit behavior coordination and "rudimentary forms of collective intentionality" (135). Humans have words, though, and combined with collective intentionality we can form shareable concepts; we can affect other people's actions; and we can perform mental inference or Theory of Mind (136-37). Emotion concepts are most easily learned via emotion words (138).

Functions of emotion concepts. 1) make meaning or interpret extero and interoceptive signals; 2) prescribe action; 3) regulate body budget; 4) communicate emotions, via synchronized concepts in different people; 5) social influence as you can influence other people's body budgets by provoking emotion concepts in them (138-39).

From 140-151, discussion of cultural variation in emotion.

CHAPTER 8: new theory of human nature. Barrett opposes her bio-cultural constructivist view to the "classical" or essentialist view. Constructivism is brain shaping our perceptions / our reality, but culture shapes our brains as well (153). See also: "we are not blank slates ... but neither is biology destiny. When we peer into the ... brain, we don't see mental modules. We see core systems that interact continuously in complex ways to produce many sorts of minds, depending on culture" (170).

The rest of the chapter is Barrett's reading of the history of essentialist / classical view of emotion. She criticizes Darwin for backtracking from the variation or population thinking of *Origin* to the essentialist theory of *Expression*. She blames Dewey for "James-Lange" or footprint theory of emotion as reading of body states – what James actually said is that each instance of emotion, not each category, is a singular body state (160). She provides an analysis of how essentialism comes about – ironically, from brain capacity to categorize and conceptualize (163). She criticizes brain localists, like Broca, and the "triune brain" idea of

reptilian brain, limbic system, and neocortex (these map onto Plato's triune soul of appetites, emotions, and reason) (169).

CHAPTER 9: mastering your emotions. Three good pieces of advice: 1) regulate your body budget with sleep, food, exercise; 2) increase the granularity of your emotional concepts (learn new emotion words, read literature); 3) in the moment, you can change your setting, and / or learn to pay attention to physical sensations underlying your affect, and / or recategorize those feelings as good signs of effort (moving from "suffering" to "discomfort"). You can also "deconstruct" the self by dissolving it into singular instances that are not expressions of an essence, but that nonetheless form patterns (192). You can also improve your emotion perception of others (194ff).

Skip to chapter 12.

CHAPTER 12. non-human animals Barrett has stringent criteria for emotional experience: interoception and affect; emotion concepts; social reality for sharing emotional concepts with next generation (254). She says non-human animals have affect, but not emotions. What looks like emotions they are undergoing are really just behavior circuits (freezing, snarling, etc), but not proper emotion. She cites LeDoux's backtracking from saying rats experience "fear" to saying they exhibit survival behavior [and presumably negatively valenced affect] (274-75).

CHAPTER 13: RECAP. The brain is a complex system due to experienced-based microwiring and continual (though somewhat declining) plasticity, emergence via neurotransmitters, degeneracy, and multipurpose networks (281-82). Brain complexity was selected for in NS, though we have some innate preset concepts: positive and negative valence, low and high arousal, but overall, we need population thinking both within and across humans.

Three commonalities of mind: affective realism, concepts, social reality.

Affective realism from importance of body-budgeting regions of interoceptive networks being powerful predictors, with sensory regions are "eager listeners" (283). This is the "gut feeling" of certainty; should be leavened with curiosity (284).

Concepts: predictions being unpacked by context clues; corrected by perception and preparing for action: what's happening and what should be done about it? (285) These are coupled in social interaction with the prediction, correction, and preparation systems of social partners. Danger here is essentializing our self-understanding (classical emotional view) and ignoring reality that doesn't match our desires (which guide predictions).

Social reality: aka, ideology or bio-cultural affective-cognitive bodies politic. What works for the group will be predicted but it gets locked in via affective realism as natural or only possible way to do things (286-87). From there LFB gets into epigenetic bio-cultural constructivism: social interactions cause micro-wiring via plasticity (and chromatin etc for gene expression) (289).

BARRETT AND MORAL PSYCHOLOGY

A strong strand of moral psychology is skeptical of the explanatory power of moral reasoning in many cases, seeing it as very often (or even always, for the hardcore), simply an *ex post facto* rationalization of an intuited preference: in modern times, we find Hobbes, Spinoza, and Nietzsche in this line of thought. Contemporary scientific formulations of this school include Damasio (1995, 1999) and Haidt (2001).

According to this corporeal school of thought the intuited preference arises from a state of the subject's body – a certain relation of neural firing patterns, neurochemical releases, bodily information, gut feeling, and so on – expressed as an affective state of liking and attraction, or disliking and repulsion. From this bodily state and its affective expression stem a judgment. In Hobbes canonical formulation, "I judge something good because I like it." Positive or negative emotion here is thus an efflorescence of the body, and moral reasoning seeks to rationalize judgments stemming from that intuited feeling.

This corporeal position differs from Stoic philosophy, in which emotions are cognitive judgments: "If I am appropriately virtuous, my liking a thing is consistent with my judgment that it is good relative to the nature of rational beings." The neo-Stoic Martha Nussbaum puts it like this: emotions are "intelligent responses to the perception of value."

Between corporealism and cognitivism we find constructivist theories of emotion, which will insist upon the contribution of semantic factors alongside somatic changes, as in Lisa Feldman Barrett or Joseph LeDoux. LeDoux's allows some reference to specific neural circuits, such as the subcortical defense reactive circuits that are added to other inputs in his "recipe" for fear and anger (LeDoux 2015, 93-112). Barrett (2017), however, insists on a strong neural globalism, which, with her insistence on holism, emergence, and degeneracy (same outcome from different mechanisms), results in a strong nominalism, such that no "fingerprint" of necessary circuits can be identified for either emotion instances or even emotion categories (2017, 35-41; see also Pessoa 2017 for a similar distributed network approach to emotions).

Barrett's work shows the relevance of Deleuzian ontology. For Barrett, emotion concept construction occurs via bottom-up summarizing of singular experiences, drawing on neural inputs from multiple brain sites mapping the body and other higher and lower intra-brain regions; each of these "core affect" experiences is tagged with culturally specific emotion terms. Hence there is a high-level, cortical / semantic component to emotion concepts, which are constructed from these multiple inputs. Such summarizing produces concepts as abstract but non-essential capacities that don't exist as enduring, locatable, actual firings, but only insist as potentials for actualization. Given her strong holism, emergence, and degeneracy, concept creation is the progressive construction of a virtual field: virtual, because emotion concepts do not exist, but insist as potentials.

An emotional episode is the actualization of the potential concept. It occurs as prediction, a top-down simulation that "unpacks" concepts, constructing an instance of the concept that assembles

its components from occurrent inputs and checks the assemblage against the prediction. This actualization occurs in a degeneracy mode, such that no single set of neural firings is necessary for each instance of the concept. Hence the concept is a virtual diagram with multiple mechanisms for the actualization of instances. In Deleuzian terms, it is an "abstract machine" with multiple machinic assemblages for its actualization / individuation / integration / differentiation.

This is not all that different from the corporealists if we allow for cortical / semantic brain firings to count as "states of the body." Where does that body state come from? From the history of encounters of the body, that is, from the interaction of the particular character of the body – its "conatus," to speak like Hobbes and Spinoza or its "system of drives" to use Nietzsche's terms – and the people and things it encounters. The history of encounters patterns the body by modulating its characteristic responses, deepening the affective response upon repetition of similar objects, that is, in Barrett's terms, the formation of more and more deeply and finely grained emotional concepts.

Here the brain / body is neither a blank slate nor a preformed set of responses. Encounters are neither imprintings nor are they mere triggers; rather, the body has characteristic dispositions that set up a range of stylized responses. Similarly, this is neither a nominalist position of sheer idiosyncratic – and hence unpredictable – responses, nor is it a universalist human nature in which we can predict responses. Rather, if we get to know someone, and see an encounter with a relatively familiar object, we have a good sense of what will happen. But bodies are complex and internal changes arise, as do situations unfamiliar to the agent, such that prediction of the response are less reliable, both to the onlookers and even to the experiencing agent: we can surprise ourselves with changes in our emotional patterns.

Whatever your position in theory of emotions, as a corporealist (Damasio 1995, 1999; Haidt 2001) or a constructivist (LeDoux 2015; Barrett 2017), singular encounters are the genesis of emotional structures. And whatever your position in moral psychology, as an emotion-first intuitionist or someone according some weight in some circumstance to moral reflection (interestingly, Jonathan Haidt [2001], although most closely identified with the former, also allows for the latter), emotions are important, either as generators of opinion, as obstacles to proper judgment, or as indicators of virtue or vice.

From the importance of encounters flows the importance of institutions to political affect. If one can set up ways in which some patterns of encounters can be made more likely than others, one can influence brain / bodily histories, and thus emotional responses, and thus the impulses toward moral action (whatever you think of the efficacy of later rational reflection). When people live in a society with firmly set institutions, their brain / bodies, intuitions, emotions, and impulses toward actions are more or less predictable, and, if the institutions are well-attuned to each other, social reproduction occurs more smoothly than if the institutions clash. (Although I've insisted that no society is so simple and integrated that clashes of patterns never occur, I don't want to gainsay differences in frequency and intensity of such clashes across cultures.)

In a multicultural society, however, there are many institutions, hence many different and not necessarily concordant patterns of encounters, many patterned bodies, many intuitions as to right action, many emotional reactions, and many different ex post facto rationalizations. Hence Plato's desire for cultural control, evidenced by the expulsion of the poets from the city of the *Republic*, or the complete cultural control of the city in the *Laws*.