

# Desire, Reward, Feeling: Commentary on *Three Faces of Desire*

ANDREW BROOK

*Carleton University*

Tim Schroeder has written a book full of rich insights from neuroscience. It offers a wealth of telling examples and wonderful (sometimes a bit too wonderful) analogies. As a result, it is a great read.<sup>1</sup>

The three faces of desire are, in a nutshell, that desires are motivating, that satisfying desires is usually pleasurable, and that desires determine what will count as rewards and punishments. Schroeder's view of how these knit together goes like this. Pleasure and displeasure are a result of desires so cannot be what desire consists in. We can desire without being motivated and vice-versa; this double dissociation entails that providing motivation cannot be what desire consists in, either. That leaves the link to rewards and punishments. Schroeder maintains that, because it is the reward system in the brain that determines what is pleasurable and what is motivating, "desires are realized . . . by the biological reward system" (p. 37). As he says in the crucial fifth chapter, "to desire is to be so organized that tokened representations of . . . well-being, if they occur, will contribute to the production of reward signals" (p. 134). More generally, determining what will count as rewards and punishments is "the essence of desire" (p. 178). Even more generally, what a desire *is* is determined by what it *does*.

One of the great strengths of this book is that it develops plausible accounts of motivation, pleasure, and reward, puts them on top of a huge helping of neuroscience, and then integrates the results into a complex, powerful theory. I have some problems with some of the bits but I want to say that this is one book where the whole is very much more than the sum of its parts. I cannot begin to reflect in a few short pages the cumulative power of Schroeder's analysis. That said, there are well-known problems with ultrafunctionalist accounts of mental kinds of the sort that he develops.

## Desire and Reward

Let us start with the crucial link between desire and reward. At the end of chap. 3, on rewards and punishments, Schroeder concludes that there are two good theories of reward: the desire theory and the learning theory. The desire theory of reward claims that satisfying our desires is rewarding, not a very controversial claim. The learning theory of reward claims, roughly, that for an event to be a reward it must contribute to reinforcement learning (p. 66).

Questions can be raised about the latter theory. For example, if being a reward is simply contributing to reinforcement learning, then being a reward has nothing essential to do with giving pleasure. Now Schroeder has some excellent examples to show that rewards do not *always* have to give pleasure, but could we make sense of the idea that events could be rewarding and yet *never* give pleasure (p. 61)? I can't. The dissociation that Schroeder wants to claim between being a reward and pleasure appears to run too deep. This is a specific case of a general problem. Pleasure has a phenomenology (it is like something to feel pleasure). Schroeder tends to be over-zealous about disconnecting all his targets from phenomenology.

He also *connects* his main target, desire, too strongly to some other things. At the end of the chapter, we read that *both* theories of reward can be correct. One blinks. Where did that come from? From introducing a new move. Schroeder suddenly adopts (here, provisionally) a *learning* theory of desire. Ahhhhh! If that were true, of course, it would indeed follow that the desire theory of reward and the learning theory of reward would be compatible -- they would be parts of a single theory! (As Schroeder recognizes, [p. 70].) Desire determines what is rewarding, and being rewarding contributes to learning.

But what about the new move? Even if he is right about reward, that **P** being a reward is simply a matter of **P** contributing to reinforcement learning, how could a learning theory of *desire* go?

Preliminary question: What does Schroeder mean when he talks about desires? Is he talking about what everyone calls desires or does he have in mind some artefact of neuroscience or whatever?

Certainly it is important to know some neuroscience. For one thing, it helps us to avoid neurologically implausible theories. Furthermore, neuroscience has revealed some fascinating constraints on cognition. Of the huge number of such constraints that Schroeder describes, one of the most interesting to me is that the motivation system and the pleasure system are quite different systems in the brain. This by itself is enough to make theories that link the two conceptually a bit suspect.

But could neuroscience tell us that something no one would recognize as a desire really is a desire? Could neuroscience tell us that something everyone agreed is a desire really is not a desire? Could neuroscience tell us that there are no such things as desires (eliminativism with a vengeance)? For Schroeder, the answer is no. Perhaps neuroscience could tell us that the events we group under “desire” are less homogenous than we thought -- though Schroeder thinks it does not even do that. Neuroscience can constrain, clarify, redescribe, reclassify, but it cannot change the subject. For Schroeder, the target of his analysis is desires as we encounter them in ourselves and others.

So what are desires like that we encounter in ourselves and others? For Schroeder, desiring **P** has two elements: a representation of **P** and use of that representation (or the capacity for it) to “constitute **P** as a reward” (p. 131), where “constitute” means something like “cause to be” or “bring it about that.” To constitute **P** as a reward, the representation of **P** leads one to produce reward signals (p. 134). This is a learning theory of desire because by desiring we learn what is rewarding.

*Prima facie*, there is something very peculiar about this theory. One is inclined to say that, by the time we desire something, we do not have to *learn* that getting it would be rewarding; we already *believe* that. Even more importantly, Schroeder's theory does not connect very well to desires as we encounter them. Suppose we found a mechanism in the brain that reliably generates reinforcement-learning signals, reliably contributes to things becoming rewarding for us. Why could this circuit not generate learning signals in the absence of all desire? More precisely, why could we not represent **P** and bring it about that getting **P** is rewarding without *desiring P* at all? Let us spell this objection out.<sup>2</sup>

Recall the statement that we quoted earlier: if we desire something like pie or well-being, “tokened representations of pie or well-being would contribute to the production of reward signals” (p. 134). Why could all this not happen without my desiring pie or well-being at all? Perhaps Schroeder means this production business to be a necessary condition of desire (so that desiring is sufficient for contributing), but not the reverse, not a “sufficient” condition? No, he says, if reward is the essence of desire, then, if **P** leads to reward signals, it entails that **P** is desired (reconstructed from p. 129), i.e., both necessary and sufficient, presumably. Yet, generating reward signals is patently not sufficient for having a desire. The first time that many people consume an alcoholic beverage, they have no particular desire to drink it. They just go along with the crowd. Then the alcohol hits and they discover that the feeling is very pleasant -- in other words, rewarding. Something being rewarding is a good way for me *to come* to desire it but it can be rewarding *prior* to my desiring it. There can be reward without desire.

In addition, as Schroeder himself allows in chap. 2, desiring **P** is a good way *to make* the getting of **P** rewarding. If so, it would again follow that desire and reward are two different things.

## Phenomenology and Desire

So where does Schroeder go wrong? One thing missing from his account is phenomenology. Desires have a distinctive phenomenology, to be sure one that is hard to describe in a noncircular way, but distinctive nonetheless -- occurrent desires at any rate.<sup>3</sup> I crave this substance, long to see that person again, would die for another piece of something I just ate. This does not mean that the desireability of desires is a good guide to anything else about them. As Schroeder says, even the actual strength of a desire might be misrepresented in how it feels to the person who has it (p. 143). But phenomenology is part of the essence of desire. If one called a desire something that had no phenomenology, was phenomenologically neutral, or even worse came with a conscious aversion (negative phenomenology), what could one have in mind? Similarly, how could anything be a desire without one having an urge to obtain the object of desire? But urges have a phenomenology. At any rate, most of them do.<sup>4</sup>

If Schroeder's story were all there was to say about (what we call) desire, it would be more natural to say that he has shown there are no desires, that the concept of desire does not refer to anything real, rather than saying he has offered us a new theory of desire. He offers a sensitive argument for why his account is not eliminativist in this way (pp. 164-73). Though he recognizes that phenomenology being merely a result in his picture diverges from the way desires as we encounter them seem to be, by my lights he attaches too little significance to the conceptual impact of the divergence.

The source of both the phenomenology gap and Schroeder's too light treatment of it is, I suspect, his ultra-functionalism, the view that desires *are* what they *do*. Contrary to this view, desire is a two-headed phenomenon, something that is also true of many other mental kinds. One head is having certain effects, the other is having a certain feel. Schroeder's theory of desires has only one of these elements. Yet the effects without the feel are not enough to make something a desire. (The same is likely true in reverse -- the feel without the effects is likely not enough to make something a desire. An example might be a craving-like feeling with no craving, no inclination to seek the craved object.)

Could a desire currently active in shaping my intention forming and decision making not be like anything to have, or carry the opposite phenomenology to what it seeks? Maybe in a few oddball cases, but not, I think, in general. Schroeder holds that characteristic phenomenology is merely an effect of desires. I hold that, in general, it is part of what they are. The relationship between the effects and feeling are not just causal. The two are both constitutive elements of desire.

Now Schroeder himself is well aware that desire often goes with a distinctive phenomenology. His book is full of references to it: felt urges to act (p. 144) (a felt urge to jump off an eleventh-floor balcony, for example), the urge in a person with Tourette syndrome to repeat a tic (pp. 125, 144), and so on. So, what divides us? Possibly two things.

Schroeder insists that "desires, like all robust entities, can exist without being represented" (p. 173). Maybe. But even if desires can exist without being represented, it does not follow that they can exist without feeling like something (for which *being* a representation might be enough, no need to be represented).

He also believes that desires can exist without being conscious. He may also believe that if a state feels like something, its bearer must be conscious of what it feels like (this is not clear). Since we clearly have desires of which we are not conscious, the conclusion would be that such states do not feel like anything. However, this inference is forced. A better idea is that states can have felt qualities of which their bearer need not be conscious.

Long ago Dennett distinguished two elements in belief, which he called opinion and belief. Opinion is how a belief appears to us, what (if you like) we think we believe. Belief is the effect a belief has on

us: on behaviour, other beliefs, and so on (Dennett 1978, p. 304). I think desire has a dual nature of much the same kind.<sup>5</sup> One response to this duality is eliminativism: we should stop using such concepts and replace them with better behaved ones (Stich 1983). Personally, I think that these dual-headed creatures play a crucial role in our understanding of ourselves as cognitive creatures. Desire has this duality.

## **Desire and Pleasure; Desire and Motivation**

If Schroeder's attempt to reduce desire to representation plus production of reward signals does not work, what about his attempt to divorce it from the other two faces, pleasure and being motivated? I have no problem with divorcing desire from pleasure. Desire is often for pleasure and satisfying desires is often extremely pleasurable, but the very possibility of a causal relationship between the two speaks to them being different from one another.<sup>6</sup> Motivation, however, is another matter. There may well be, as Schroeder urges, forms of motivation that do not involve desire (p. 21); such motives may be a feature of some addictions. However, how could desiring not be motivating? Even if we grant that desire is not necessary for all motivation, how could it fail to be sufficient?

Let us examine Schroeder's arguments concerning pleasure and motivation. Against the idea that motivation is constitutive of desire, he mounts counter examples such as Tourette syndrome victims who have an urge and therefore a motive to produce tics (bits of repeated, stereotypical behaviour) even when they desire not to do so. Against the idea that pleasure is constitutive of desire, he offers examples in which one is moved but one's "hedonic tone" is neutral -- shaving every morning, for example. The trouble is, these examples show at most only that desire is not necessary for motivation and is not sufficient for pleasure. About the Tourette example one could urge, "Well, even if desire is not necessary for motivation, it is still sufficient." About the example of shaving, one could argue that even if pleasure is not necessary for desire, it is still sufficient. And the link between motivation and desire, pleasure and desire will have been only half severed, not severed entirely.

Notwithstanding this objection, I agree, as I said, with Schroeder about desire and pleasure. But not about motivation. He does offer thought-experiments designed to show that desire is not sufficient for motivation. He cites Galen Strawson's "weather watchers," people who can watch but not act, for example (p. 20), stroke victims who suffer from akinetic mutism (they are unable to move or speak), severe Parkinson patients who are unable to move (pp. 173-74), and so on. It is plausible, indeed compelling, to say that they still have desires, including the desire to move. But they do not have motives. So desire is not sufficient for motive.

Unfortunately, while *movement* may be impossible in all these cases, it is not clear that being *moved* is impossible. Schroeder says it is: "Their motivational systems are damaged" (p. 174). But this is at the very least unclear. Grant that they are unable to move and also that they have desires, but what should we say about what is going on inbetween? I can see two possibilities, and no way to choose between them. One is to accept Schroeder's story and say that these people do not move and they have no disposition, no motive to move. The other is to say that, since they desire to move, they have a motive to move and that what they lack is the capacity to translate the desire/motive into a plan of action or into a pattern of muscle commands or something of that sort. On the second option, we have no reason to deny that desire is sufficient for motivation. I need not opt for either option. A mere impasse would be enough to undermine Schroeder's analysis. Why? Because it would leave what he calls the standard theory, that desires are a form of motivation, still standing. As I believe occurs.

## Score So Far

Here is how Schroeder's claims have fared:

Reducing desire to representation plus learning signals does not work.

Divorcing desire from phenomenology does not work.

Divorcing desire from motivation does not work.

Divorcing of pleasure and desire works.

## Three Smaller Issues

### *Trying*

As I said, one of the excellent features of this book is that it not only offers a theory of desire, it also offers theories of each of the three “faces” of desire: pleasure, motivation, and reward. In connection with motivation, Schroeder discusses trying, making an effort. This is excellent as the literature on free will, for example, mostly ignores trying and some writers render it impossible (some compatibilists, in particular). However, Schroeder treats trying as a *source* of motivation (p. 108). That cannot be right. Trying enters when what we are moved to do is settled and we have set about to do it. We just saw that Schroeder sometimes seems to conflate motive and motion. Maybe that is happening here. Trying is a main source of motion, but it is not at a source of motivation at all.

### *Moral Motives*

Near the end of chapter 5, the crucial chapter, Schroeder takes up a number of issues downstream of that of the nature of desire. Two of them are moral motivation and how to explain desire's mind-to-world direction of fit. He does not intend his treatment of either to be definitive but there are problems. The problem with Schroeder on moral motives is not what he says but what he does not say. He says that he has not found a structure of motivation adequate to capture moral motives. However, he does not consider one of the most plausible ideas, Harry Frankfurt's higher-order approach. It is not clear why. On this approach, what makes moral motives possible is the idea of higher-order motives. We can assess not just particular acts against motives but particular motives against broader and deeper motives. It is this capacity to assess our motives that allows us to be moral and also to be free. I wonder if Schroeder would find this approach as unsatisfactory as the ones he did consider.

### *Direction of Fit*

Schroeder twice considers direction of fit, the first time to determine whether a pleasure-based account of desire can capture it, and then to see whether a reward-generation account of desire can capture it. In the first instance, Schroeder says a pleasure-based account cannot capture direction of fit because such an account cannot explain why it is that, when the world does not satisfy our desires, it is the world we want to change, not our desires. He acknowledges that the issue is not clear (pp. 34-35). Some times it *would* be rational to change one's desire -- if, for example, there is no hope whatsoever of satisfying it. So let us put it this way: When the world can easily be changed to satisfy a desire, why on a pleasure-based account would it be more “rational” (Schroeder's word) to change the world rather than abandon the desire?

I think that a pleasure-based account actually has a good answer to this question. If one abandons the desire, one does not end up in a state as good as the one that would result if one changes the world. Changing the world so that it satisfies a desire would give one pleasure. Changing the desire would

leave one with nothing, neither pleasure nor displeasure. So the “hedonic approach” can explain why the mind-to-world direction of fit is rational in the case of desire.

Schroeder thinks that an reward-based account can do the trick. But can it? If desires determine what will be rewarding, he tell us, “it is part of the essential nature of desire to regulate the mind” (p. 156), but not vice-versa. Perhaps. Is this an explanation of direction of fit? I do not think so. It just tells us that desires *do* change mind and world. However, there is also a normative dimension to direction of fit. When a belief turns out to be false, it is *rational* to change the belief, not the world, and when a desire is not satisfied, it is *rational* to change the world, not the desire. This asymmetry of rationality needs to be explained. An appeal to the power of desires to shape the mind does not do so.

## Conclusion

The score of 3 to 1 against and the smaller problems notwithstanding, this is a splendid piece of scientifically informed philosophy. That I do not accept some of its arguments is par for the course in philosophy -- and not just philosophy.

## Notes

- 1 The text is also the cleanest I have ever seen. I spotted only three mistakes in 180 pages!
- 2 I hate to pull this move because an exactly parallel move is so often used to argue that consciousness is something weird and wonderful, noncognitive and even nonphysical. Unfortunately, here the move has force.
- 3 Occurrent desires are desires at work in my volitional life at the moment. The contrast is with desires that are currently just dispositions. Dispositions need not have a phenomenology.
- 4 Here we should address the question of unconscious desire. I do not have space to do so.
- 5 Nor do the two heads need to fit neatly together in all cases.
- 6 The only felt quality that Schroeder considers in any detail is pleasure. This may have helped to make it easy for him to deny that phenomenology has a constitutive role in desire.

## References

- Dennett, D. C.  
1978      How to Change Your Mind. *Brainstorms*. Montgomery, VT: Bradford Books
- Stich, S.  
1983      *From Folk Psychology to Cognitive Science: The Case against Belief*. Cambridge, MA: MIT Press/A Bradford Book.