

Dust, ashes, and vice: on Tim Schroeder's theory of desire

for CPA symposium on *Three Faces of Desire*, May 28 2005

Ronald de Sousa

Ronald.de.Sousa@utoronto.ca

La jouissance ajoute au désir de la force.

Désir, vieil arbre à qui le plaisir sert d'engrais....

Baudelaire

0. Introduction

Tim Schroeder's book offers a beautiful convergence of philosophical, analytical, folk theoretic, and neuroscientific approaches. It presents a fine example of a scientifically informed way of doing philosophy, which anyone in the philosophy of mind should emulate. Many philosophers, however, are still reluctant to do so. For this resistance, I can think of three causes. It may be due to a high-minded commitment to the ancient ideal of pure conceptual analysis as traditionally performed: from an armchair, without any technological equipment more complex than a pipe. Or it may stem from darker motives. There is here, perhaps, much the same protectiveness about turf as motivated outrage among social scientists when biologists first attempted to remind us of the fact that we are primates, for whom sociality, like everything associated with nurture, has its roots in nature shaped by natural selection. Finally, the reluctance to accept science as a partner in philosophy may simply be a matter of sloth. Finding out the relevant neurological facts takes work. Yet sloth can motivate gratitude as well as mistrust. For myself, I readily confess I have not done the hard work. But as I regard my profession as a license to be a dilettante, I'm happy that someone else has done it. For in my view the result, in the present case, has much to teach any philosopher of mind.

Though the strategy is ingenious and the details fairly intricate, Schroeder's basic argument can be sketched without excessive distortion. He begins by showing that neither *motivational theories* nor *hedonic theories* of desire will do. In both cases, the candidate's connection to desire is neither necessary nor sufficient. He then shows the plausibility of two theories of reward: the "Contingency-based Learning Theory of Reward," and the "Desire Theory of Reward," and finds that both coincide both in their advantages and in their disadvantages. He concludes that the two theories—the CTL theory of reward and the desire theory of reward—can plausibly be merged. (p. 69)¹ The core idea of that merged theory is that the essence of desire as a natural kind is captured by the claim that "To be a desire is to be a representational capacity contributing to a reward or punishment signal," in the sense of the mathematical theory of learning, and "realized in human beings... by the biological reward system, centered around the dopanime-releasing neurons of the SNpc and VTA." (p. 168).

1. Method.

Despite my sympathy with the general approach, I begin with some qualms about method.

Schroeder compares the quarry of his quest—summed up in his *Precis* as "to go looking for a natural kind suitable for identifying with desire"—to the discovery that water is H₂O. But what kind of kind is in question here? Desire isn't any kind of substance, so the comparison can have only analogical force. Is it a sufficient condition in support of the answer *A* that *A* explains the causal properties of desire? If we subtract the fact that water is a substance, that seems to be the payload of the fact that water is H₂O. Or is it also required that when all is said and done we be entitled to make the claim that *desire = A* is *analytic* in the light of the scientific facts? It doesn't quite seem to work that way, because, as Schroeder admits, when we pull the rug a bit more one

¹ Unadorned page numbers are to *Three Faces of Desire* (Oxford 2004).

way, or smooth this or that wrinkle, we find some exposed facts or displaced wrinkles elsewhere. The justification for the theory is ultimately going to be subject to a holistic judgment.

These reflections raise a couple of questions:

First, why must there be a *single* answer to the question about the nature of the natural kind *desire*? Note that in the case of water, while it is often assumed that H₂O is *the* right answer to *What is water?* the appropriate answer actually depends on the purpose for which the question is asked. Heavy water is physically different from regular water; yet the familiar formula ignores isotopes. Water includes about 1 in 20million parts of D₂O: we don't usually distinguish between hydrogen ¹H and deuterium, D or ²H. Nor does the familiar formula distinguish between ¹⁶O and ¹⁸O. Is there any reason that we need a single unequivocal answer for all purposes, rather than different ones for different explanatory purposes?

My second question is grounded in the observation that the connections we are interested in are all causal. What exactly is the justification for endowing some of them with an analytic aura, while others are left as "merely" causal? One argument is found on p. 32:

A defender of a hedonic theory of desire might try to argue that desiring only requires dispositions to pleasure or displeasure under normal conditions, and when an individual is depressed, conditions are not normal. But making such a claim stick is very difficult when one's theory of desire makes tendencies to pleasure and displeasure the essence of desire, rather than a causal consequence of desiring.

This argument is designed to undermine the claim of pleasure to be essentially linked to desire. But it doesn't answer my question: why do we need anything to play that role? The passage just quoted suggests a criterion: *To count as an essence, a property must be a necessary and sufficient condition.* But why should there be any necessary and sufficient condition for anything in biology or psychology? The presumption that there must be such a thing may be both otiose and pernicious. It's what leads people to disputes about whether we are controlled by nature or

nurture, or whether there are differences of aptitude between men and women, or gays and straights. Essentialist thinking in biology retreated under the impact of Darwin, because it seemed absurd to plant strict dividing walls into smoothly and multi-dimensionally, and possibly chaotically morphing ground. That doesn't mean that in such ground there won't be attractors, homeostatically maintained, but those attractors will exist in terms of processes rather than single event types or isolated mechanisms. The connection between desire, pleasure, reward, and motivation illustrates this particularly well. Folk-psychology understands that connection as a dynamic cycle: desire motivates us to pursue a goal; successful pursuit secures the object of desire; attaining the object of our desire produces pleasure; pleasure rewards the pursuit; and the reward increases the probability that the desire will recur.

As I will stress below, I see real advantages in Schroeder's idea that we should give a privileged position to the reward system in the operation of this ancient treadmill. But I am sceptical of the presumption that we need to privilege some part of that dynamic cycle. One might, instead, view the whole as a system in which every component plays its role and may break down, but is likely to be called back to order by natural selection when it does fail.

2. Counter-examples to the classical views

Schroeder details a number of counter-examples to the motivational view. These are plausible enough, but not incontrovertible, since it's always possible to claim that while the standard position doesn't *appear* to be true in those cases, this is due to interfering factors.

Some instances of wishing, for example, (such as wishing that π were rational) is held to provide an example of desire without motivational consequences. But it might simply belong to a different category than desire altogether. Or else it might conform to the classical view, yet trivially fail ever to manifest itself in behavioural dispositions, simply because the appropriate means are always lacking. (But magic, prayer, and other forms of superstition might actually count as "trying", *modulo* a dose of irrationality. Cardinal Ratzinger, for example, just before his elevation to the Papacy, is said to have wished not to be Pope and tried to avert his own election

by praying God to let him off.)

Because the learning induced by the reward system modifies the perceptual and association capacities as well as behavioral dispositions, some learning can occur without affecting motivation. Schroeder argues that this situation is best construed as involving desire without motivation. In support he adduces, in addition to the case of wishing, dissociations of motivation and desire such as are found in such bizarre pathologies as akinetic mutism. This condition is described as follows in a British medical dictionary:

Akinetic mutism is a variety of stupor in which the patient is unable to talk or carry out purposeful behaviour but may lie with eyes open, seemingly unaware of what is going on around him. It results from bilateral damage to the orbital surface of the frontal lobes. The patient appears awake and has normal ocular movement but does not speak, is incontinent, and has minimal motor response to painful stimulation.²

Now I'm not sure what common-sense would say about these cases, but I find akinetic mutism rather unconvincing as an example. "Locked-in Syndrome," so eloquently described in the autobiographical account by Jean-Dominique Bauby³, might better serve his purposes, since in that case it is indeed the ability to move that is completely suppressed, but clearly not the capacity to feel, desire, and describe one's desires. Schroeder comments:

...it is not particularly plausible to see such people as lacking in desires. Would one treat such a person as having no interests that could be served by being cured? If I genuinely have no desire to get better, and my illness is no burden to society, a doctor will not be justified in interfering with my body without my consent.....Yet people suffering strokes

² *General Practice Notebook*, at <http://www.gpnotebook.co.uk/simplepage.cfm?ID=483393567>

³ Bauby, J.-D. 1997. *The diving-bell and the butterfly*. J. Leggatt. New York: Knopf.

who are thereby rendered akinetic and mute seem to be reasonably treated without consent. (p. 173).

But this strikes me as beside the point. It confuses desires with *interests*. Whatever its actual incidence, it is not difficult to imagine an incapacity to form the kind of representations that could serve to stimulate a reward signal. Yet we would be equally reluctant to treat someone suffering from this condition as lacking interests: interests are to some extent, objective in a way that doesn't correlate strictly with desires. Furthermore, interests have nothing to do with capacities for representations that can trigger reward signals.

Nevertheless, Schroeder has built a solid case for the possibility of a disconnect between motivation and reward, between motivation and pleasure, and between pleasure and reward. But might there also be counter-examples to Schroeder's own alternative theory? If the link between desire and reward is analytic, then presumably there could not be any such counterexample. But that move, as Russell once remarked, has all the advantages of theft over honest toil. If there is evidence against viewing that link as tight enough to warrant a decree of analyticity, then putative counter-examples turn into mere exceptions. In any case, in science all analyticity is *pro tem*.

Rather than pursuing the counter-example game, then, let me invoke a couple of thought experiments pertaining to the cycle of interrelated mechanisms working together. First, could we conceive of an organism that has a reward system, but no motivating function? Schroeder rightly points out that the reward function doesn't merely modify the probability of future behaviour (which one might regard as functionally equivalent if not phenomenologically identical to motivation.) It affects salient perceptions, inferences, and emotions as well as the purely physiological signs typically attendant upon emotions. But without modifications of motivation, there would be no useful learning, regardless of the remaining salience of pleasure. So a reward system functioning without influence on motivation would be severely pathological.

Second, can we conceive of an organism that has the reward system but no hedonic

function? It is difficult to conceive of pleasure that isn't conscious, but it's certainly possible (indeed demonstrably true) that learning, activated by the reward system, can take place without any conscious experience. Furthermore, we assume—if only for the comfort of our conscience as killers and consumers of other living organisms—that pleasure and pain only make an appearance in our relatively closer cousins on the phylogenetic scale, mammals, perhaps, or at least (since nociceptors have recently been found in fish)⁴ in vertebrates. But it is implausible to suppose that pleasure as such has no direct causal role to play in influencing our behaviour.

Could conscious pleasure act on future motivation without the intervention of the reward system? Surely this seems *possible*, though one would have to concede that an organism so designed (or so deprived) would not be capable of the sort of responsiveness to changing circumstances that are crucial to our capacity to adapt in everyday life. It would be an unwieldy life indeed that was lived entirely on the basis of conscious memory of pleasures and satisfactions, grounding the elaboration of conscious intentions to bring about similar experiences in the future.

These considerations seem to allow for the possibility that all three of the "faces of desire" are actually "essential", in the sense that the disconnection of any one of the links in the cycle would severely impair the workings of the whole. On the other hand, each could be lacking in some respects on some occasions, resulting in pathological conditions, perhaps, but not definitively crippling. So in a stronger sense of 'essential' *none* of the three would actually be essential.

⁴ See Sneddon, L. U., V. A. Braithwaite, and M. J. Gentle. 2003. Do fish have nociceptors: evidence for the evolution of a vertebrate sensory system. *Proceedings of the Royal Society London B* 270(1520):1115-21.

3. Some pleasant consequences of Schroeder's view.

Let us return to the dynamic picture of the cyclic relation between desire, pleasure, reinforcement and motivation. Here, once again, is how it is generally assumed to work:

- desire motivates us to pursue a goal
- pursuit secures the object of desire
- the object of desire produces pleasure
- pleasure rewards the successful pursuit
- the reward increases the probability that the desire will recur.

When all goes normally, each part of the cycle plays its own causal role in the perpetuation of the cycle, like the Lion King's Circle of Life: the antelopes eat the grass that feeds on the soil processed by the worms that eat the lions that eat the antelopes. It gets interesting only when the cycle is broken. What happens then?

What Schroeder has shown is that this picture conflicts both with phenomenology and neuroscience. Of the "three faces of desire", only the reward system is linked *essentially* to desire. Hence the other links in the chain are looser: there can be motivation without desire, and successful pursuit without pleasure.

I want now to ask how this proposal might throw light on four phenomena—or more exactly two phenomena and two somewhat eccentric but plausible principles of rationality—which I believe to be of intrinsic interest. The two phenomena are: (a) vice and (b) Platonic advertising; the two principles or rationality are (c) the Philebus Principle, and (d) Aspectual Adequacy.

(a) Vice

The word is often used to designate some habit of which one vaguely disapproves, or as a contrary, in some rather obscure sense, of "virtue". (Obscure, in part, because as Aristotle noted most virtues represent a happy medium between opposite excesses, so in that sense the opposite

of a vice ought to be not virtue but some other vice.) I propose instead that we define vice as *something you can no longer stop choosing to do although it no longer brings any pleasure*. On the standard view, vice in this sense ought to be impossible; and yet notoriously it not infrequently plagues our lives. Against the background of Schroeder's view, that sorry fact is much less puzzling. For since the reward system, and hence desire, is only contingently linked to the production of pleasure, it can retain its link to motivation (as well as to other modification of mental habits) without generating any pleasure, or for that matter without being negatively affected by the absence of pleasure.

The confirmation of the way this works rests in part on a recent triumph of science: the production in the laboratory of vice in rats. Kent Berridge and Elliot Valenstein manipulated rats' brains in such a way as to produce a desire and motivation to eat, but were able to ascertain by the rats' physiological and expressive behaviour that while they wanted to eat, they did so without enjoying it: the very paradigm of vice.⁵

Since Schroeder rejects the "hedonic view", for which pleasure is criterial of the genuine presence of pre-existing desire, he is undisturbed by vice, as by any other counter-examples to that view. Such counterexamples arise wherever there is a dissociation between pleasure and desire. A first such dissociation was noted by Plato, who took a pessimistic view of the virtues of pleasure and its relation to desire. He likened desire to a leaky vessel in perpetual need of replenishment. Replenishment is the pleasure, but the very condition of its existence is the pain of deprivation which is desire. Yet Plato noted an exception in pure pleasures of smell. You might be walking along without any pain of desire, when your nostrils are blessed with the perfume wafting, unexpected, from a neighbouring rose garden. (*Philebus* 51b). Now that's a

⁵ Berridge, Kent C. and Elliot S. Valenstein, "What psychological process mediates feeding evoked by electrical-stimulation of the lateral hypothalamus", *Behavioral Neuroscience* 5/1 (1991): 3-14.

good counter-example against the universal applicability of Plato's own leaky vessel view, but it's not necessarily an objection to a modern "hedonic theorist of desire." For a hedonist needn't insist that all desires are conscious, and can plausibly enough claim that sweet smells and other pleasant sensations are no less intrinsically desired for being unsought. Schroeder, indeed, asserts just that. So it seems the existence of vice is actually equally friendly to the reward theory and to the hedonic theory.

The harder cases of dissociations go the other way: not pleasure without desire, but desire satisfied that fails to yield the expected pleasure. That can happen in several different ways, all of which should act as reminders of the fact that the *semantic* satisfaction of desire—the mere fact that desire for *p* is followed by the realization that *p* has actually come to be—does not guarantee *emotional* satisfaction. Emotional satisfaction is none other than the pleasure represented in the experience of desire. One way this can occur is because the representation that went with my desire was insufficiently specific. When I said I wanted a strawberry, I didn't mean that kind of odourless, flavourless, chewy monster. But more careful framing of the propositions constituting the object of desire won't necessarily guarantee emotional satisfaction. For when I get the object of my desire I never get *just that*, and some of the attendant circumstances can be such as to annul the expected pleasure altogether. This is the "monkey's paw" phenomenon.⁶ Less dramatic but perhaps no less baneful are the cases where the neurotransmitters just don't do their job at the appropriate time. The most obvious manifestation of this to common-sense is the "dust and ashes" phenomenon: the reason the semantic satisfaction of a desire will not guarantee emotional satisfaction, which on the standard view is the experience of pleasure in the attainment of desire, is that the relation between semantic satisfaction and emotional satisfaction is a

⁶ From the story by W.W. Jacobs in which a monkey's paw has the power to make wishes come true, but only in unimagined and horrible ways. You can read this story on line at <http://gaslight.mtroyal.ca/mnkyspaw.htm>.

contingent, causal one. On Schroeder's view, the only thing that could establish that I didn't really desire something would be that some adequate representation of it altogether failed to produce any reward signal.

One *could*, however, insist that my failure to enjoy what I said I wanted showed that I didn't really want it.

(b) Platonic advertising.

That is actually Plato's strategy, when he tells us, *passim*, that what we really want is always the Good. In a sense, then, our ordinary desires, such as that I might feel for a bright and beautiful boy, are always actually for something *else*. Although this seems to contradict the Socratic doctrine that all wrong choices are the result of a failure of knowledge not desire, it is actually better seen as a corollary of that doctrine. All of us really just desire the Good. What happens when we choose wrongly is that we have misidentified it. But now if our desires are grounded in our imagination of pleasure, we can blame that imaginative representation for the misidentification, and that's getting close to Schroeder's theory that the *representation of pleasure* has failed to match up with *the representation of what would trigger a reward signal*. That's what actually makes it a mistake, and on Schroeder's view it is also the case that the mistake is not in the desire itself (which can't misidentify the reward signal by definition) but in the representation of pleasure.

The way Schroeder illustrates this in neurological terms is a nice example of the way that neurology can illuminate a philosophical point. The brain, he suggests, "normally stimulated... by structures carrying information about whether unexpectedly rewarding states of affairs ... are being represented at present" is actually "hijacked" by "directly infused chemical agents", which "naturally can be expected to lead to misrepresentation"—"hence the unreality of the pleasures of euphorigenic drugs" (pp. 92-3.)

Transposed into a somewhat more Freudian vein, this idea is at the heart of an advertising strategy which, judging by its frequent implementation, must at least pass among professionals

for being highly successful. That consists in giving you the impression that buying a Buick will get you sex. If the Buick were what you really wanted, you would stop there and run it into the ground. But it isn't, since sex—the post-Freudian equivalent of The Form of the Good—is what you really wanted. By association with sex, the Buick triggered the reward system before you got it, but the pleasure it procured didn't live up to the expectation. When getting the Buick fails to please (as much as it should), the common-sense view would predict that reinforcement will fail, extinguishing any desire for another Buick. Yet it doesn't, providing the advertising campaign has cemented the association sufficiently firmly. Your failure to get satisfaction will merely be interpreted as signifying that you need another Buick. So you buy another Buick.

That all seems to fit in rather well with Schroeder's scheme. (On the other hand, I'm not entirely sure that advertising works: it may merely be superstition on the part of advertisers who persist in tying Buicks to sex.)

(c) The Philebus Principle

Plato argued, in the *Philebus*, that pleasure can be false not merely by association with false beliefs but in its own right. One of the ways he explicates this pertains to *pleasures of anticipation*. These pleasures can be thought of as second order representations, distinct from beliefs. A pleasure of anticipation represents the anticipated pleasure. And if we think in terms of the role such pleasures might play in the planning of an organized life, it's clear that the pleasure of anticipation, insofar as it has any influence on our decisions (which the RTD doesn't deny in general), will work to our long term advantage if it is a reliable indication of the future satisfaction. The space allowed between the pleasure of anticipation and the anticipated pleasure to which Schroeder has given a specific neurological interpretation in the case of drug addiction, is also a space in which one can insert a *principle of rationality* peculiar to the relation of pleasures of anticipation, anticipated pleasures, and desire. The principle says nothing about the actual extent to which one should experience the anticipation of pleasure as pleasurable, but prescribes, more weakly, a relation of proportionality between the two pleasures:

(PP) Pleasure of anticipation should be *proportional* to the anticipated pleasure.

Schroeder doesn't say anything about any principle governing the rationality of emotion and pleasure that might relate to, let alone be derivable from, his characterization of pleasure. But I'm keen to know whether he would be sympathetic to the idea that once we reconfigure, as he proposes, the relation of motivation, pleasure, and desire, we might also find ways in which we might reconfigure, or indeed discover afresh, principles of *rationality* that pertain specifically to emotion and desire.

(d) *The Principle of Aspectual Adequacy.*

In that spirit, I'd like to ask him whether he can fit into his scheme another principle of *sui generis emotional* rationality that may seem somewhat arcane, but which I hold to be both sound and important.

To explain this principle, I must start with the grammatical concept of an *aspect*, as it applies to verbal forms. Aspects are easy to confuse with tenses or moods, but unlike moods they do not correspond to modal distinctions (such as the indicative, optative, subjunctive, and imperative moods, all of which have to do with possibility, actuality, desirability, or the counterfactual nature of some proposition envisaged). Unlike tenses, an aspect does not carry information about whether an event is past, present, or future. Instead, it has to do with the way the event is envisaged as taking place in time in respect of its duration. If you learned classical Greek, you will have learned about the Aorist, which often relates to the past, but can equally well relate to the present. The aorist is punctual: it envisages an event as taking place at some particular *point in time*. By contrast, the imperfect envisages it as taking place *over a stretch of time*, and the perfect envisages it as having been, now or at some other point of time referred to, *completed*. In English, it is seldom noticed that the present tense is actually not a tense at all, but connotes a *frequentative aspect*. Thus if you ask someone "What do you do?" you are not interested in finding out what they are engaged in at the present moment. (The continuous present, "What are you doing?" would be used to ask that question.) Instead, "What do you do?" is

after something like a description of the interlocutor's job, or what they *usually* do.

That should suffice to convey the idea of aspect. To this grammatical concept, Aristotle pointed out that there correspond different sorts of human behaviour, or at least different ways of envisaging human behaviour. Aristotle distinguished, in particular, between *kineseis* and *energeiai*. (*Met.* 1048b 18-35). The former are processes, aiming for a natural end that gives a process its *point*, achieved only as the process ceases. Such are running a race, or writing an exam. *Activities*, by contrast, have their point as they unfold. They have no specific natural end, and carry on for as long as the agent feels like it rather than until they have achieved a particular purpose. Such are going out walking, or contemplating a landscape, or having a conversation.

Now among desired human activities, just as some can be envisaged in memory or imagination under one or the other of the perfect, continuous, or punctual aspects, so they can also be *desired* under one or the other aspect. Sometimes, an activity is desired under the punctual or perfect aspect, despite being more intrinsically suited to the continuous. The male orgasm, for example, is a natural end of sex, and the focus on male orgasm lends to sexual activity the air of a process, completion of which constitutes an achievement. But it may be more suitable to the intrinsic worth of sexual relations to be regarded as activity rather than achievement, and so it may be irrational to desire sexual achievement as opposed to sexual activity. Or so, at least it is enjoined by the *Principle of Aspectual Adequacy (PPA)*:

(PAA) An object of desire should be desired in the temporal aspect most suited to its nature as a source of pleasure or satisfaction.

The contemplation of art, the reading of poetry, and the enjoyment of music may be other instances of things which are often approached in an intrinsically inappropriate way.

Can the reward theory of desire make sense of the idea of aspects? Could we discover, in some further refinement of the way the brain processes representations and gauges their influence on learning, on motivation, and on pleasure, a difference that will reflect aspectual

difference? It doesn't seem altogether impossible. But I have no clear answer, so I'm keen to take advantage of the opportunity to ask whether he can throw any light on it.

Epiphenomenal Pleasure

I began these comments with some skeptical reflections on the quest for an essence in the context of biology. It seemed to me a curiously unbiological preoccupation to import into an investigation the most notable virtue of which was that it sought to get away from the standard obsessions of analytic philosophy, in favour of a more integrated, scientifically informed approach. I end with a puzzle that itself expresses a worry more characteristic of philosophy than science, but that would not really arise were it not for the biological point of view that I take Schroeder and myself to share. In the course of arguing against the hedonic theory of desire, Schroeder observes that it is "something of a surprise to find that neuroscience has documented very few strong links between pleasure and motivation." (p. 121). Given that fact, the hedonic theory of desire is "forced to give desire a much more diminished role in its account of action production. This, I think, is a serious blow against the hedonic theory." (p. 127). But here is the worry. Isn't this, more seriously, a blow to the idea that pleasure has any function *at all*? And shouldn't this worry someone who thinks of the enterprise of philosophy of mind as one with biology?

Actually, this problem is part of a more general problem about the function of consciousness. Other neurological findings, such as those of Benjamin Libet, have cast doubt on the role of conscious states in the origination of action, since physical movements are apparently initiated in the brain some 200ms before the consciousness of an intention appears.⁷ This threatens to make conscious intention, no less than pleasure, an epiphenomenal consequence rather than a cause of action. Some philosophers believe that philosophical zombies are logically

⁷ Libet, B. 1999. Do we have free will? *Journal of Consciousness Studies* 6:47-57.

possible, capable of passing the most exigent form of Turing test and yet wholly devoid of any states of consciousness. If so, then consciousness is unlikely to have played any role in differentiating those individuals privileged from those eliminated by natural selection.

For anyone who finds it incredible that consciousness might be strictly functionless, that is a problem. The non-efficacy of conscious pleasure is a special case of that general problem. Perhaps pleasure affects behaviour indirectly, insofar as it serves to influence the reward system by affording, in memory, a basis for the imagination of future pleasures. That would mean that a zombie that could learn and modify its behaviour wouldn't be doing so in the way envisaged by those who endorse the usual thought experiment of philosophical zombies. For in the usual conception, a zombie would be in every respect identical with humans except for the existence of conscious states. As envisaged here, on the contrary, there would be an additional difference, not reducible to the trait of consciousness, in the functioning of their learning mechanisms. If rewards, not pleasure, activate learning and cause normal motivation, then what is the point of pleasure, and how could the capacity to experience it have contributed sufficiently to the survival of our ancestors for it to have been selected to fixation?

Again it seems to me plausible to think this is a false problem. If the fundamental biological phenomenon we are dealing with is the cycle of desire, motivation, action, pleasure and reward described above, then all the elements of this cyclical process are equally "essential", even if some of the links involved are weaker than others. But that involves giving up the core idea that desire is a natural kind privileged by its essential ties to the reward system.

Even so, by showing that the bonds that tie desire to the reward system are tighter, and those that tie it to pleasure or motivation weaker, than common sense presumed, Schroeder has effected a brilliant reconfiguration of our understanding of desire. It will never look quite the same again.