

## The lower bounds of desire

---

**Author's note:** this essay explores the question of whether animals have simple forms of desire, and if so, which forms of behavior might provide evidence of this fact. It was a winning entry in the 2016 Annual Essay Prize at the Centre for Philosophical Psychology, University of Antwerp.

### (1) Introduction

One influential philosophical account of desire treats it as a species of propositional attitude, possessing broadly the same kinds of content as belief while differing in direction of fit. However, this arguably neglects more basic forms of desire. It seems an open possibility, for example, that animals that lack propositional attitudes might still have simple desires mediated by sensations like hunger and thirst. In this essay, I will argue the case for the existence of these basic desires, and suggest a strategy for understanding them. I begin by briefly discussing some different approaches to the psychological ascription of desires, and examine some candidate forms of behavior that might be constitutive of a creature's having a basic desire. I go on to argue that a form of behavior known as motivational tradeoff offers a particularly promising framework for attributions of basic desire.

### (2) Towards a science of desire

My cat is crouched low, stalking a sparrow in a tree. A few feet away, tadpoles teem in my garden pond, feasting on fish food. It is tempting in this case to say that my cat desires to catch the bird, and the tadpoles desire the food. However, the history of psychology and animal behavior has taught us these kinds of attributions should not be made too glibly: very often, seemingly sophisticated behavior in animals has been successfully explained in terms of surprisingly simple mechanisms.

Nonetheless, it might seem that the question of which creatures *really* have desires is one that cognitive science can meaningfully engage with, and which may admit of a properly scientific answer. Yet a large number of theorists, most notably Daniel Dennett, have argued that this approach may be misguided. Rather than thinking of desires as a kind of psychological entity whose presence must be inferred by a science of the mind, Dennett and others have argued that attributions of desire should be seen pragmatically as an explanatory strategy that may be more or less useful in a given situation. Considering a case in which someone correctly predicts that a plant will grow towards light on the assumption it *wants* to get out of the shade, Dennett states that “[s]ince it works, some plants are very low-grade Intentional Systems” (Dennett, 1976).<sup>1</sup>

Dennett's pragmatic approach to psychological explanation has long met with resistance, famously in the work of Fodor and more recently in the work of Tyler Burge (2010). Burge argues that we should clearly distinguish between states that are strictly *representational* such as perception, belief, and desire and more basic forms of information processing like sensory registration. These assumptions, as he puts, “are sufficiently well entrenched, not only in common sense but in serious scientific theory, not to require extensive support. What they need is explication, sharpening, delineation” (2010: 27).

---

<sup>1</sup> Dennett's view is more complex than this brief description conveys. See “Real Patterns” (1991) for a more nuanced discussion.

In what follows, I will adopt a Burgean strategy towards desires, treating them as a scientific kind in good standing whose nature psychologists and philosophers should aim to illuminate. While I will not attempt to give a worked-out defense of this approach, I would suggest that even if we are broadly sympathetic to the Dennettian strategy, we can nonetheless agree there may be important differences among the states we call desires in different contexts of psychological explanation. These differences, as I will argue below, license us to draw some fairly sharp distinctions between different kinds of systems with desires. Even if theoretical considerations ultimately push us to use the term “desire” fairly liberally across these systems, the distinctions themselves may still matter for the purposes of scientific explanation. Thus, much of what I will say below may nonetheless be able to be reconstructed on a Dennettian approach.

If we are to treat desires as distinctive psychological kinds, the first question we face is what sort of psychological processes or mechanisms underpin them, and whether desire should be regarded as having a single distinct structure or as admitting of different forms. One prominent approach has been to treat desires as a species of propositional attitude (Fodor, 1981). There are many reasons why this approach is attractive. First, desires seem to interact in systematic ways with beliefs: if I desire raspberry ice-cream, and I come to learn that Vincenzo’s Ice Cream Parlor sells raspberry ice-cream, then *ceteris paribus* I will pay Vincenzo a visit. Second, desires seem to be capable of having sophisticated contents of a kind that are found elsewhere in the mind only at the level of full blown cognitive as opposed to perceptual or sensory states. These sophisticated contents include conditionals, counterfactuals, and negations. While I might perceive something as a cat, I cannot perceive something as “a cat if it’s raining but only on Tuesdays, weather permitting” (cf. Burge, 2014). By contrast, I can desire all sorts of highly abstract and logically complex states of affairs (Dennett, 1987: 20). Finally, note that in the human case, desires seem on the face of it to satisfy the generality constraint, widely held to be constitutive of representations with conceptual content (Evans, 1982). For example, if I can desire that it be sunny in New York tomorrow, and desire that it not rain in London next week, I can also desire that it not rain in New York next week.

These considerations certainly militate in favor of the view that *some* desires are full-blown propositional attitudes. However, it does not follow that all desires fit this profile. For one, many paradigmatic forms of desire, and in particular the ‘primordial emotions’ of hunger, thirst, the desire to breathe, and so on (Denton et al., 2009), have distinctive characters that do not admit of the same degree of representational sophistication as typical propositional attitudes. For example, feeling desperately hungry arguably constitutes a case of a desire for food, but no sensation of hunger could all by itself express the content “desire for food but only if it’s a Tuesday”.

Of course, one might deny that feelings of hunger, thirst, and so on could constitute desires all by themselves; perhaps it is only once these feelings feed into the appropriate propositional attitudes that they warrant the title of desires. I suggest we resist this move however. First, it would mean that the compelling tendency to ascribe desires to non-human animals would, in many cases at least, be deeply misguided (Thagard 2006). And while most of us might allow that we may be anthropomorphizing in attributing desires to woodlice, it is far harder to deny that a hungry dog really wants to eat the food in front of him.<sup>2</sup> Second, as noted above, these primordial feelings are *canonical* desires within folk psychology: experiences of intense hunger or thirst are manifestly

---

<sup>2</sup> Another response to this worry is to be more liberal about the range of animals that possess propositional attitudes (Carruthers, 2009). However, I take it as a more controversial hypothesis, both intuitively and by the lights of philosophical psychology, to claim that a frog has propositional attitudes than to claim that it wants to eat.

instances of wanting something, even when their behavioral effects are overridden by more reflective desires not to indulge them. While we may of course recommend revision of folk psychology, if our aim is to identify the scientific basis for our existing forms of psychological explanation, we will be well advised to pay careful attention to our central pretheoretical commitments.

### **(3) A motivational tradeoff framework for desire**

I would suggest, then, that we should be open to a kind of pluralism about desires, allowing for both propositional desires and more basic forms of desire anchored in the satisfaction of core bodily needs (unified perhaps, by their direction of fit). With this move in place, we are now in a position to ask what the conditions for having these basic desires might be.

One common approach among both philosophers and scientists is to relate desires to a capacity for pleasure and displeasure (Cabanac, 2010; see also Aristotle, *De Anima*, 2.3.1). On one simple version of this proposal, it is constitutive of a state's being a desire that it result in pleasure (or the relief of displeasure) once its object is attained. There is much to be said for this proposal, but I will briefly note two reasons why it may not be the most promising option. First, the notions of pleasure and displeasure are themselves somewhat theoretically mysterious, even if they are phenomenologically familiar. To explain what is involved in a creature's having a desire by introducing the notion of feeling pleasure is thus not by itself hugely explanatory. Second, there is a growing body of empirical work that suggests that intensity of pleasure can be dissociated from intensity of desire. This is familiar in the human case from phenomena such as addiction, but also seems to be true of simpler desires in hyperdopaminergic mice (Berridge, 2004).

An alternate strategy might attempt to ground basic desires in the capacity for some form of behavior or learning. A natural candidate in this regard is operant conditioning. According to one simple proposal, we might say that to have a basic desire is to be in a state the satisfaction of which can positively reinforce a behavior. Thus, a creature has a basic desire for food at time  $t$  if and only if the presentation of food at  $t$  can positively reinforce a behavior performed at  $t$ . A similar story could account for desires with negative content, such as a desire for the cessation of pain. A major advantage of this approach would be that it would be able to operationalize basic desires, allowing us to quickly establish the range of creatures that possess them. Indeed, there is already a good body of evidence suggesting, for example, that insects and gastropods can undergo operant conditioning, but, at least going by existing evidence, Cnidarians cannot (Perry, Barron, & Cheng, 2013).

An important objection, however, is that operant conditioning can also occur within subpersonal systems. Thus, rats with transected spinal cords are still capable of being conditioned just within the isolated section of spinal tissue. Specifically, they can be trained to keep their rear legs elevated to avoid electric shocks, even though their rear legs cannot communicate with the rat's brain at all (see Allen et al., 2009 for a review). A view that simply identifies desire with negative reinforcement must seemingly say that rats' spinal cords can want things all by themselves. While not impossible, this is again sufficiently at odds with our folk psychological conception of desire as to warrant looking for a more constrained account of desire.

I now wish to suggest one such alternate option in the form of motivational tradeoff behavior. This is the observed tendency of some organisms to willingly undergo one normally avoided state in order to avoid another that is even more repellent, or to come to seek out a normally neutral stimulus if it will relieve a negative stimulus. Motivational tradeoff behavior in

various forms has been demonstrated in a wide range of animals and occurs spontaneously without extended schedules of reinforcement. For example, rats, in a reversal of normal preferences, will prefer a light chamber to a dark chamber in order to avoid unpleasant mechanical stimulation of an injured paw (LaBuda & Fuchs, 2000). Broiler chickens are normally disinclined to jump over high barriers unless food-deprived, but will do so spontaneously without any food deprivation in order to get away from highly-crowded enclosures (Buijs, Keeling, & Tuytens, 2011). Zebrafish prefer to swim to an environmentally enriched chamber rather than a barren, brightly lit one, but their preferences are reversed if the fish are injected with an irritating acid and the bright chamber filled with an analgesic (Sneddon, 2013). Even hermit crabs will more slowly vacate a shell following electric shocks if the species of shell is highly preferred (Elwood & Appel, 2009).

The capacity of organisms in these paradigms not merely to respond to bodily damage or other negative states but to tailor their behavior flexibly in accordance with the relative intensity of such states bespeaks a cognitive sophistication that makes attributions of genuine desires seem far from scientifically absurd. Moreover, the states in motivational tradeoff paradigms go beyond operant conditioning insofar as they do not merely negatively reinforce individual behaviors, but are poised to influence selection between behaviors at the whole organism level. Finally, a creature's ability to display motivational tradeoff satisfies a key pretheoretical characteristic of desire insofar as it bespeaks a capacity to assign different weightings to different outcomes and decide between them, just as in daily life, I might prioritize thirst over hunger, or my desire for alleviation of pain over the unpleasant taste of medicine. Finally, I will note that motivational tradeoff behavior is apparently confined to vertebrates and a few species of invertebrate, and is seemingly absent in creatures like plants or bacteria that are unlikely targets for psychological explanation in the first place.

#### **(4) Conclusion**

The notion of motivational tradeoff has previously been suggested as a scientific criterion for assessing conscious pain in non-human animals (Sneddon, et al., 2014) and even consciousness (Cabanac et al. 2009), but may also constitute an even better candidate as a signature of creatures with basic desires. However, a good degree of caution is of course in order. It may emerge, for example, that we should draw further distinctions among different types of motivational tradeoff (perhaps the mechanism used by crabs is distinct from that used by vertebrates) and it may turn out that some forms but not others should be identified with basic desires proper.

Nonetheless, I think that motivational tradeoff merits close attention from philosophers and scientists seeking to identify the lower bounds of desire. Additionally, the phenomenon raises a number of further philosophical issues which it is worth flagging. One central question concerns what sort of mental content is possessed by those states that feed into the decision-making underpinning motivational tradeoff behavior. Insofar as creatures in these paradigms are able to select flexibly among states relating to different bodily needs, it is natural to interpret these states as possessing a shared representational component common to them all which determines the weighting they are afforded by a creature's decision-making mechanisms. This in turn might suggest that this component is *amodal*, insofar as it is not tied to any single form of sensory representation. A very familiar class of amodal representations are, of course, concepts: I can token the concept 'glass' in response to tactile, auditory, or visual stimuli. Thus, one way to understand the proposed amodal component involved in motivational tradeoff would be via conceptual encoding: perhaps animals encode bodily sensations in respect of a primitive motivational concept, the content of which served

to indicate the degree to which the state was to avoided or sought out.

However, this is not the only viable approach. One might also allow that there are amodal forms of *nonconceptual* representation that could be bound to or otherwise indexed to sensory states. Broadly similar proposals have been advanced, for example, in giving accounts of the content of the mental maps seemingly utilized by many non-human animals in navigation (Burge, 2014).

Despite these many unanswered questions, I would suggest that motivational tradeoff is a promising behavioral signature of basic desires. It captures a key pretheoretical component of desire, namely the idea that desires come in different degrees and are commensurable with one another; it is closely connected to sensations corresponding to basic bodily needs; and it encompasses many organisms to which we may be inclined to attribute desires without collapsing into excessive liberalism. In seeking an account of which animals have desires, then, motivational tradeoff looks a good place to start.

## REFERENCES

- Allen, C., Grau, J., Meagher, M.W. (2009). The Lower Bounds of Cognition: What Do Spinal Cords Reveal? In Bickle, John (ed.) *The Oxford Handbook of Philosophy and Neuroscience*. Oxford University Press.
- Berridge, K. 2004: Motivation concepts in behavioral neuroscience. *Physiology and Behavior*, 81, 179–209
- Buijs, S., Keeling, L. J., & Tuytens, F. A. (2011). Using motivation to feed as a way to assess the importance of space for broiler chickens. *Animal Behaviour*, 81(1), 145-151.
- Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press.
- (2014). Perception: Where Mind Begins. *Philosophy* 89 (3):385-403.
- Cabanac, M. (2010). *The fifth influence: The dialectics of pleasure*. Bloomington, IN: IUiverse.
- Cabanac, M., & Cabanac, J., & Paren, A. (2009). The emergence of consciousness in phylogeny. *Behavioural Brain Research*, 2(198), 267–272.
- Carruthers, P. (2009). Invertebrate concepts confront the generality constraint (and win). In R.W. Lurz (ed.), *The Philosophy of Animal Minds*. Cambridge University Press 89-107.
- Dennett, Daniel C. (1987). *The Intentional Stance*. MIT Press.
- (1991). Real patterns. *Journal of Philosophy* 88 (1):27-51.
- (1976). Conditions of personhood. In Amelie Oksenberg Rorty (ed.), *The Identities of Persons*. University of California Press
- Denton, D., Mckinley, M., Farrell, M., & Egan, G. (2009). The role of primordial emotions in the evolutionary origin of consciousness. *Consciousness and Cognition*, 18(2), 500-514
- Elwood, R. W., & Appel, M. (2009). Pain experience in hermit crabs? *Animal Behaviour*, 77(5), 1243-1246.
- Fodor, J. A. (1981). *Representations: Philosophical essays on the foundations of cognitive science*. Cambridge, MA: MIT Press.
- LaBuda, C. J., & Fuchs, P. N. (2000). A Behavioral Test Paradigm to Measure the Aversive Quality of Inflammatory and Neuropathic Pain in Rats. *Experimental Neurology* 163: 490–494.
- Perry, C. J., Barron, A. B., & Cheng, K. (2013). Invertebrate learning and cognition: Relating phenomena to neural substrate. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(5), 561-582.
- Sneddon, L. (2013). Do painful sensations and fear exist in fish. In T. A. van der Kemp, & M. Lachance (Eds.), *Animal suffering: From science to law, international symposium*. Toronto, Canada: Carswell, 93-112.
- Sneddon, L. U., Elwood, R. W., Adamo, S. A., & Leach, M. C. (2014). Defining and assessing animal pain. *Animal Behaviour*, 97, 201-212.
- Thagard, P., 2006. “Desires are not propositional attitudes,” *Dialogue*, 45: 151–156.