Are Emotions Psychological Constructions?

Charlie Kurth
Department of Philosophy
Western Michigan University

Abstract: According to psychological constructivism, emotions result from projecting folk emotion concepts onto felt affective episodes (e.g., Barrett 2017, LeDoux 2015, Russell 2004). Moreover, while constructivists acknowledge there's a biological dimension to emotion, they deny that emotions are (or involve) affect programs. So they also deny that emotions are natural kinds. However, the essential role constructivism gives to felt experience and folk concepts leads to an account that's extensionally inadequate and functionally inaccurate. Moreover, biologically-oriented proposals that reject these commitments are not similarly encumbered. Recognizing this has two implications: biological mechanisms are more central to emotion than constructivism allows, and the conclusion that emotions aren't natural kinds is premature.

This paper challenges the psychological constructivist account of emotions that is gaining prominence among neuroscientists and psychologists (e.g., Barrett 2017, 2012, 2009; LeDoux 2015; Russell 2004). According to constructivism, emotions result from projecting culturally-fashioned concepts onto felt affective episodes. Fear, for instance, just is a feeling of negative arousal as viewed through the lens of one’s folk concept FEAR. This proposal is novel in taking felt experience and cognitive projection to be essential elements of what emotions are. Moreover, while constructivists acknowledge that there’s a biological dimension to emotions (e.g., neural mechanisms are responsible for generating the conscious feelings that we project our emotion concepts on to), they deny that emotions are, or necessarily involve, anything like an affect program. Thus, constructivism is philosophically significant in two ways. First, in denying an essential role for biological mechanisms, it challenges influential, affect-program-oriented accounts of emotion (e.g., Scarantino & Griffiths 2011; Ekman & Cordaro 2011). Second, in understanding emotions as projections of folk emotion concepts, it takes emotions to be social-psychological constructions, not natural kinds.
But despite constructivism’s appeal among cognitive scientists, the role that it gives to felt experience and folk concepts leads to an account of emotion that’s both extensionally inadequate and functionally inaccurate. Moreover, biologically-oriented proposals that reject constructivism’s problematic commitments are not similarly encumbered. Recognizing all this reveals that an adequate account needs to give greater place to the biological mechanisms that underlie emotions than constructivism allows. This, in turn, suggests that the constructivists’ conclusion that emotions are not natural kinds is premature.

1. Psychological Constructivism and Its Appeal

Constructivism sees emotions as having two elements: a felt affective experience and a cognitive projection or labeling. Taking these in turn, the felt experience component—or “core affect” as it’s often called—is a neurophysiological state that manifests as a consciously experienced combination of valence (i.e., feeling good or bad) and arousal (i.e., feeling activated or deactivated) (Barrett 2006: 48; Russell 2004; LeDoux 2015: 226-232). Importantly, constructivism’s focus on core affect looks just to the amalgamated experience of these two components—valence and arousal. What causes this felt experience is irrelevant to the nature and individuation of emotions. In fact, and as we will see, allowing that particular sensations (instances of core affect) can be produced by a range of distinct neural circuits or somatic events is taken to be a point in favor of the constructivist proposal.

Given this account of the felt dimension, constructivism maintains that “discrete emotions emerge from a conceptual analysis of core affect. Specifically, the experience of feeling an emotion…occurs when conceptual knowledge about emotion is brought to bear to categorize a momentary state of core affect. … [These] [c]ategorization processes enact the rules, [that guide] the emergence of an emotional episode” (Barrett 2006: 49; also LeDoux 2015: 225-232). This talk of “conceptual analysis,” “conceptual knowledge,” and “categorization” should be understood thinly.
The underlying process needn’t involve some full-fledged, conscious judgment. Rather, all that’s necessary is an unconscious or implicit recognition that one’s sense of one’s situation, and one’s felt physiological state, fall under a particular folk emotion concept.

These emotions concepts, in turn, should be understood as folk theories or culturally-shaped behavioral scripts that detail the nature and function of the particular mental states picked out by specific emotion labels (‘fear,’ ‘joy,’ ‘anger,’ etc.). Moreover, the fact that folk emotion concepts engage these folk theories and behavioral scripts entails that the projecting of a particular label onto an instance of core affect not only imbues one’s situation with the associated, emotionally-colored meaning, but also shapes one’s subsequent thoughts, physiological responses, and behaviors (Barrett 2012; LeDoux 2015).

Formalizing this a bit, we can see psychological constructivism as committed to four theses:

(PC1) Each emotion type/category is constituted by the projecting of a specific folk emotion concept (e.g., FEAR, JOY) onto a felt affective experience.

(PC2) Token emotion episodes (e.g., a given instance of fear) are cognitive acts where one (implicitly) labels an occurrent conscious feeling with a particular folk emotion concept and so comes to see the feeling through the lens of that concept.

(PC3) There is no unique (set of) neural circuit(s) or psychological mechanism(s) responsible for the conscious feelings that get categorized with particular folk emotion concepts.

(PC4) The act of labeling a feeling with a particular folk emotion concept affects one’s subsequent thoughts, physiological responses, and behaviors.

According to its advocates, much of constructivism’s appeal lies in its explanatory power. In comparison to more biologically-oriented theories, it provides a better explanation of empirical research on the biological mechanisms and correlates associated with emotions (e.g., neural circuits, patterns of physiological change, and expressive behavior). Since the discussion that follows will build
from the contrast between constructivism and competing biologically-oriented theories (BTs), it will be useful to briefly sketch the BT approach and the constructivists' case against it.

As a generalization, BTs maintain that emotions are, or necessarily engage, affect programs—that is, largely encapsulated systems that automatically prompt stereotyped patterns of physiological changes, expressive behavior, motor routines, attentional shifts, and forms of higher-cognitive processing in response to (evolutionarily-relevant) threats and opportunities. So, for example, fear is (or essentially involves) an affect state that consists of automatically engaged tendencies for inter alia increases in arousal, narrowing of attention, and the cueing of fight/flight/freeze behavior in response to the perception of some danger.

But since BTs take affect programs to be essential (even identical) to emotions, constructivists argue they cannot explain two well-documented sets of findings.1

(F1) One can feel a given emotion without engaging what science suggests is the best candidate for its underlying biological drivers (or their correlates)—e.g., activation of particular neural circuits, a distinctive physiological response, characteristic expressive behavior.

(F2) The relevant biological drivers/correlates can be engaged though one does not report feeling the associated emotion.

So, for instance, though the central nucleus of the amygdala (CeA) is thought to be central to fear, research shows both that individuals will report being afraid when the CeA is not engaged (F1), and that the CeA can be active though individuals report not feeling fear (F2).

BT proponents have sought to address these explanatory limitations by insisting that we must narrow our understanding of what, say, FEAR is. More specifically, they maintain that the folk emotion concepts that the above research relies on (in, e.g., the self-reports of emotions (not) felt) are too

---

1 See, e.g., Barrett 2012 for a review of the relevant empirical work.
coarsely grained for scientific investigations like these. The BT advocates’ expectation is that a more refined account of what ‘fear’ refers to will reduce, even eliminate, dissociations of the sort noted above (e.g., Scarantino & Griffiths 2011; Kurth 2018). But constructivists respond that any effort to narrow or otherwise refine our emotion concepts along these lines will result in an account of (e.g.) fear that is troublingly stipulative or excessively revisionary with regard to our ordinary understanding of these emotions (Barrett 2012: 415-6; LeDoux 2015: 234).

Two aspects of this debates are particularly important for our purposes. First, central to the constructivist complaint is the move to take a failure to accommodate our ordinary emotion talk as the standard for what counts as stipulative or excessively revisionary account. Second, given our ordinary emotion talk as the standard, the above four theses appear to give constructivism the resources and flexibility it needs to explain not just (F1)-(F2), but also the richness and cultural variation of emotional life more generally (e.g., Barrett 2012, 2009). However, I will argue that investigating the extensional adequacy and functional accuracy of constructivism’s core theses provides us with reason to doubt each of (PC1)-(PC4).

2. **Is Constructivism Extensionally Adequate?**

As we’ve seen, a central feature of the debate between constructivism and BTs is the charge that BTs cannot accommodate dissociation data without committing to a stipulative or excessively revisionary account of what emotions are. In what follows, I give three examples that suggest constructivism faces a similar problem. More specifically, a closer look at the constructivists’ dual claim that emotions are cognitive labelings of felt experiences reveals that the account is both under- and over-inclusive with regard
to our ordinary understanding of things like: what emotions are, when we experience them, and how they differ from moods, feelings, and other categories of affect.²

First consider the constructivist’s commitment to understanding emotions as felt experiences—that is, changes in core affect that we’re consciously aware of. An implication of taking felt affective experience as essential to being an emotion is that it rules out the possibility of unconscious emotions. Some constructivists appear to embrace this result. For instance, Joseph LeDoux maintains that claims about unconscious emotions are “oxymoronic” (2015: 234; also, 19). But LeDoux’s acceptance of this implication aside, the thought that there cannot be unconscious emotions fits poorly with our everyday experiences and our ordinary emotion talk.

For instance, if there aren’t unconscious emotions, then how do we explain situations where we don’t realize that we were (say) afraid until after the danger has passed? Pressing further, notice that we not only regularly speak of unconscious emotions, but also appeal to them in order to explain our behavior. For example, we say things like, “Bill won’t discuss the book he is working on. He says it’s not ready yet—but he doesn’t realize that he’s really just afraid about getting negative feedback.” While ordinary talk like this is easy to make sense of on the assumption that Bill is unconsciously fearful, such an explanation isn’t available to a constructivist like LeDoux—our ordinary talk to the contrary, Bill isn’t unconsciously afraid, but rather experiencing some other psychological blockage.

But the constructivists’ trouble with unconscious emotions runs deeper—the case for their existence also has empirical support. For instance, recent experimental work has shown that subliminally presented emotion faces can produce affective responses that bring emotion-specific behaviors even though the subject denies feeling an emotion. In particular, subliminally presented happy

² Thus the strategy I employ here—one that grants constructivists’ their criterion for assessing when an account is excessively revisionary—is distinct from standard defenses of BTs noted in §1.
faces bring increased “liking” behavior (e.g., greater consumption of a novel beverage), while subliminally presented angry faces have the opposite result (Winkielman et al. 2003; also, Kihlstrom 1999). Since these patterns of behavior mesh with our understanding of both joy as an emotion that tends to increase interest/engagement, and anger as an emotion that brings avoidance/rejection tendencies, these results are taken as evidence of unconscious emotions.

While the constructivist might try to pass these findings off as cases where unconscious changes in core affect (not emotion) produce the behaviors, the plausibility of the proposal is undercut by the fit we find between the subliminally presented happy (angry) face, the resulting liking (avoidance) behavior, and our ordinary understanding what happiness (anger) involves (Winkielman et al. 2005). The upshot, then, is that constructivism’s insistence that felt changes in core affect are essential to what emotions are has revisionary implications with regard to our ordinary (and scientific) understanding of emotional life.

But even if we’re willing to grant that our talk of unconscious emotions is merely metaphorical—an elliptical way of talking about some non-emotion form of (unconscious) affect—the constructivist’s second core commitment brings additional problems. In particular, the claim that emotions are the product of our cognitive labelings/projections makes facts about when we are experiencing an emotion—and what emotion it is—too sensitive to random situational features and framing effects. To draw this out, consider the following case.

Coffee. I order a cup of decaf coffee and sit down to read a magazine cover story about Trump’s latest foreign policy provocations. But unbeknownst to me, the barista confuses my order and I get a cup of regular coffee. As the caffeine works its way into my system, it brings a (consciously experienced) change in my arousal. As a result, I start reading the article with jittery attentiveness.

Given the scenario, it seems my jittery, attentive reading is best understood as a bout of caffeine-induced hyperactivity. But notice: there’s nothing in the constructivist account to rule out the
possibility that I’m actually having an emotional experience—I’m afraid. After all, on the constructivist account, this experience could be a change in core affect that I’ve (implicitly) labeled ‘fear.’ While that possibility alone seems odd (to my ear, at least, the case is best understood as emotionless hyperactivity, not fear), there’s more trouble.

To draw this out, consider the constructivist’s likely response to the case. Given the setup, she would likely maintain that whether this is an instance of fear depends on whether I see it that way—what sort of meaning do I attribute to my situation (e.g., Barrett 2017: 126; 2012: 419-420; 2009: 1293)? For instance, if I assent to the barista’s remark that I seem really uneasy about the article that I’m reading, then—by (implicitly) labeling my behavior through my assent—I imbue my situation with the meaning carried by my FEAR concept. I am, therefore, feeling fear. While this move might seem to allow the constructivist a way to account for the case, it comes at a high cost. For notice, had the barista instead said something like, “Whoops, I messed up and gave you regular, not decaf—no wonder you’re so hyper,” I’d likely assent to that too. And so I wouldn’t be afraid—just hyperactively aroused.

But that’s odd. Our ordinary thinking about emotions suggests that whether I’m experiencing a particular emotion, and what emotion I’m experiencing, should not be so sensitive to random situational features like what questions the barista—or anyone for that matter—just happen to ask me. To be clear, the claim here is not that emotions are immune to situational and contextual factors. Rather, the point is that on the constructivists’ account emotions turn out to be too sensitive to them. The radical situational sensitivity entailed by constructivism makes it not only too easy to experience an emotion, but also ties facts about what emotion we’re experiencing to irrelevant situational factors.

Together, the difficulties raised by unconscious emotions and incidental situational features call the extensional adequacy of the constructivist account into question and do so in a way that
pinpoints the commitments of (PC1) and (PC2) as the source of the trouble—after all, these claims posit feelings of core affect and projections of folk concepts as essential to what emotions are. Of equal note is the fact that biological theories are less vulnerable to these difficulties. For one, irrelevant situational features should have less influence on what emotion one happens to experience since, according to BTs, emotions are (or are principally driven by) affect programs, not contextualized cognitive labelings. Moreover, since affect programs are things that can operate below that level of conscious awareness (Kurth 2018), taking emotions to be driven by affect programs provides BTs with the resources needed to explain unconscious emotions.

While the above discussion raises worries about the first two constructivist theses (PC1-PC2), it also provides the makings for worries about the third. In particular, because constructivism denies (via PC3) that emotions are underwritten by affect programs, it has trouble making plausible distinctions between emotions and similar states like moods. To draw this out, notice that the coffee case from above can be easily extended to show that constructivism makes it too easy to flip between moods and emotions. All we need to do is substitute “being in a worried mood” for “hyperactive” in the presentation of the case. Once we do this, we see that mere changes in the question the barista asks me can change whether I’m worried (a mood) or afraid (an emotion).

So we again see that constructivism has problematic explanatory limitations—this time with regard to preserving the thought that there’s a substantive difference between moods and emotions. On the constructivist account, this distinction is just a matter of how we happen to label our felt experiences. While some constructivists appear willing to accept this conclusion (e.g., Barrett 2017, 2009), it highlights another place where the constructivist proposal has revisionary implications—after all, moods and emotions are generally thought to be distinct forms of affect (e.g., Ben-Ze'ev 2000: Chap. 4). Moreover, here too we have a difficulty that’s easily avoided by biological accounts. Since
BTs take emotions to be (driven by) affect programs, they can appeal to the engagement of these mechanisms as the basis for the emotion/mood distinction (e.g., Kurth 2018; Wong 2017).

Stepping back, then, although constructivism purports to be less stipulative with regard to capturing our ordinary understanding of emotions, the above examples call this into question. For starters, the constructivists’ commitment to (PC1)-(PC3) has revisionary implications for our ordinary understanding of what emotions are, when we experience them, and how they differ from moods. Moreover, we have also seen that biologically-oriented accounts—in eschewing this trio of problematic theses—are better equipped to provide a plausible account of these features of our everyday emotion talk.

3. Is Constructivism Functionally Accurate?

The challenges to the constructivist picture extend beyond concerns about its extensional adequacy. The account also makes predictions about how projecting emotion concepts onto felt experience should shape subsequent behavior that are poorly supported by the empirical record. Two examples will draw this out.

First consider emotion misattribution research. In this work, a feeling that is typically associated with a particular emotion (e.g., feelings of unease and anxiety) is subtly induced, but the individual is lead to believe they are not, in fact, experiencing that emotion but rather something else (e.g., the effects of caffeine). Constructivism predicts (via PC4) that individuals in these experiments should display different behaviors depending on whether they are in the control or misattribution conditions. For instance, individuals led to believe that the unease they’re feeling is not anxiety, but something else (caffeine) should display diminished anxiety-related behaviors in comparison to controls who were not misled about their unease. But on this score, the experimental findings are decidedly mixed.
First, while there is a sizable body of findings showing misattribution manipulations attenuate subsequent emotion-related behavior, there is also a sufficiently large set of non-confirmations to raise concerns. For instance, while some research on public speaking anxiety suggests that attributing unease to a pill you just took rather than anxiety about a public talk you must give leads to a reduction in anxiety-related behaviors—stuttering, apprehension, and the like (Olson 1988), other studies have failed to find any differences in these behaviors (Slivkin & Buss 1984; Singerman, Borkovec & Baron 1976).

Moreover, even in cases where emotion-related behavior is reduced in the manipulation condition, it’s not clear how much support this brings to the constructivist. This is because it’s often unclear whether the reductions in emotion-specific behavior are (i) the result of the misattribution or (ii) a consequence of directing subjects’ attention away from the emotion eliciting stimuli (for a review, see, e.g., Reisenzein 1983). This potential confound is problematic for constructivists since only possibility (i) provides direct support for the claim of (PC4)—namely, that the act of labeling itself affects subsequent behavior.

The second problematic set of results comes from work in political science. This research investigates how negative emotions shape public policy decision making among voters (e.g., MacKuen et al. 2010; Brader et al. 2008; Valentino et al. 2008). The core hypothesis of this research is that negative emotions (especially, anger and anxiety) affect subsequent behavior in different ways. In particular, anger—as a response to challenges to what one values—should tend to bring behavior geared toward defending the threatened values. By contrast, since anxiety is a response to uncertainty, it should tend to bring caution and information gathering aimed helping one work through the uncertainty one faces.
To test these predictions, the experimental set up works as follows. First, individuals are asked to read a (fake) news story designed to provoke anger or anxiety by challenging the individuals’ pre-existing views about contentious policy issues like immigration, affirmative action, and economic policy. After reading the story, the participants are given the opportunity to use a website containing links to additional information, both for and against, the policy issue at hand. They are also asked how the original news story they read made them feel (e.g., angry, anxious). So by tracking what kinds of information the participants looked at through the website, experimenters can identify differences in how the anger and anxiety provoked by the story shaped subsequent behavior.

In the present context, these experiments allow us to test a pair of predictions that follow from the constructivist theses (PC1) and (PC4):

(P1) Labeling felt experiences with distinct folk emotion concepts should bring different patterns of behavior.

(P2) The behaviors that result from labeling a felt experience with a particular concept should map to our folk understanding of the emotion in question.3

More specifically, given (P1) and (P2), we should see different behaviors based on whether the participants in the experiment label their emotion ‘anger’ or ‘anxiety’ (P1). Moreover, the different behaviors should map to the above, ordinary understanding of these emotions—e.g., angry individuals should look for information that helps them defend their preferred policy position, while anxious individuals should engage less in motivated inquiry and more in open-minded forms of investigation (P2).

3 As evidence of constructivism’s commitment to these predictions, consider Lisa Feldman Barrett’s comment that “when a person is feeling angry…she has categorized sensations from the body and the world using conceptual knowledge of the category ‘anger’. As a result, that person will experience an unpleasant, high arousal state as evidence that someone is offensive. In fear…she will experience the same state as evidence that the world is threatening. And, either way, the person will behave accordingly” (2009: 1293, emphasis added).
However, whether we find support for these predictions turns—surprisingly—on what the policy issue used in the experiment was. More specifically, in experiments where the policy question that was challenged by the fake news story concerned immigration, the results fit poorly with constructivism’s predictions. That is, participants behaved in the same angry way regardless of whether they reported feeling anger or anxiety (Brader et al. 2008). By contrast, if the policy issue at hand concerned affirmative action or economic policy, the results are more in line with (P1)-(P2): anger and anxiety provoked by the news stories not only brought different patterns of behavior, but the resulting behaviors mesh with our ordinary conception of how these emotions function (MacKuen et al. 2010; Valentino 2008).

While this second set of results might appear to be good news for constructivists, the trouble lies in explaining why we get the different results between the immigration and affirmative action/economic policy experiments. After all, other than the content of the issue at hand, the experimental designs were identical. In response, the constructivist might argue that content and context matter (e.g., Barrett 2012, 2009): the similar behaviors that subjects display in the immigration version of the study suggest that the cultural scripts associated with ‘anger’ and ‘anxiety’ are highly sensitive to negative stereotypes about minorities. More specifically, the thought would be that there’s something about the combination of immigration debates and racial stereotypes that changes the standard behavioral scripts associated with ‘anger’ and ‘anxiety’ so that, while they typically generate different behaviors, they now bring the same ones.

But setting aside concerns about the ad hoc nature of this proposal, without more of a backstory, it’s unconvincing. After all, affirmative action debates are also framed in racial stereotype provoking ways. So here too we should see anger and anxiety generating similar patterns of behavior. But we don’t.
Moreover, notice that, on this front, biological accounts have an easier time explaining the experimental findings. For instance, as one possibility, the BT advocate could argue that only participants in the immigration study are likely to be experiencing both anger and anxiety: anger about the harms immigrants will bring and anxiety given their uncertainty about the likelihood of these harms. Given this, the BT advocate could then add two claims about what happens when both these emotions are engaged. First, since anger is a more powerful emotion than anxiety, it tends to win out with regard to shaping individuals’ subsequent behavior. Second, given the high degree of overlap in the felt experiences produced by the anger and anxiety affect programs (e.g., both bring increased, negatively valenced arousal), when prompted to state what emotion they are feeling, some subjects happen to interpret their feelings as anger, while others see it as anxiety. Thus, the BT advocate can explain both why we get mixed results when subjects are prompted to state what emotion they are feeling and why, despite these differences in self-reports, the individuals nonetheless respond with behavior characteristic of anger, not anxiety. Moreover, because this proposal allows anger to drive behavior regardless of how subjects happen to label it, the explanation is unavailable to constructivists.

All told, we have two independent sets of experimental findings showing (at best) equivocal support for constructivism’s predictions about how projecting emotion concepts onto felt experience should shape subsequent behavior. Moreover, we’ve also learned that more biologically-oriented accounts are better able to handle the experimental findings we’ve reviewed.

4. Conclusion: Emotions, Biology, and Natural Kinds

As we’ve seen, constructivism’s purported advantage over more biologically-oriented theories lies in its ability to better explain the richness and diversity of emotional life (§1). But we have also seen that a crucial premise in this argument is the move to take accommodating our ordinary emotion talk as the standard for assessing a theory’s explanatory power. Not only are there familiar problems for
adopting such a standard (e.g., Scarantino & Griffiths 2011, Kurth 2018), but—even if we accept it—we’ve learned that there’s trouble for constructivism. In particular, the explanatory “success” constructivism secures come by way of a highly revisionary account of what emotions are, when we experience them, how they differ from moods, and the way that they shape behavior (§§2-3). Moreover, our critical observations also implicate the four constructivist theses (PC1-PC4) as the source of these difficulties. Thus it’s not surprising that more biologically oriented proposals—accounts that reject these commitments—do not face similar explanatory limitations.

Taken together, then, the arguments of this paper suggest a pair of larger lessons. First, even if we agree that constructivists are correct about what the relevant standard for assessing a theory of emotion is, we’ve learned that an adequate account must give greater place to the biological mechanisms that underlie emotions than constructivism allows. This, in turn, indicates that the constructivists’ conclusion that emotions are not natural kinds is premature. After all, if we must posit something like an affect program in order to (i) explain everyday talk and empirical findings about unconscious emotions, (ii) capture the thought that emotional experience is not radically sensitive to random situational features, and (iii) accommodate research regarding how emotions shape behavior, then we have evidence that (at least some) emotions are underwritten by mechanisms that make them plausible candidates for being natural kinds.

References


