

Some Further Thoughts on Emotions and Natural Kinds

Andrea Scarantino

Department of Philosophy and Neuroscience Institute, Georgia State University, USA

Abstract

In this brief reply, which cannot do justice to all of the valuable points my commentators have raised, I defend the view that the notion of natural kind I have introduced satisfies the ontological independence criterion and is in keeping with the commitments of realism. I also further clarify the scope of my argument against basic emotion theory, and reiterate that we should stop looking for universal theories of discrete emotions.

Keywords

affective neuroscience, basic emotion, emotion, James, natural kinds

I am grateful to the commentators for their thought-provoking remarks. Here, I shall only tackle a handful of the issues they have raised. Shweder (2012) remarks that in my account a category such as “Brahman widow in rural India” would qualify as a natural kind in anthropology, and wonders whether this entails dropping the ontological independence criterion for natural kinds entirely. It does not.

The assumption of ontological independence distinguishes *realist* from *conventionalist* accounts of kinds. The realist thinks there are naturally privileged divisions humans can discover, whereas for the conventionalist all divisions are arbitrarily made up. Being a Brahman widow in rural India undoubtedly depends on human social activities. But it does not follow that the category is arbitrarily made up (as the category of “things taller than 1 meter” would be).

Rather, it emerges as a means to the ends of predicting and explaining the social structure of rural India. Thus, formulating categories apt for induction and explanation—natural kinds in Boyd’s (1991) sense—is compatible with ontological independence. What remain ontologically independent of our classificatory practices are the causal structures (natural or social) that categories must “accommodate” to be inductively and explanatorily apt.

One consequence is that a mental state category can be both culturally specific and natural. The category of *amae*, which

designates an emotional state of child-like attachment towards authority figures, constitutes a natural kind in anthropology on account of its rich inductive and explanatory potential, but it is only instantiated in Japanese culture.

Majid (2012) worries about overreliance on the English language in emotion science. If the methodology I proposed is correct, this worry is misplaced. Emotion scientists, I argued, are in the business of discovering natural kinds of emotions. But to determine whether “fear” or “anger” are natural kinds we must only consider the set of things comprised in the respective categories in English, and figure out whether such sets are theoretically homogeneous. This determination is not affected one way or the other by whether some non-English-speaking culture bundles into a unique lexical concept what we call “fear” and “anger.”

Deigh (2012) finds the phrase “theoretically homogeneous” ambiguous between an empiricist interpretation (homogeneous as observationally similar) and a realist interpretation (homogeneous as similar in a theoretical sense). I side with the realist interpretation, which underlies Boyd’s (1991) account. In advanced stages of science, scientists group instances of phenomena together not by observationally available similarities, but by theory-based similarities grounded in theoretical definitions.

But in a realist interpretation, Deigh (2012, p. 374) continues, my argument against basic emotion theory becomes “baseless.” As he understands it, my argument is that since anger, fear, disgust, etcetera, are not natural kinds, basic emotion theory is deficient. Deigh suggests that this inference gets things the wrong way around. I’d have to first demonstrate that basic emotion theory is “unsound” on grounds other than its failure to capture natural kinds, and then conclude that the theoretical definitions it provides do not capture natural kinds.

This is because, if natural kinds are defined by theory-based similarity, it is unclear on which theory-independent grounds instances of anger, fear, disgust, etcetera, can be judged to be too dissimilar from one another to form a natural kind. Since I am a realist, I cannot base this dissimilarity claim on empiricist

grounds. But this leaves me in effect nowhere to go, barring some yet-to-be-specified notion of natural kinds that is neither realist nor empiricist.

This criticism misses a key point: membership to the categories designated by the English words “fear,” “anger,” etcetera, is not theory-based, but rather grounded in folk practice. Instances are grouped together under the folk categories of fear, anger, etcetera, by ordinary-language users in light of pretheoretical similarities (alternatively, in light of folk psychological similarities). When I say that instances of anger, fear, disgust, etcetera, are too dissimilar from one another to form a natural kind, what I mean is that the items sharing the pretheoretical similarity that makes them members of folk emotion categories do not share any further theory-based similarity.

For instance, they do not share the theory-based similarity proposed by basic emotion theorists: Some instances of fear, anger, etcetera, lack the requisite automatic elicitation, brief duration, facial expressions, neural underpinnings, etcetera, supposedly characteristic of basic emotions. My criticism of basic emotion theory, then, is not that it is deficient because fear, anger, etcetera, are not natural kinds, but because its proponents assume it to apply to all things we call “anger,” “fear,” etcetera, in ordinary English. This is the task at which basic emotion theory fails.

The conclusion that fear, anger, and so on are not natural kinds comes later in the argumentative order. It results from an inference to the best explanation for why neither basic emotion theory nor any other discrete emotion theory has managed, despite centuries of attempts, to unveil some theory-based similarity—whether based in biological or social properties—shared by all members of the relevant folk categories. The inference is that there is no theory-based similarity to be unveiled.

Deigh (2012) seems to think that emotion scientists go about their definitional business as biologists do with respect to the concept of *species*, or astronomers do with respect to the concept of *planet*. They definitely should, but the whole point of my paper was to argue that they do not. Whereas an astronomer would not dream of criticizing another astronomer’s definition of a planet because it does not apply to what the folk call “planet,” it is customary practice in affective science to criticize a scientific definition of anger—for example, the definition of anger offered by basic emotion theorists—because it does not apply to something the folk call “anger.”

This is the practice my paper aims to stop. To this effect, I have proposed we distinguish between the Folk Emotion Project, which aims to unveil pretheoretical similarities that ground ordinary usage of folk emotion categories, and the Scientific Emotion Project, which aims to unveil natural kinds not coextensive with, but similar to, folk emotion categories.¹

LeDoux’s (2012) remarks are illuminating with respect to how the Scientific Emotion Project should be carried out. In his “one emotion at a time” approach, there are a variety of systems in the brain dedicated to specific types of fear, specific types of anger, specific types of disgust, etcetera.² The job of an affective

neuroscientist is to discover which neural circuits underlie different *species* of the same folk psychological *genus* (e.g., the fear genus). This job description fits perfectly the methodological recommendations of my paper.

LeDoux (2012) points out that the circuits dedicated to the type of fear he studies—learned fear elicited by Pavlovian conditioning—cannot be extrapolated to other forms of fear. For example, the type of innate fear elicited by conspecific or predator odors involves different neural circuits than learned fear, as probably do other types of more cognitively demanding fears such as fear of falling in love or fear of being afraid. Thus, *fear in general* is not a natural kind in affective neuroscience, that is, not a theoretically homogeneous category for purposes of prediction and explanation. This is compatible with there being multiple natural kinds of fear, each served by distinct neural circuits.

The neural data, jointly with the data on phenomenology, physiology, expressions and behavior, suggest that the time has come for a new era of modesty in emotion science. The grand Jamesian dream of finding a general causal principle—“the goose which lays the golden eggs” (James, 1890, p. 449)—that can do for the understanding of emotions what the principle of natural selection has done for the understanding of species, should make room for the more modest dream of piecing together a mosaic of mini-theories of natural kinds of fear, natural kinds of anger, natural kinds of disgust, and so on.

Each natural kind of emotion will be defined by dedicated causal mechanisms, but there will be no single causal mechanism shared by all instances of fear, all instances of anger, etcetera. A fortiori, there will be no single causal mechanism shared by all instances of emotion. To put it otherwise, there is no single goose that lays all the golden eggs. We can still account for all the golden eggs, but we must do it one goose at a time.

Notes

- 1 The theoretical definitions basic emotion theorists have provided may in principle individuate natural kinds. If they do, such kinds should *not* be designated by unqualified folk emotion categories such as “fear” and “anger,” because they are not coextensive with them.
- 2 I disagree that Panksepp is a good example of a neuroscientist supporting what LeDoux (2012) calls the “basic emotions in the brain” approach. Panksepp (1998) has indeed proposed that there are innate circuits in the brain dedicated to discrete emotions, but he has been careful to designate such circuits with capitalized letters like FEAR and RAGE. This terminological choice is meant to indicate that the circuits unveiled do not map perfectly with the folk categories of fear and anger/rage. On this interpretation, Panksepp is also a “one emotion at a time” theorist. A great many basic emotion theorists, however, are less careful, and they liberally refer to the “fear circuit” and the “anger circuit,” as suggested by LeDoux.

References

- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61, 127–148.
- Deigh, J. (2012). Comments on Dixon, Scarantino, and Mulligan and Scherer. *Emotion Review*, 4, 371–374.

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

LeDoux, J. (2012). A neuroscientist's perspective on debates about the nature of emotion. *Emotion Review*, 4, 375–379.

Majid, A. (2012). The role of language in a science of emotion. *Emotion Review*, 4, 380–381.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York, NY: Oxford University Press.

Shweder, R. A. (2012). The cultural psychology of natural kinds and the deconstruction of the emotions: A comment. *Emotion Review*, 4, 382–384.