

Desires Are Not Propositional Attitudes

Paul Thagard
University of Waterloo
pthagard@uwaterloo.ca

Thagard, P. (forthcoming). Desires are not propositional attitudes. *Dialogue: Canadian Philosophical Review*.

Tim Schroeder's *Three Faces of Desire* is an excellent contribution to naturalistic philosophy of mind, in the grand tradition of Aristotle, Hume, Quine, and Patricia and Paul Churchland. He develops a rich and plausible theory of desire with strong connections both to neuroscience and to everyday cases. This theory ties desire to reward rather than just to behavioral dispositions or the experience of pleasure. Far from abandoning the ordinary concept of desire, he enhances and expands it through the development of a generally convincing neuropsychological theory.

However, there is one crucial respect in which Schroeder has failed to throw off the intellectual shackles of linguistic philosophy. He assumes, in keeping with the understanding of mental states that derives from Frege and Russell, that desires are propositional attitudes. He states his Reward Theory of Desire as follows: "To have an intrinsic (positive) desire that *P* is to use the capacity to perceptually or cognitively represent that *P* to constitute *P* as a reward." (Schroeder, 2004, p. 131). All the desires discussed by Schroeder are desires *that* something.

Taking desires as propositional attitudes is implausible for humans, and even more implausible for animals with more limited representational capacities. Here are some examples:

Andrew desires a beer.

Ron desires ice cream.

Brad desires Angelina.

The lab rat desires a food pellet.

Grammatically, none of these sentences requires a “that” clause; the object of desire is a thing or kind of thing, not a state of affairs.

In order to maintain the view that desires are propositional attitudes, one would have to paraphrase these sentences along the lines of construing “X desires O” as “X desires that X should have O”:

Andrew desires that he drink a beer.

Ron desires that he eat ice cream.

Brad desires that he sleep with Angelina.

The lab rat desires that it eat a food pellet.

But these paraphrases are at best awkward, and even more awkward are paraphrases of sentences involving words that dictionaries relate closely to “desire”, such as “want” and “long for”. Only a non-native speaker of English would paraphrase “Andrew wants a beer.” by “Andrew wants that he should have a beer.” Ben-Yami (1997) provides general arguments against paraphrasing sentences involving mental states such as wanting, knowing how, and trust in terms of propositional attitudes.

It would be folly, however, to try to derive conclusions about psychology from the structure of English grammar. Considerations of comparative psychology and neuroscience provide stronger reasons for supposing that desires are directed toward things rather than toward states of affairs. Schroeder makes a strong case that desires are related to reward, but it is remarkable that the reward system in human brains is very similar to structures found, not only in smaller mammals such as rats, but even in reptiles.

This makes it as plausible that a lizard can desire a fly in much the same way that a person can desire a steak. But lizards, rats, and even chimpanzees are highly limited in their ability to represent complex situations, compared to humans. Holyoak and Thagard (1995, ch. 3) review evidence that only in primates is there clear evidence of explicit knowledge of relations, and even chimpanzees are limited in their ability to think about higher order relations such as *cause*. Desiring food requires only the ability to represent food, which any animal can do. But an animal cannot desire *that* it should have food to eat without a more complex representation that includes something like the relational expressiveness that humans gain from language.

Moreover, the construal of “X desires O” as “X desires that X should have O” introduces another representational problem: in order to desire anything, an animal would have to represent itself. Self-representation is not a problem for humans and the great apes such as chimpanzees and orangutans. Experiments such as mirror recognition tasks suggest that these apes have at least a limited sort of self concept, so perhaps they are able like humans to desire something about themselves. But for less cognitively endowed animals there is no evidence of a capability of self representation that would be required for a propositional construal of desire. Animals desire food, not that they should have food. Humans are usually the same.

It is possible that the increased self-representational capacity of apes and humans derives from a recently discovered kind of cell, the spindle neuron, which is found in the anterior cingulate cortex of humans and great apes, but not in other primates (Allman et al, 2001). Spindle neurons are large cells with only two dendrites, one extending upward and the other downward, which establish long-distance connections with many areas of

the brain. The concentration of spindle neurons in humans is much greater than in chimpanzees, gorillas, and orangutangs. It is known that the anterior cingulate cortex is important for focusing attention and processing rewards, and Allman et al. (2001) conjecture that spindle cells may serve to coordinate the activity of diverse parts of the brain to achieve focus on solving problems and self-control. Whether self-representation depends on these special kind of neurons or merely on the much greater cortical capacity of humans and great apes, the inadequacies of other animals undermine any propositional-attitude account of desire that requires that they be able to represent themselves.

Accordingly, we should construe the non-propositional representation of desire as fundamental and the conception of desires as propositional attitudes as applicable only to some animals some of the time. People are capable of desiring that a state of affairs come into being, for example that there be a NHL hockey season next year. But the neuropsychologically plausible paraphrase runs in the opposite direction from the one that presumed by any defender of the propositional-attitude account: “X desires that object O has feature F” becomes “X desires O to be F”. People, like animals, often have the more direct, non-propositional kind of desire.

Abandoning the overly linguistic account of desires as propositional attitudes is not at all fatal to the core of Schroeder’s reward theory of desire, which can be rephrased as: To have an intrinsic (positive) desire for *Y* is to use the capacity to perceptually or cognitively represent *Y* to constitute *Y* as a reward. Here *Y* can range equally over states of affairs, objects, or features of objects, depending on the representational capabilities of

whatever animal has the desire. There is psychological evidence for many kinds of mental representations besides propositions (Thagard, 2005).

Although the concept of representation is crucial to Schroeder's reward theory of desire, he has little to say about it. The theories of linguistic representation familiar to philosophers and even the broader theories of mental representation found in cognitive science are of little help in characterizing how an animal represents the object of its desire. An alternative can be found in recent neurocomputational theories of representation that describe how groups of neurons can encode features of the world (Eliasmith and Anderson, 2003). We can define a *neural structure* as a complex <neurons, connections, spiking behaviors> that consists of a set of neurons, a set of synaptic connections among them, and a set of behaviors of individual neurons that specifies their patterns of spiking as determined by the spiking behaviors of all those neurons to which they are connected (Thagard, forthcoming). A neural structure is a representation if its connections and spiking behavior enable it to relate to perceptual input and/or the behavior of other neural structures in such a way that it can be construed as standing for something, which can be a thing, concept, or whatever. Inferences are transformations of neural structures.

Such representational neural structures suffice for understanding how representation of Y can be part of desiring that Y. More complicated is seeing how, in Schroeder's terminology, this representation can constitute Y as a reward, but we can conjecture that the answer depends on there being synaptic connections and temporal coordination of spiking behaviors between neural structures representing Y and neural structures in the brain's reward centres.

Such neural theories of mental representation are in their infancy, much like genetic theory after the discovery of the structure of DNA. In the 1950s, it became clear that investigation of DNA would reveal much about the structure and function of genes, but working out the relation has taken decades. Similarly, I expect it will take decades to work out how neural structures represent objects, properties, and states of affairs. Nevertheless, a neural-structure view of how the brain represents is much more compatible with the kind of naturalistic theory of desire that Schroeder offers than the propositional attitude view that still survives in language-oriented philosophy of mind.

Another advantage in thinking in terms of neural structures and processes is that it becomes possible to link Schroeder's theory of desire directly with neural processes of decision making. Part of the appeal of the propositional-attitude conception of desire is that it melds with a propositional conception of belief to produce an intuitive explanation of action. Why does X do action A? Because X desires that P and X believes that A will contribute to accomplishing P. Belief and desire are both propositional attitudes, and combining them explains why people do what they do. For everyday purposes, belief/desire explanations are often adequate, but scientific explanations of behavior require much deeper accounts of the neural representations and processes that generate the actions of humans and other animals.

Wagar and Thagard (2004) propose a neurocomputational model of decision making that fits well with Schroeder's theory of desire. The model is called GAGE because it explains behavioral problems of Phineas Gage and other patients who have decision making deficits as the result of brain damage. GAGE construes decision as a process of cognitive-affective integration, which requires dynamic interconnections

between areas of the brain involved in representing things and their properties and areas involved in representing their emotional value. The key emotional brain areas in the model are the amygdala and nucleus accumbens which are known to be very important in animals' assessments of positive and negative value. As Schroeder mentions, the nucleus accumbens is crucial for reward processing. The key cognitive areas are the hippocampus, which provides context through memories of previous situations, and the ventromedial prefrontal cortex, which connects emotional regions such as the amygdala with areas of the prefrontal cortex responsible for high level reasoning. In keeping with the arguments of Damasio (1994), this model shows how effective decision making depends on integrating cognitive and emotional information by means of dynamic, reciprocal interconnections among multiple brain areas. Because a crucial part of the model is the assessment of value in the nucleus accumbens, the model is highly compatible with Schroeder's reward theory of desire when it is broadened beyond propositional attitudes.

The resulting neurocomputational account of decision making has much greater explanatory potential than previous philosophical or psychological accounts. The everyday belief/desire explanation of action often suffices for ordinary social purposes, but it fails to give anything like the kind of deep mechanistic explanations that have contributed so much to physics, chemistry, and biology. Since the 1940s, economics has presupposed a theory of decision making based on utility and probability, but a wealth of behavioral experiments have shown that its fundamental assumptions are psychologically unrealistic. Now, thanks to new scanning technologies that enable neuroscientists to observe brains in the process of making decisions, we see the

emergence of neuroeconomics as an interdisciplinary field that uses knowledge about brain mechanisms to inform economic theory (Camerer, Loewenstein, and Prelec, 2004).

We can schematize these three approaches to decision making as:

(1) desires + beliefs -> actions

(2) utilities + probabilities -> actions

(3) emotional brain areas + cognitive brain areas + interactive mechanisms -> actions.

It is another advantage of Schroeder's reward theory of desire that it shows how the rather crude everyday theory of action (1) can be related to the burgeoning scientific theory sketched by (3).

REFERENCES

- Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The anterior cingulate cortex: The evolution of an interface between emotion and cognition. *Annals of the New York Academy of Science*, 935, 107-117.
- Ben-Yami, H. (1997). Against characterizing mental states as propositional attitudes. *Philosophical Quarterly*, 47, 84-89.
- Camerer, C., Loewenstein, G. F., & Prelec, D. (2004). Neuroeconomics: Why economics needs brains. *Scandinavian Journal of Economics*, 106, 555-579.
- Damasio, A. R. (1994). *Descartes' error*. New York: G. P. Putnam's Sons.
- Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation and dynamics in neurobiological systems*. Cambridge, MA: MIT Press.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press/Bradford Books.
- Schroeder, T. (2004). *Three faces of desire*. Oxford: Oxford University Press.
- Thagard, P. (2005). *Mind: Introduction to cognitive science* (2nd ed.). Cambridge, MA: MIT Press.
- Thagard, P. (forthcoming). Abductive inference: From philosophical analysis to neural mechanisms. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Cognitive, mathematical, and neuroscientific approaches*. Cambridge: Cambridge University Press.

Wagar, B. M., & Thagard, P. (2004). Spiking Phineas Gage: A neurocomputational theory of cognitive-affective integration in decision making. *Psychological Review*, *111*, 67-79.