

## Chapter 2

### Anti-Reductionism

---

There is an old bit of advice which says: Watch your friends; your enemies will take care of themselves. In the scientific *métier*, this saying goes: Suspect the obvious; the obscure truths will elude you anyway.

— Kline, *Mathematics in Western Culture*

The sting is only removed from a system of thought when the particular conditions under which it makes sense are described.

— Bhaskar, *A Realist Theory of Science*

Over the last three hundred years the scientific materialist understanding of the world has been moulded by two pictures, two *intuitions*, and I want to convince you that a third is at least possible. The first intuition is *reductionism*. Societies are made of people, people are made of cells, cells of molecules, and molecules of atoms. These parts come together to form wholes, and the behaviour of the wholes are determined by the parts. The logical conclusion of reductionism is that all science — and, in extreme cases, all art, ethics, and politics — is just ‘physics plus abbreviations’ as the logical positivists put it.

Reductionism has been one the most powerful ideas of the modern age, but many are reluctant to reach its seemingly brutal conclusions. ‘Indeed, it seems to be a little-known law governing the behaviour of contemporary philosophers that whenever they profess faith in any form of materialism or physicalism they must make it absolutely clear that they are, of course, in no way endorsing anything as unsophisticated, reactionary, and generally intolerable as reductionism’ (Melnyk, 1995).

The alternative intuition is less of a militantly cohesive picture than reductionism, and more of an understandable reaction to it. Let us call it *pluralism*, though the same intuition goes under different names in different contexts. The central point is that objects — including people, and their art ethics, and politics — must be understood on their own terms. As Fodor, in a review of E.O.Wilson’s reductionist manifesto *Consilience: The Unity of Knowledge*, puts it

everything is physical perhaps, but surely there are many kinds of physical things. Some are protons; some are constellations; some are trees or cats; and some are butchers, bakers or candlesticks. For each kind of thing, there are the proprietary

generalisations by which it is subsumed, and in terms of which its behaviour is to be explained. For each such generalisation there is the proprietary vocabulary that is required in order for our discourse to express it. Nothing can happen except what the laws of physics permit, of course; but much goes on that the laws of physics do not talk about.

It is important to realise that there is no necessary contradiction between pluralism and reductionism, and most modern philosophies combine elements of both. For example, many scientists and philosophers are increasingly unhappy with reductionist claim that the higher levels are *determined* by the lower, and one attractive alternative is what we may call *pragmatic anti-reductionism*. This argues that although reductionism may be correct in principle, it can rarely be used in practice: it is simply not feasible to collect all the data, and perform the calculations necessary, for all but the most trivial systems. According to pragmatic anti-reductionism, properties of wholes may be determined by those of the parts, but this does not imply that they are necessarily derivable from them. If you want to understand the world then the only possible strategy is to investigate each phenomenon on its own terms rather than start from the physics. In my experience most practising scientists would agree with some form of pragmatic anti-reductionism.

Basic pragmatic anti-reductionism can be strengthened in various ways. For example, we can borrow from chaos theory and argue that aggregate properties of the system may be sensitive to some properties of a part, such as the infamous sensitivity of weather systems to a butterfly's wing. If this is the case then accurate predictions about the higher level depend on knowing the properties of the parts with unbounded accuracy, and there are various reasons, such as the Uncertainty Principle, why this is not possible.

A pragmatic anti-reductionist can also argue that just knowing the properties of the parts is not enough to derive higher level properties; we also have to know the composition of the higher level entities that we are interested in. Thus although the set of valid higher level descriptions may be *determined* by the lower level properties, they cannot be *discovered* or *derived* without additional knowledge. Thus we find that, with a few exceptions in astrophysics, there are virtually no cases in the history of science in which a higher level scientific law or description has been derived from a lower one; rather such phenomena are discovered by investigation at the appropriate level and only subsequently related to lower level properties.

For a non-realist — for whom properties *only* exist within, or with respect to, our knowledge — these pragmatic objections to reductionism are also ontological ones. Thus if objects and properties must be discovered at their appropriate level (rather than the higher being derived from the lower) then they will have the same epistemological status; and this means that there will also be an ontological symmetry between levels of organisation. Indeed it seems plausible that it is this ability to avoid reductionism that attracts many materialists to some form of pragmatism or instrumentalism in the first place: instrumentalism allows one to be a materialist without implying reductionism.

(It is sometimes said that an abstract philosopher, in contrast to a hard-headed scientist, is one who will complain that 'it may work in practice, but does it work in principle?' We *know* that anti-reductionism is a vital strategy in scientific practice, so why should we care whether it works in principle? The problem is that there is no Chinese Wall between practice and principle. A scientist strives to get their truth-in-practice as close to the truth-in-principle as possible; their assumptions

of principle guide empirical work. Therefore if we want to avoid a reductionist scientific practice then we need anti-reductionist principles.)

Another way of avoiding the implications of reductionism is to express the relationship between levels of organisation in terms of *supervenience* (Kim, 1984). The idea is that a property of a whole, *P*, is supervenient on some properties of the parts, *p*, iff there can be no change in *P* without a change in *p*; or if when two entities are indiscernible with respect to *p* they are indiscernible with respect to *P*<sup>1</sup>. Thus the concept of supervenience can be used to describe the relationship between levels of organisation without mentioning the reductionist bogey-word ‘determination’.

The notion of *emergence* has also been used to do a similar job to that of supervenience. The idea is that higher levels of organisation ‘emerge’ out of the lower, rather than being determined by it. But the problem is then to define emergence in a way that does not involve determination. One way is a kind of mystical holism in which wholes are blessed with properties that are not dependent on parts. But the more common way is a kind of pragmatic reductionism in which properties of wholes are considered emergent if they are in some way novel or surprising (Nagel, 1961, p374-80)(Crutchfield, 1994). Hydrogen and oxygen gas are not wet, for example. Indeed there is nothing about them which even *suggests* wetness. But if you put them together and spark a chemical reaction then you get water, which quite clearly is. Nonetheless it still seems the case that the properties of water are determined by those of its constituent molecules, even if we have problems deriving them.

Another way of combining reductionism and pluralism is to hold that each doctrine is true of its own separate domain. The usual split is that reductionism holds for biology downwards, whereas pluralism applies to humans and their cultures ‘from the neck up’, as it were. Thus we find that many philosophers are relatively uncritical of reductionism in natural science whilst strenuously denying that the same methodology can be applied to human affairs<sup>2</sup>. Nonetheless it is hard to avoid the blunt fact that our human experience is in some way linked to our ‘lower’ properties. If you push me, for example, then, as a physical object, I will fall. If I fall then, as a biological object, I will be injured. If I am injured then, as a psychological object, I will be in pain. And if I am in pain then, as a social object, I will sue. Thus we have many ways of being, and all these ways of being — these levels of organisation — are linked.

There is evidently some connection between levels of organisation, but what? The obvious answer is the reductionist one, that higher properties are dependent on lower ones. But in the rest of this chapter I argue that, although reductionism may be true in one sense, it only gives us half the picture. Reductionism is true in the sense that the properties of an object are indeed dependent on those of its parts; but it is *also* true that the properties of parts are dependent on wholes. Consider this example. In the nineteenth century Britain twice went to war with Manchu China in order to free up her trade in opium. So the foreign policy of the British Empire resulted in an increase in the concentration of opiates in the brains of millions of Chinese peasants. If a Chinese neurologist wanted to know why there was such a high concentration of endorphins in the

<sup>1</sup>These two versions of supervenience are not strictly equivalent, but the differences between them are not important for this discussion.

<sup>2</sup>The same combination of intuitions recurs in Margaret Thatcher’s infamous claim that ‘there is no such thing as society, just individuals and their families’. After all she did not claim that there are just *atoms* and their interactions, as a more consistent reductionist would. Individuals are the only things that exist for Thatcher for the bluntly pluralist reason that they are the actors in the political discourse that she was concerned with.

synapses of the brain cells she studied, then part of the answer would point to British Imperialism. Moreover, one of the effects of this mass addiction was to increase Chinese support for the Taiping rebellion which called for the prohibition of opium and the expulsion of the foreign powers that supported the trade. So in order to understand this piece of history we must trace the connections from the social level, down to biology, and back up to the social again; and it is precisely this kind of analysis that both reductionism and pluralism rule out. The pluralist would argue that you can understand the ebb and flow of historical tides without bothering with biology. The reductionist would disagree, but would argue that the biologist should stick to biochemistry, because politics cannot tell us anything about brains.

Or consider this other example. We are now all familiar with the increasing power and sophistication of psychoactive drugs that are able to relieve the symptoms of various conditions, such as the effects of SSRIs on depression. The reductionist interpretation of this success is that the psychological depression is *caused* by a neurochemical imbalance which the drugs correct. Many pluralists are unhappy with this interpretation and prefer to emphasise the psychological and/or social causes of the condition, but seem to believe that this requires they deny that the drugs have any beneficial effects at all. But there is no contradiction between the two explanations. It is perfectly possible for social pressures to have effects on the neurochemistry of our brains, with depression being the result. Drugs can break one link in this chain — and the relief can be welcome — but this does not imply that the condition was ultimately biological.

We are not determined by our biology, as the more ‘greedy’ reductionist would argue. But nor are we independent of it, as the more ‘idealistic’ pluralist would argue. What we do as humans depends on our biology; but it is equally true that what we do as humans *affects* our biology. The rest of this chapter is an attempt to outline an alternative way of understanding the relationship between levels of organisation that can accommodate this simple intuition.

## 2.1 Reductionism and Materialism

The basic premise of reductionism is that the world is made of objects, each of which has properties. These objects come together to form larger objects, and the properties of these wholes are dependent on the properties of their parts. We may then argue whether or not descriptions involving those larger objects are eliminable, or whether the properties of the wholes are strictly derivable from those of the parts, but the basic logic seems to be an irrefutable and inevitable consequence of materialism (Melnyk, 1995)— hence the suspicion in some quarters that anyone who espouses any form of anti-reductionist holism must be some kind of ‘flaky’ anti-materialist. Let me put the same point another way: how, precisely, can the whole be greater than the sum of the parts? If one is to remain a consistent materialist then it is a bit of the problem to explain where the extra comes from.<sup>3</sup>

Reductionism seems like one of the most obvious and basic truths in science. But it is precisely

---

<sup>3</sup>“The whole is greater than the sum of the parts” is a useful way of summing up the basic intuition of anti-reductionism, but it is strictly inaccurate. The idea that the whole is greater than a linear sum of its parts is perfectly compatible with even the strictest reductionism. The gravitational force between two masses, for example, is equal to the *product* of the parts, and yet this example is a triumph of reductionist analysis rather than being any kind of threat to it. A more accurate way of expressing the problem is to ask how the whole can ever be more than a *function* of the parts.

because it seems so obvious and basic that I want to put it under suspicion. The aim is to show how it is a peculiarly one-sided way of looking at the world, and to suggest another perspective in an attempt to redress the balance. Unfortunately the reductionist intuition runs so deep that it is difficult to know where to start challenging it. It is tempting to start from metaphysical first principles — this is, after all, the philosopher's natural strategy — but I have found that presenting the argument in this way rarely convinces. It seems that the intuition is just too deep. The alternative strategy that I pursue here is much more pragmatic. I consider three simple, familiar, examples of systems made of many interacting parts, and show how the anti-reductionist perspective can do useful work in making sense of aspects of these systems that the reductionist perspective neglects. I hope that the very mundane familiarity of these examples will convince where metaphysical generalities would not: if the anti-reductionist perspective can do some useful work on such well-worn examples then perhaps it should be given a chance on the more obscure ones discussed later in the thesis?

The first example is the Boyle-Charles Gas Law, which states that temperature of a closed container of gas is inversely proportional to its temperature. In this case we have an object (the container) that has properties (temperature and pressure) that behave in a particular way (they vary inversely). Maxwell and Boltzmann, in a triumph of reductionist analysis, proved how the behaviour of the gas could be explained by the motion of the individual molecules that make it up: each of these molecules collide elastically with each other and the container walls, and as we heat the gas the velocity of the molecules increases and they exert an increased force on the container walls. Thus the temperature and pressure of a gas are determined by the motion of its molecules, and the gas laws governing the properties of the whole container are determined by the laws governing the behaviour of its parts.

The second example is Conway's Game of Life. Suppose we have a large grid of square cells, each of which can be 'on' or 'off'. The state of each cell at the next step in the life-cycle of the grid is determined by simple rules defined over the current state of the cell and those of its neighbours. Out of these simple rules emerge a rich 'eco-system' of higher order patterns that may glide across the grid, blink between two states, generate gliders, and so on<sup>4</sup>. The Game of Life is often used as an illustration of how systems of interacting parts obeying simple rules can produce novel and interesting behaviour. But nonetheless it is still the case that the appearance and behaviour of the objects in the system (i.e. the higher order patterns) is determined by the arrangement and properties of their parts. Gliders glide and blinkers blink because of the rules governing the cells. The Game of Life is an example of how the reductionist approach can make sense of the emergence of complexity, not a challenge to it (Faith, 1998).

The third example is a car engine. A car engine is made of many different parts — driveshafts, pistons, cam-belts, and so on — each of which are carefully engineered to have very precise properties. None of these parts produce any power on their own, but when they are put together in the right way then we have a complete engine that does. Power is thus a property of the whole object that is dependent on the properties of the various parts.

It should be noted at this point that none of these three examples are biological or social. I do not, for example, consider how thought processes can emerge out of the interacting neurons in

---

<sup>4</sup>For more on the ontological and epistemological status of these patterns see Dennett (1991).

our brains, or how social systems can arise out of the interactions of free agents. This omission is deliberate for I want to break decisively with the intuition, mentioned above, that reductionism may be true of the lower sciences, but not for us higher, sophisticated, biological beings. This intuition only serves to enforce the gap between natural science and the philosophy of mind that it is my purpose to break down. I hope to show, by considering such ‘mechanical’ examples, that reductionism is not enough to understand examples from what is usually taken to be its strongest ground, and hence that its extension to other areas should be viewed with suspicion.

## 2.2 Anti-Reductive Materialism

So, how can an anti-reductionist perspective help us understand such simple systems? Consider the derivation of the gas laws from the kinetic theory of molecular collisions. Why do these derivations work so well? Feynman argues that

we shall find that we can derive all kinds of things — marvellous things — from the kinetic theory, and it is most interesting that we can apparently get so much from so little. . . . How do we get so much out? The answer is that we have been perpetually making a certain important assumption, which is that if a system is in thermal equilibrium at some temperature, it will also be in thermal equilibrium with *anything else* at the same temperature. (1963, p40-1)

So what happens if the gas is *not* in equilibrium?<sup>5</sup> The easiest way to find out is to compress it. As soon as we push on the walls of the container the measured pressure will rise, and as we continue to push we do work in compressing the gas. This energy diffuses through the container, raising the mean molecular momentum per unit volume, and those molecules nearer the compressed surface will be affected before those further away. Thus the properties of the parts are affected by what happens to the whole. The constituent molecules have the momentum that they do *because* of the pressure on the cylinder. The dependency only appears to run the other way when the system is static or in thermal equilibrium. Or suppose that we cool the container until the gas reaches its dew point where the molecules stop rebounding and start to stick together as the gas condenses into a liquid. Thus the molecules only collide elastically *because* they are in a gas at a certain temperature. Changes to the whole can affect the rules governing the behaviour of the parts.

The molecules of the gas are causally affected by what happens to the whole container, but there is another way in which the properties of parts of a system are dependent on the whole. This is *conceptual* dependence. Suppose, for example, we wanted to know the power of the car engine, and in order to measure this property we connected a measuring device to the main drive shaft. Now the drive shaft is clearly a *part* of the engine, and yet the power we measure is described as a property of the *whole*. We would not usually say that the power of the *shaft* was *X* Watts, but that the power of the *engine*, at the shaft, was *X*. Power is a property *of* the whole engine, but is located *at* one of the parts. On the other hand, when we measure the power at, say, the camshaft we would not normally describe this as ‘the power of the engine measured at the camshaft’. Why the difference in the two cases? What makes one a property of the whole and the other a property

---

<sup>5</sup>Non-equilibrium systems have been largely neglected in physics, with the notable exception of the work of Prigogine (1962).

of a part? The reason depends on the relationship between the engine and the rest of the car — i.e. the larger object of which the engine is itself a part. Now the output of the engine is connected to the rest of the car *via* the drive shaft, so the properties of the cam shaft and piston heads and all the other parts do not effect the rest of the car directly, but only through that single output. The power of the drive shaft is dependent on the properties of the other parts, not *vice versa*. In other words something is described as being a property ‘of the whole’ *because* it is dependent on other parts. Therefore the dependence of wholes on parts is built into the way we define ‘property of whole’ and ‘property of part’; it is a conceptual assumption that we make, not an empirical result about the way systems actually work.

The distinction between ‘higher’ and ‘lower’ properties in such cases is purely epistemic. It is an artefact of how we view the system. The speed of the camshaft, or the temperature inside the cylinders, is no less a property of the whole system than the power of the drive shaft. Each of these properties is at the same level. The distinction is rather between properties whose direct effects are felt outside the system and those whose effects are internal. We designate the former as higher properties simply because we cannot see under the skin of the system, when in strictly ontological terms there is nothing to choose between them.<sup>6</sup>

What about our third example, the Game of Life? The problem with the usual reductionist picture is that it treats the Game as a formal system. However all actual Games of Life — as opposed to the Platonic Ideal of the formal definition of the game — exist on computers<sup>7</sup>. And all computers — as opposed to the Platonic Ideal of formally-defined Turing Machines — exist in a physical and social context. They have power supplies, human users and programmers, cooling fans, manufacturers, and so on. This context forms a larger system of which the computer, along with the Game of Life that runs on it, is just a part. Moreover this context can causally effect the running of the game: the power supply may fail and interrupt it; the user may get bored and switch it off; the programmer may start to hack at the code; or the manufacturer may force an upgrade of the operating system which renders the old code obsolete. Therefore the behaviour of the gliders and blinkers in the Game are *not* determined solely by the rules governing the parts, but are also dependent on the physical context in which the Game runs.

However it may still be argued that in all cases in which the Game is running then the supervenience of the higher patterns on the cellular rules is maintained; i.e. as long as the Game is running ‘normally’ then the properties of the gliders and blinkers are determined by the rules governing the cells, even though the Game as a whole is dependent on what happens in its environment. But this is not quite true. Suppose we ask why the cells obey the rules that they do? The simple answer is that the computer was programmed in a certain way such that the cells obey the rules defined by Conway. But why was the computer configured in that way rather than any other? Now the rules of the Game were not revealed to Conway from on high, but were the result of experimentation; he tried many different versions until he found a set of rules that generated interesting behaviour. This process of experimentation is still going on. Most copies of the Game available on the Web, for example, allow the user to play with the rules themselves, and so there are many different versions

---

<sup>6</sup>This point will become important when we consider the distinction between theoretical and observational terms in general (section 3.3), and the status of mental representations in particular (5.2).

<sup>7</sup>It is also possible to run the game using pen and paper though this is time-consuming and tends to rob the game of its interest. The same arguments apply in either case, so I will just discuss the computer-based form.

of ‘the’ Game in existence. Thus, in all existing instantiations of the Game, the rules governing the interactions of the cells have the form that they do *because* they generate interesting higher patterns, as much as *vice versa*. If the rules did not generate interesting patterns then they would be changed.

(Of course within the space of all possible cellular automata there exists one, call it *L*, which has the same rules as Conway’s Game; and the rules of *L* are prior to, and not dependent on, the behaviour they generate. But unless we invoke the axiom of choice then *L* is picked out as The Game *because* of the higher behaviour. Therefore, even within the space of possible CAs, the property of ‘being the rules of the Game of Life’ is not prior to the property of ‘being the emergent patterns of the Game of Life’. The rules may be defined independently of their emergent behaviour but this does not imply that they are ontologically prior.)

Systems like the Game of Life certainly exhibit rich and fascinating behaviour. And in some cases such systems can be successfully used to model natural biological phenomena — such as in the work of Thom, Waddington, Kauffman, and Goodwin, and in Turing’s diffusion-reaction model of morphogenesis. But we should be careful about deriving general philosophical conclusions about the relationship between levels of organisation in nature from such artificial systems. These models embody certain assumptions about how physical systems work. In particular, they assume that there is a set of prior lower level entities whose behaviour is determined by fixed laws. Therefore, when we find that their higher level behaviour is only non-reducible in a weak, pragmatic, sense we should not assume that this is a correct understanding of emergent phenomena in nature. Dennett once noted that, for philosophers, the attraction of experiments such as the Game of Life is that one gets to make up the facts. But we should be aware of the cost of such factual liberalism.

Where did the reductionist go wrong? Where is the flaw in their argument about the inevitability of reductionism? The problem was that the reductionist starts by considering the properties of objects in isolation, and then asks what happens when those objects come together to form wholes. The reductionist metaphysical intuition is that objects are in a strong sense *independent* of their environment, in the sense that they need nothing else in order to be. According to this intuition properties are *intrinsic* to objects, they are *essential*, they belong to the *object-in-itself*. But no object has ever existed ‘in itself’. All objects exist *in the world*. Although we can imagine objects on their own — we can leave the mental background blank, as it were — all objects that have ever actually existed have done so in environments. All things are, on every occasion, surrounded by other things. All objects are parts of larger wholes. Molecules are parts of gasses, engines are parts of cars, and Games of Life exist within computers. The same is true of stars in galaxies, individuals in societies, cells in bodies, neurons in brains, and right down the tertiary structure of proteins in their enzymatic environment. In none of these cases are objects born in isolation and subsequently come together to form wholes, rather the object comes into being as part of the whole.

Moreover the properties and rules governing the behaviour of objects *depend* on the properties of those larger wholes, and they do so in two ways. First, properties of parts are *causally* dependent on the properties of wholes: cooling the container causes the molecules to stop colliding, and the gliders and blinkers in the Game of Life stop gliding and blinking if the program is interrupted.

Second, properties of parts are *conceptually* dependent on wholes: the power measured at the drive shaft ‘is’ the power of the engine because of the way in which it is connected to the rest of the car, and the rules of a particular CA ‘are’ the rules of the Game of Life because of the types of pattern that they generate<sup>8</sup>. Wholes are often described as being emergent products of their parts. This is true, but it overlooks the fact that *parts* are also emergent products of *wholes*.

### 2.3 Downwards Causation

Downwards causation (i.e. the causal dependence of parts on wholes) has had a disreputable history in the philosophy of science ever since it was proposed as a solution to the mind-body problem by Sperry, Popper and Eccles (1977). It has also become unpopular in the context of social studies by its association with strongly structuralist analyses of history, in which the actions of the individual are determined by higher social structures. Szentágothai admits that defending downwards causation will confirm his image as a ‘crazy Hungarian and an impossible romantic adventurer’ (1984), whilst Bedau notes that

although [downward causation] is logically possible, it is uncomfortably like magic. How does an irreducible but supervenient downward causal power arise, since by definition it cannot be due to the aggregation of the micro-level potentialities? Such causal powers would be quite unlike anything within our scientific ken. This not only indicates how they will discomfort reasonable forms of materialism. Their mysteriousness will only heighten the traditional worry that emergence entails illegitimately getting something from nothing. . . . But the most disappointing aspect of [downwards causation] is its apparent scientific irrelevance. . . . We should avoid proliferating mysteries beyond necessity. To judge from the available evidence, [downward causation] is one mystery which we don’t need. (Bedau, 1997, p377)

But downward causation is neither mysterious nor superfluous. For example, suppose we measure the mean momentum of a particular gas molecule over a period of time, and then compress the container by 20%. From just this information we can accurately predict that the mean momentum of the molecule will rise by a proportionate amount. This is an example of prediction using downwards causation that is both easy and reliable. The reductionist would claim that we could have produced a similar prediction given enough information about the exact trajectories of the other molecules. But this ignores the fact that molecules are not perfectly elastic billiard balls. If they were then gases would never condense. The determined reductionist would also need information about the exact structure of the electronic orbits of the molecules in order to predict the results of their interactions. In contrast the analysis that uses downwards causation is easy, reliable, and theoretically sound. If one ignores the kind of empirical regularity on which it is based then one has missed an important fact about the behaviour of the system. In short, if one is inclined towards pragmatism — with a small ‘p’ — in science, then downwards causation is as pragmatically useful and theoretically respectable as any other sort.

The same argument can also be cashed out in terms of counterfactuals (Lewis, 1986). Suppose one holds that *C* causes *E* iff *E* would not have happened if *C* had not. In our example the stochastic

---

<sup>8</sup>Conceptual dependence will turn out to be important when considering the relationship between mental states and the world (section 5.3), and between genes and organisms (9.3). But for most of the rest of this chapter I concentrate on the strictly causal relationship between parts and wholes.

nature of the system ensures that the effect of compressing the gas on the average momentum of the individual molecule would have been the same no matter what particular sequence of collisions occurred: the lower level events are overdetermined by their immediate causes. Therefore it is the higher, rather than lower, events that should properly be described as the cause in this case.

One objection to the possibility of downward causation is that it seems to imply causal overdetermination: if the momentum of a molecule is caused by its own history of collisions then how can it also be caused by the gas being compressed in a cylinder? But the idea that there must be a unique efficient cause of any event is an unnecessary hangover from Aristotle. Any real phenomenon is a dense web of causal processes. We can usefully pick out certain of these processes as being more important for statistical or discursive reasons, but this does not require that those properties be regarded as the *unique* cause of an event. Of course all instances of downwards causation will be mediated by local, lower, mechanisms. But once we abandon Aristotle's insistence on a unique efficient cause, then there need be no contradiction between regarding both higher and lower events as proper causes.

However there is one notion of cause that does necessarily exclude the possibility of downwards causation. This is to demand that observed correlations are only causal if they are instances of a covering law. So, for example, compressing the cylinder may effect the motion of particular molecules, but the laws governing their motion will remain the same. According to this view it is the frequency and order of micro-events which are the target of downwards causation, rather than the laws that govern them (Schröder, 1998). This version of anti-reductionism concedes that objects may be affected by their environment, but the *way* they are affected is a fixed and intrinsic property of the object. For example, moving the piston of a compressed gas will change the motion of the constituent molecules, but the *laws* that govern their collisions remain unaffected. According to this analysis downwards causation can always be eliminated — at least in principle — in favour of a causal story written in exclusively lower-level terms. The causally effective higher level property is just a shorthand description of a state properly defined at a lower level. Thus we can re-write the claim that 'the movement of the piston affected a molecule' in terms of the collision of individual particles. But as we saw above, gas molecules do *not* always obey the laws of elastic collisions. If we cool the container then the molecules stop rebounding and start to stick. Therefore changes in the environment of a part do not just affect the part. They can also change the *way* the part is affected by its environment.

There are two ways of dealing with such examples. The reductionist reaction is that if a property turns out to be dependent on the context then it should be eliminated in favour of one at a lower level. So if molecules are going to stop rebounding and start sticking then perhaps the gas should be understood at the lower, and presumably surer, level of atoms and electronic orbits. The reductionist axiom that objects and their properties are in a strong sense independent of their environment is built into their definition of a 'real object' and 'real property'. Reductionism is an *a priori* assumption about how the world is: if properties and entities are dependent on their environment then their reality is questioned; they are second-rate entities, just approximations that we find convenient. But it is highly doubtful that the reductionist strategy would *ever* yield properties that satisfy their criteria. The reductionist ends up in free fall, tumbling through the levels of description, looking for one that fits the ideal of a billiard ball universe. Until the revolution

of quantum mechanics there seemed to be a bottom level safety net, but now even that has disappeared. However the philosophical problems around quantum mechanics are eventually resolved — whether by the use of action at a distance, or the role of the conscious observer in the collapse of the wave-function, or some other equally exotic solution — it seems more than likely that it will involve some form of radical environmental-dependence of properties and fail to fit the classical reductionist picture. Quarks, for example, are currently assumed to be the ultimate components of nature, and yet they cannot even exist outside of the protons and neutrons they form. The reductionist criterion of objecthood is that an object does not require anything else in order to exist, but quarks do not meet this criterion. The persistent reductionist is in danger of being unable to find anything to reduce higher level descriptions to<sup>9</sup>.

The alternative strategy, the *anti-reductive* strategy, is to accept that properties are held by objects-in-environments, not objects-in-themselves. There are no such thing as truly intrinsic (i.e. environment-independent) properties since nothing ever exists in isolation. The situations that science usually describes as ‘isolated’ — such as a vacuum with flat electromagnetic and gravitational fields, or a laboratory arena, or a test tube or Petri dish — are themselves environments as much as occupied space, natural environments, or living organisms are; they are just different kinds of environment. (Similar considerations force us to conclude that one does not find one’s ‘true’ self in a Buddhist retreat, just a different one.) When we say that an object weighs *X*, or has a mass of *Y*, or that it has the colour *Z* — i.e. whenever we predicate a property of an object — we actually mean that the object *on earth* weighs *X*, or has a mass of *Y in our inertial frame of reference*, or has colour *Z at room temperature*. If we put the object on the moon, or in a rocket, or in a kiln, then its properties will change accordingly. Objects only ever exist in environments, and the environment can affect even the seemingly most fundamental and intrinsic of properties. An object may be made of gold, but put it in *aqua regia* and it soon dissolves; therefore its continued existence as a gold object is dependent on its environment *not being aqua regia*.

Science proceeds by trying to find objects and properties that are ‘robust’ — i.e. that are constant across environments — and it is often successful. But we can draw two different conclusions from this success. The first is that those properties are essential and intrinsic to the object, and held independently of the environment. Or we can conclude that those properties are held in, and due to, that range of environments. Although the two positions account for any given set of empirical evidence equally well, the former position is far stronger in that it implies (or, rather, assumes) that the same properties will be held in other, future, environments. Science is built on extrapolating from observed cases to unobserved situations. We assume that the physical constants measured in the accelerator are the same outside our light cone, that the biochemistry of the Petri dish proceeds in the same way in the living cell, or that the psychological behaviour exhibited in the laboratory would happen in everyday life. Sometimes this is true but, as Feynman points out, in each case this is an assumption and should be acknowledged as such:

Of course this means that science is uncertain; the moment that you make a proposition about a region of experience that you have not directly seen then you must be uncertain. But we must make statements about the regions that we have not seen, or the whole business is no use . . . We have to make guesses in order to give any utility

---

<sup>9</sup>B.C. Smith (1996, ch5) also argues that modern field-theoretic physics does not supply us with ready-made objects in the way the reductionist fondly imagines it does.

at all to science. In order to avoid simply describing experiments that have been done, we have to propose laws beyond their observed range. There is nothing wrong with that, despite the fact that it makes science uncertain. If you thought that science was certain — well, that is just an error on your part. (1965, p76)

Or, as Hume put it 200 years before,

Even after the observation of the frequent conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience. (*A Treatise of Human Nature* II.3.12)

These are simple truths, but they are ones that the reductionist seems unwilling to acknowledge.

## 2.4 Conclusion

The debate between reductionism and anti-reductionism is a debate over what kinds of questions one should ask in order to understand how physical systems work, and the kinds of answers one should expect. Now the reductionist certainly asks valid questions, and their answers are certainly true and useful. But we should not infer from this success that their answers are complete. In particular they ignore the way in which objects are dependent on their surroundings, and not just their insides. This simple intuition will turn out to have important consequences, especially when the objects under consideration include intelligent agents.

## Bibliography

- Adami, C., Belew, R., Kitano, H., & Taylor, C. (Eds.). (1998). *Artificial Life 6: Proceedings of the Sixth International Conference on Artificial Life*. MIT Press.
- Anastasi, A. (1958). Heredity, environment, and the question *how?*. *Psychological Review*, *65*, 197–208.
- Anscombe, G. (1958). Modern moral philosophy. In *Ethics, Religion and Politics*. Blackwell. (Collection published 1981).
- Ariew, A. (1997). Innateness and canalization. *Philosophy of Science*, *63* (Proceedings), S19–S27.
- Ashby, W. (1952). *Design For a Brain: The Origin of Adaptive Behaviour*. Chapman Hall.
- Babbage, C. (1864). *Passages From the Life of a Philosopher*. Longman.
- Barkow, J., Cosmides, L., & Tooby, J. (Eds.). (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press.
- Barlow, H. (1972). Single units and sensation: a neuron doctrine for perceptual psychology?. *Perception*, *1*, 371–394.
- Barlow, R., Powers, M., Kass, L., Fioralice, R. W., Camara, M., & Howard, H. (1984). Vision in *limulus* mating behavior during the day and at night. *Biological Bulletin*, *167*.
- Barlow, R., Powers, M., Kass, L., & Howard, H. (1986). Migration of *limulus* for mating: relation to lunar phase, tide height, and sunlight. *Biological Bulletin*, *171*.
- Barnett, L. (1998). Ruggedness and neutrality: The *NKp* family of fitness landscapes.. In Adami et al. (Adami, Belew, Kitano, & Taylor, 1998).
- Bateson, P. (1987). Biological approaches to the study of behavioural development. *International Journal of Behavioural Development*, *10*, 1–22.
- Bedau, M. (1997). Weak emergence. In Tomberlin, J. (Ed.), *Mind, Causation, and World*, Vol. 11 of *Philosophical Perspectives*, pp. 375–399. Blackwell.
- Beer, R. (1990). *Intelligence as Adaptive Behaviour: An Experiment in Computational Neuroethology*. Academic Publishers, Boston.
- Beer, R. (1995a). Computational and dynamical languages for autonomous agents. In Van Gelder, T., & Port, R. (Eds.), *Mind as Motion*. MIT Press.
- Beer, R. (1995b). A dynamical systems perspective on autonomous agents. *Artificial Intelligence*, *72*, 173–215.
- Beer, R., & Chiel, H. (1997). The brain has a body: Adaptive behaviour emerges from interactions of nervous system, body and environment. *Trends in Neuroscience*, *20*(12), 553–557.
- Beer, R., & Gallagher, J. (1992). Evolving dynamical neural networks for adaptive behaviour. *Adaptive Behaviour*, *1*(1).

- Bennett, A. (1996). Do animals have cognitive maps?. *Journal of Experimental Biology*, 199, 219–224.
- Berg, D., & Howe, M. (1989). *Mobile DNA*. American Society of Microbiology.
- Bermúdez, J. (1995). Nonconceptual content: From perceptual experience to subpersonal computational states. *Mind and Language*, 10(4), 333–369.
- Bhaskar, R. (1978). *A Realist Theory of Science*. Harvester.
- Bickhard, M. (1998). Robots and representations.. In Pfeifer et al. (Pfeifer, Blumberg, Meyer, & Wilson, 1998).
- Bigelow, J., & Pargetter, R. (1987). Functions. *Journal of Philosophy*, 84, 181–196.
- Bishop, J. (1983). Cellular oncogenes and retroviruses. *Annual Review of Biochemistry*, 52, 301–354.
- Bishop, J. (1985). Viral oncogenes. *Cell*, 42, 23–38.
- Block, N. (1981). Psychologism and behaviourism. *Philosophical Review*, 90, 5–43.
- Bonner, J. (1988). *The Evolution of Complexity*. Princeton University Press.
- Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*. MIT Press.
- Brandon, R. (1982). The levels of selection. In Asquith, P., & Nickles, T. (Eds.), *PSA 1982*, Vol. 1. The Philosophy of Science Association.
- Breathnach, R., & Chambon, P. (1981). Organization and expression of eukaryotic split genes coding for proteins. *Annual Review of Biochemistry*, 50, 349–383.
- Brenner, R. (1986). The social basis of economic development. In Roemer, J. (Ed.), *Analytical Marxism*. Cambridge University Press.
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Brooks, R. (1995). Intelligence without reason. In Steels, L., & Brooks, R. (Eds.), *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*. Lawrence Erlbaum.
- Brown, J., & Burton, R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155–192.
- Bullock, S. (1998). A continuous evolutionary simulation model of the attainability of honest signalling equilibria.. In Adami et al. (Adami et al., 1998).
- Buss, L. (1987). *The Evolution of Individuality*. Princeton University Press.
- Calabretta, R., Nolfi, S., Parisi, D., & Wagner, G. (1998). Emergence of functional modularity in robots.. In Pfeifer et al. (Pfeifer et al., 1998).
- Callinicos, A. (1987). *Making History: Agency, Structure and Change in Social Theory*. Polity, Cambridge.
- Carew, T., Hawkins, R., & Kandel, E. (1983). Differential classical conditioning of a defensive withdrawal reflex in *Aplysia californica*. *Science*, 219, 397–400.
- Carrier, M., & Roubertoux, P. (1984). Genotype and postnatal maternal effects on early development. *Behaviour Genetics*, 14(6).

- Carlier, M., Roubertoux, P., & Pastoret, C. (1991). The Y chromosome effect on intermale aggression in mice depends on the maternal environment. *Genetics*, *129*, 231–236.
- Carling, A. (1993). Analytical marxism and historical materialism: The debate on social evolution. *Science and Society*, *57*(1), 31–65.
- Carnap, R. (1953). Testability and meaning. In Feigl, H., & Brodbeck, M. (Eds.), *Readings in the Philosophy of Science*. Appleton-Century.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford University Press.
- Chong, A., & Zanforlin, L. (1999). Technology and epidemics. IMF working paper 99/125, International Monetary Fund, [www.imf.org](http://www.imf.org).
- Chrisley, R. (1995). Taking embodiment seriously: Nonconceptual content and robotics. In Glymour, C., Ford, K., & Hayes, P. (Eds.), *Android Epistemology*. AAAI Press/MIT Press.
- Clark, A., & Wheeler, M. (1998). Bringing representation back to life.. In Pfeifer et al. (Pfeifer et al., 1998).
- Cliff, D. (1991). Computational neuroethology: A provisional manifesto. In Meyer, J., & Wilson, S. (Eds.), *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour*, pp. 29–39. MIT Press.
- Cliff, D., Husbands, P., Meyer, J., & Wilson, S. (Eds.). (1994). *From Animals to Animats 3: Proceedings of the Third International Conference on the Simulation of Adaptive Behaviour*. MIT Press.
- Cliff, D., & Noble, J. (1997). Knowledge-based vision and simple visual machines. *Philosophical Transactions of the Royal Society: Biological Sciences*, *352*, 1165–1175.
- Cohen, G. (1978). *Karl Marx's Theory of History: A Defence*. Clarendon Press.
- Cohen, G. (1980). Functional explanation: Reply to Elster. *Political Studies*, *28*(1), 129–135.
- Crane, T. (1990). The language of thought: no syntax without semantics. *Mind and Language*, *5*, 187–212.
- Crane, T. (1992). The non-conceptual content of experience. In Crane, T. (Ed.), *The Contents of Experience. Essays on Perception*. Cambridge University Press.
- Crow, J. (1979). Genes that violate Mendel's rules. *Scientific American*, *240*, 104–13.
- Crutchfield, J. (1994). Is anything ever new?: Considering emergence. In Cowan, G., Melzner, D., & Pines, D. (Eds.), *Integrative Themes*, Vol. 19 of *Santa Fe Institute Studies in the Sciences of Complexity*. Addison-Wesley.
- Cummins, R. (1975). Functional analysis. *Journal of Philosophy*, *72*, 741–765.
- Cummins, R., & Schwarz, G. (1987). Radical connectionism. *The Southern Journal of Philosophy*, *26* *supp*, 43–72.
- Cushing, J. (1982). Models and methodologies in current theoretical high-energy physics. *Synthese*, *50*(1), 5–101.
- Cussins, A. (1992a). Content, embodiment and objectivity: The theory of cognitive trails. *Mind*, *101*(404), 651–688.

- Cussins, A. (1992b). The limitations of pluralism. In Charles, D., & Lennon, K. (Eds.), *Reduction, Explanation, and Realism*. Oxford University Press.
- Davidson, D. (1980). *Actions and Events*. Clarendon Press.
- Dawkins, M. (1995). *Unravelling Animal Behaviour*. Longman Scientific and Technical.
- Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.
- Dawkins, R. (1982). *The Extended Phenotype*. Freeman.
- Dawkins, R. (1993). Viruses of the mind. In Dahlbohm, B. (Ed.), *Dennett and His Critics*. Blackwell.
- Dawkins, R. (1994). Burying the vehicle. *Behavioral and Brain Sciences*, 17(4), 616–7.
- Dennett, D. (1978). Why not the whole iguana?. *Behavioral and Brain Sciences*, 1(1), 103–104.
- Dennett, D. (1987). *The Intentional Stance*. MIT Press.
- Dennett, D. (1990). The interpretation of texts, people, and other artifacts. *Philosophy and Phenomenological Research*, L, supplement.
- Dennett, D. (1991). Real patterns. *The Journal of Philosophy*, 88(3), 27–51.
- Dennett, D. (1994). *E pluribus unum?*. *Behavioral and Brain Sciences*, 17(4), 617–8.
- Dennett, D. (1995). *Darwin's Dangerous Idea*. Simon and Schuster.
- Dennett, D. (1996). The scope of natural selection. *The Boston Review*, October/November, 34–38.
- Dennett, D. (1997). The evolution of evaluators. International School of Economic Research, University of Sienna. Volume on Evolutionary Economics.
- Dennett, D. (1998a). *Brainchildren*. MIT Press.
- Dennett, D. (1998b). Snowmobiles, horses, rats, and memes: A comment on *A critique of Evolutionary Archeology* by James L. Boone and Eric Alden Smith. *Current Anthropology*.
- Dennett, D. (1999). The evolution of culture.. Charles Simonyi Lecture, Oxford University.
- Desmond, A., & Moore, J. (1992). *Darwin*. Penguin.
- Diamond, J. (1997). *Guns, Germs, and Steel*. Jonathan Cape.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press.
- Dreyfus, H. (1972). *What Computers Can't Do*. MIT Press.
- Dreyfus, H., & Dreyfus, S. (1988). Making a mind versus modelling the brain: Artificial intelligence back at a branch point. In Boden, M. (Ed.), *The Philosophy of Artificial Intelligence*. Oxford University Press.
- Dummett, M. (1981). *Frege's Philosophy of Language*. Duckworth. (Second edition).
- Dummett, M. (1991a). *Frege's Philosophy of Mathematics*. Duckworth.
- Dummett, M. (1991b). *The Interpretation of Frege's Philosophy*. Duckworth.

- Elman, J. (1995). *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press.
- Elster, J. (1980). Cohen on Marx's theory of history. *Political Studies*, 28(1), 121–8.
- Elster, J. (1983). *Explaining Technical Change*. Cambridge University Press.
- Engels, F. (1987). *The Dialectic of Nature and Anti-Duhring*, Vol. 25 of *Karl Marx and Frederick Engels: Collected Works*. Lawrence Wishart.
- Evans, G. (1982). *Varieties of Reference*. Oxford University Press.
- Ewald, P. (1994). *Evolution of Infectious Disease*. Oxford University Press.
- Faith, J. (1995). The role of oscillations in the olfactory bulb. Master's thesis, School of Cognitive and Computing Sciences, University of Sussex.
- Faith, J. (1997). In defence of functional analysis.. In Husbands, & Harvey (Husbands & Harvey, 1997).
- Faith, J. (1998). Why gliders don't exist.. In Adami et al. (Adami et al., 1998).
- Feynman, R. (1963). *The Feynman Lectures on Physics*, Vol. 1. Addison Wesley.
- Feynman, R. (1964). The motion of planets around the sun. In Goodstein, D., & Goodstein, J. (Eds.), *Feynman's Lost Lecture*. Cape. (Published 1996).
- Feynman, R. (1965). *The Character of Physical Law*. Penguin.
- Fine, A. (1984). The natural ontological attitude. In Leplin, J. (Ed.), *Scientific Realism*. University of California Press.
- Fisher, R. (1930). *The genetical theory of natural selection*. Oxford University Press.
- Fisher, R. (1951). Limits to intensive production in animals. *British Agricultural Science*, 4, 217–18.
- Fodor, J. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. MIT Press.
- Fodor, J. (1990). A theory of content II. In *A Theory of Content and Other Essays*. MIT Press.
- Fodor, J. (1991). Methodological solipsism considered as a research strategy in cognitive psychology. In Rosenthal, D. (Ed.), *The Nature of Mind*. Oxford University Press.
- Fodor, J. (1994). *The Elm and the Expert*. MIT Press.
- Fodor, J. (1998). Look! — a review of *Consilience: The Unity of Knowledge*, by E.O.Wilson. *London Review of Books*, 20(21).
- Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3–72.
- Frankel, J. (1983). What are the developmental underpinnings of evolutionary changes in protozoan morphology?. In Goodwin, B., Holder, N., & Wylie, C. (Eds.), *Development and Evolution: Sixth Symposium of the British Society for Developmental Biology*, pp. 279–314. Cambridge University Press.
- Freeman, W., & Skarda, C. (1985). Spatial EEG patterns, non-linear dynamics and perception: the neo-Sherringtonian view. *Brain Research Reviews*, 10, 147–175.

- Frege, G. (1891). Function and concept. In Geach, & Black (Geach & Black, 1960).
- Frege, G. (1892). On sense and reference. In Geach, & Black (Geach & Black, 1960).
- Gärdenfors, P. (1990). Induction, conceptual spaces and AI. *Philosophy of Science*, 57, 78–95.
- Geach, P., & Black, M. (Eds.). (1960). *Translations from the Philosophical Writings of Gottlob Frege*. Blackwell.
- Getting, P. (1989). Emerging principles governing the operation of neural networks. *Annual Review of Neuroscience*, 12, 185–204.
- Gibson, J. (1966). *The Senses Considered As Perceptual Systems*. Houghton Mifflin.
- Gibson, J. (1979). *The Ecological Approach To Visual Perception*. Houghton Mifflin.
- Glennan, S. (1996). Mechanism and the nature of causation. *Erkenntnis*, 44, 49–71.
- Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.
- Gonzales, A. *et al.* (1985). Developmental outcome at 1yr of very low-birth-weight inborn neonates of low-income families. *Pediatric Research*, 19(4).
- Goodman, N. (1955). *Fact, Fiction and Forecast*. Harvard University Press.
- Goodnight, C. (1990). Experimental studies of community evolution I: The response to selection at the community level. *Evolution*, 44, 1625–1636.
- Gould, S. (1977). *Ontogeny and Phylogeny*. Harvard University Press.
- Gould, S. (1978). *Ever Since Darwin: Reflections in Natural History*. Deutsch.
- Gould, S., & Lewontin, R. (1978). The Spandrels of San Marco and the Panglossian Paradigm: A critique of the adaptationism programme. *Proceedings of the Royal Society*, 205, 581–598.
- Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, 144, 517–46.
- Griffith, A., & Suzuki, D. (1993). *An Introduction to Genetic Analysis*. Freeman.
- Gusella, J. (1986). DNA polymorphism and human disease. *Annual Review of Biochemistry*, 55, 831–854.
- Haldane, J. (1932). *The Causes of Evolution*. Longmans, Green.
- Halder, G., Callerts, P., & Gehring, W. (1995). Induction of ectopic eye by targeted expression of the eyeless gene in *drosophila*. *Science*, 267, 1788–92.
- Hamilton, W. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7, 1–52.
- Harris-Warrick, R., & Marder, E. (1991). Modulation of neural networks for behaviour. *Annual Review of Neuroscience*, 14, 39–57.
- Harvey, I. (1992). Untimed and misrepresented: Connectionism and the computer metaphor. Cognitive science research paper 245, School of Cognitive and Computing Sciences, University of Sussex.
- Harvey, I., Husbands, P., & Cliff, D. (1994). Seeing the light: artificial evolution, real vision.. In Cliff *et al.* (Cliff, Husbands, Meyer, & Wilson, 1994), pp. 392–401.

- Harvey, I., Husbands, P., Cliff, D., Thompson, A., & Jakobi, N. (1997). Evolutionary robotics: The Sussex approach. *Robotics and Autonomous Systems*, 20, 205–224.
- Hawkins, R., Abrams, T., Carew, T., & Kandel, E. (1983). A cellular mechanism of classical conditioning in *Aplysia*: Activity-dependent amplification of presynaptic facilitation. *Science*, 219, 400–405.
- Heath, A. *et al.* (1985). Educational policy and the heritability of educational attainment. *Nature*, 314, 734–46.
- Heldin, C., & Westermark, B. (1984). Growth factors: mechanism of action and relation to oncogenes. *Cell*, 37, 9–20.
- Hempel, C. (1965). *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. Free Press, New York.
- Henderson, N. (1982). Human behaviour genetics. *Annual Review of Psychology*, 33, 403–40.
- Hinde, R. (1975). The concept of function. In Baerends, G., Beer, C., & Manning, A. (Eds.), *Function and Evolution in Behaviour: Essays in Honour of Niko Tinbergen*. Clarendon Press.
- Hinde, R. (1982). *Ethology: Its Nature and Relations to the Other Sciences*. Oxford University Press.
- Hobsbawm, E. (1962). *The Age of Capital: 1848–1875*. Weidenfeld and Nicholson.
- Hooper, E., & Hamilton, B. (1999). *The River: A Journey to the Source of HIV and AIDS*. Little Brown.
- Hoyle, G. (1984). The scope of neuroethology. *Behavioral and Brain Sciences*, 7, 367–412.
- Hull, D. (1980). Individuality and selection. *Annual Review of Ecology and Systematics*, 11, 311–32.
- Husbands, P. (1998). Evolving robot behaviours with diffusing gas networks.. In Husbands, & Meyer (Husbands & Meyer, 1998).
- Husbands, P., & Harvey, I. (Eds.). (1997). *Proceedings of the Fourth European Conference on Artificial Life*, Complex Adaptive Systems Series. MIT Press.
- Husbands, P., & Meyer, J. (Eds.). (1998). *EvoRob98: Proceedings of the First European Conference on Evolutionary Robotics*, Paris. Springer Verlag.
- Huynen, M., Stadler, P., & Fontana, W. (1996). Smoothness within ruggedness: The role of neutrality in adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 93.
- Jacob, F. (1974). *The Logic of Living Systems: A History of Heredity*. Allen Lane.
- Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3, 318–356.
- Jakobi, N. (1997). Half-baked, *ad hoc* and noisy: Minimal simulations for evolutionary robotics.. In Husbands, & Harvey (Husbands & Harvey, 1997).
- Karmiloff-Smith, A. (1994). Beyond modularity: A developmental perspective on cognitive science. *Behavioural and Brain Sciences*, 17(4), 693–745.

- Keeley, B. (1998). Artificial Life for philosophers. *Philosophical Psychology*, 11(2), 251–260.
- Kennedy, H., Martin, K., Dehay, C., Bullier, J., Killackey, H., Chalupa, L., & Kaas, J. (Eds.). (1995). *Cerebral Cortex: Function and Development*, Lyon, France.
- Kim, J. (1984). Concepts of supervenience. *Philosophy and Phenomenological Research*, 45, 153–176.
- Kim, J. (1992). Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research*, 52(1), 1–26.
- Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.
- Kirsh, D. (1991). Today the earwig, tomorrow man?. *Artificial Intelligence*, 47, 161–184.
- Kitcher, P. (1984). 1953 and all that. A tale of two sciences. *Philosophical Review*, 93, 335–74.
- Kline, M. (1954). *Mathematics in Western Culture*. Allen and Unwin.
- Kline, M. (1980). *Mathematics: The Loss of Certainty*. Oxford University Press.
- Knierim, J., Kudrimoti, H., & McNaughton, B. (1995). Place cells, head direction cells, and the learning of landmark stability. *The Journal of Neuroscience*, 15(3), 1648–1659.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Harvard University Press.
- Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago University Press.
- Kutsch, W., Schwarz, G., Fischer, H., & Kautz, H. (1993). Wireless transmission of muscle potentials during free-flight of a locust. *Journal of Experimental Biology*, 185, 367–373.
- Lansner, A., & Liljenström, H. (1994). Computer models of the brain — How far can they take us?. *Journal of Theoretical Biology*, 171, 61–73.
- Laurel, B. (1997). Interface agents: Metaphors with character. In Friedman, B. (Ed.), *Human Values and the Design of Computer Technology*, No. 72 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, Ca.
- Levins, R., & Lewontin, R. (1985). *The Dialectical Biologist*. Harvard University Press.
- Lewin, B. (1997). *Genes VI*. Oxford University Press.
- Lewis, D. (1986). Causation. In *Philosophical Papers*, Vol. 2. Oxford University Press.
- Lewontin, R. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1, 1–18.
- Lewontin, R. (1998). The evolution of cognition: Questions we will never answer. In Scarborough, D., & Sternberg, S. (Eds.), *An Invitation to Cognitive Science: Methods, Models, and Conceptual Issues*. MIT Press.
- Loehlin, J. (1979). Combining data from different groups in human behaviour genetics. In Royce, J., & Mos, L. (Eds.), *Theoretical Advances in Behaviour Genetics*. Sijthoff and Noordhof.
- Malinowski, B. (1944). *A Scientific Theory of Culture*.
- Marr, D. (1982). *Vision*. Freeman.
- Marx, K., & Engels, F. (1846). *The German Ideology*. Lawrence Wishart. (Edition published 1970).

- Marx, K., & Engels, F. (1848). *Manifesto of the Communist Party*.
- Maturana, H., & Varela, F. (1980). *Autopoiesis and Cognition*. No. 42 in Boston Studies in the Philosophy of Science. D. Reidel.
- Maynard Smith, J. (1961). Evolution and history. In Banton, M. (Ed.), *Darwinism and the Study of Society*. Tavistock.
- Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics*, 9, 31–56.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.
- Maynard Smith, J. (1989a). *Did Darwin Get it Right? Essays on Games, Sex and Evolution*. Chapman and Hall.
- Maynard Smith, J. (1989b). *Evolutionary Genetics*. Oxford University Press.
- Maynard Smith, J. (1991). Honest signalling: The Philip Sydney game. *Animal Behaviour*, 42, 1034–5.
- Maynard Smith, J., & Sondhi, K. (1960). The genetics of a pattern. *Genetics*, 45, 1039–1050.
- Maynard Smith, J. *et al.* (1985). Developmental constraints and evolution. *Quarterly Review of Biology*, 60(3), 265–287.
- Mayr, E. (1963). *Animal Species and Evolution*. Harvard University Press.
- Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Belknap Publishers.
- McGinn, C. (1989). *Mental Content*. Blackwell.
- McNaughton, B. *et al.* (1996). Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *Journal of Experimental Biology*, 199, 173–185.
- Melnyk, A. (1995). Two cheers for reductionism: Or, the dim prospects for non-reductive materialism. *Philosophy of Science*, 62, 370–388.
- Miglino, O., Denaro, D., Tascini, G., & Parisi, D. (1998). Detour behaviour in evolving robots: Are internal representations necessary?.. In Husbands, & Meyer (Husbands & Meyer, 1998).
- Miller, R. (1981). Marx and Aristotle: A kind of consequentialism. *Canadian Journal of Philosophy, Supp 7*.
- Millikan, R. (1984). *Language, Thought and Other Biological Categories*. MIT Press.
- Millikan, R. (1989). Biosemantics. *The Journal of Philosophy*, 86(6), 281–97.
- Millikan, R. (1993). *White Queen Psychology, and Other Essays for Alice*. MIT Press.
- Millikan, R. (1999). Wings, spoons, pills, and quills: A pluralist theory of function. *The Journal of Philosophy*, 96(4), 191–206.
- Mondada, F., & Floreano, D. (1996). Evolution of neural control structures: some experiments on mobile robots. *Robotics and Autonomous Systems*, 16, 183–195.
- Monod, J. (1997). *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*. Penguin. (First published 1971).

- Morris, M. (1986). Causes of behaviour. In Stevenson, L., Squires, R., & Haldane, J. (Eds.), *Mind, Causation and Action*. Blackwell.
- Nagel, E. (1961). *The Structure of Science*. Routledge.
- Neander, K. (1991). Functions as selected effects: The conceptual analyst's defence. *Philosophy of Science*, 58, 168–84.
- Needham, J. (1933). On the dissociability of the fundamental processes in ontogenesis. *Biological Review*, 8, 180–223.
- Newell, A., & Simon, H. (1972). *Human Problem Solving*. Prentice-Hall.
- Nilsson, N. (ed). (1984). Shakey the robot. Tech. rep. 323, SRI A.I. Center.
- Norman, D. (1993). Cognition in the head and in the world: An introduction to the special issue on situated action. *Cognitive Science*, 17, 1–6.
- Nosten, M. (1989). Early development in mice VI: Additive and interactive effects of offspring genotype and maternal environments. *Physiology and Behaviour*, 45, 955–61.
- Nosten, M., & Roubertoux, P. (1988). Uterine and cytoplasmic effects on pup eyelid opening in two inbred strains of mice. *Physiology and Behaviour*, 43, 167–71.
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34, 171–175.
- Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation and confirmation of numerical models in the earth sciences. *Science*, 263, 641–646.
- Orgel, L. (1979). Selection *in vitro*. *Proceedings of the Royal Society of London*, 205, 435–442.
- Papineau, D. (1984). Representation and explanation. *Philosophy of Science*, 51, 550–572.
- Peacocke, C. (1992). *A Study of Concepts*. MIT Press.
- Peacocke, C. (1994). Content, computation and externalism. *Mind and Language*, 9(3), 303–335.
- Pfeifer, R., Blumberg, B., Meyer, J., & Wilson, S. (Eds.). (1998). *From Animals to Animats: Proceedings of the Fifth International Conference on the Simulation of Adaptive Behaviour*. MIT Press.
- Phillips, K., Fulker, D., & Rose, R. (1987). Path analysis of seven fear factors in adult twin and sibling pairs and their parents. *Genetic Epidemiology*, 4, 345–55.
- Pinker, S. (1998). *How The Mind Works*. Allen Lane.
- Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annual Review of Biochemistry*, 55, 339–372.
- Plomin, R. (1986). *Development, Genetics, and Psychology*. Erlbaum.
- Plomin, R. (1990). Trying to shoot the messenger for his message. *Behavioral and Brain Sciences*, 13(1), 144.
- Plomin, R., & DeFries, J. (1983). The Colorado Adoption Project. *Child Development*, 54, 276–89.
- Plomin, R., DeFries, J., & Fulker, D. (1988). *Nature and Nurture During Infancy and Early Childhood*. Cambridge University Press.

- Plomin, R., DeFries, J., & Loehlin, J. (1977). Genotype-environment interaction and correlation in the analysis of human behaviour. *Psychological Bulletin*, *84*(2), 309–322.
- Plomin, R., DeFries, J., & Loehlin, J. (1985). Genetic and environmental components of "environmental" influences. *Developmental Psychology*, *21*(2), 391–402.
- Popper, K. (1960). *The Poverty of Historicism*. Routledge.
- Popper, K., & Eccles, J. (1977). *The Self and its Brain*. Springer.
- Preston, K. (1998). A pluralist theory of function. *The Journal of Philosophy*, *XCV*(5), 215–254.
- Prigogine, I. (1962). *Non-Equilibrium Statistical Mechanics*. Interscience.
- Pritchard, D. (1986). *Foundations of Developmental Genetics*. Taylor and Francis.
- Putnam, H. (1973). Explanation and reference. In Pearce, G., & Maynard, P. (Eds.), *Conceptual Change*. Reidel.
- Putnam, H. (1975). The meaning of meaning. In *Mind, Language and Reality*, Vol. 1 of *Philosophical Papers*. Cambridge University Press.
- Putnam, H. (1981). *Reason, Truth and History*. Cambridge University Press.
- Putnam, H. (1998). A politics of hope: Review of *Achieving Our Country* by Richard Rorty. *Times Literary Supplement*. (May 22).
- Quine, W. (1951). Two dogmas of empiricism. *Philosophical Review*, *60*(1).
- Quine, W. (1960). *Word and Object*. MIT Press.
- Quine, W. (1969). Natural kinds. In *Ontological Relativity and Other Essays*. Columbia University Press.
- Radcliffe-Brown, A. (1952). *Structure and Function in Primitive Society*. Free Press.
- Reichenbach, H. (1938). *Experience and Prediction*. University of Chicago Press.
- Reznikoff, W., Siegele, D., Cowing, D., & Gross, C. (1985). The regulation of transcription initiation in bacteria. *Annual Review of Genetics*, *19*, 355–387.
- Roitblat, H. (1994). Mechanism and process in animal behaviour.. In Cliff et al. (Cliff et al., 1994).
- Rorty, R. (1991a). Antirepresentationalism, ethnocentrism and liberalism. In *Objectivity, Relativism, and Truth* (Rorty, 1991b).
- Rorty, R. (1991b). *Objectivity, Relativism, and Truth*, Vol. 1 of *Philosophical Papers*. Cambridge University Press.
- Rorty, R. (1991c). Solidarity or objectivity?. In *Objectivity, Relativism, and Truth* (Rorty, 1991b).
- Rorty, R. (1998). *Achieving Our Country: Leftist Thought in Twentieth-Century America*. Harvard University Press.
- Roubertoux, P., Carlier, M., & Nosten-Bertrand, M. (1990). Additive and interactive effects of genotype and maternal environment. *Advances in the Study of Behaviour*, *19*, 205–247.
- Rowlands, M. (1997). Teleological semantics. *Mind*, *106*, 279–303.
- Runciman, W. (1998). *The Social Animal*. Harper Collins.

- Russell, B. (1905). On denoting. In Marsh, R. (Ed.), *Bertrand Russell: Logic and Knowledge. Essays 1901–1950*. Allen Unwin. (Collection published 1965).
- Sampson, G. (1997). *Educating Eve: The 'Language Instinct' Debate*. Cassell.
- Scheier, C., & Pfeifer, R. (1998). Exploiting embodiment for category learning.. In Pfeifer et al. (Pfeifer et al., 1998).
- Schönemann, P. (1997). On models and muddles of heritability. *Genetica*, 99, 97–108.
- Schröder, J. (1998). The concept of emergence: Non-deducibility or downward casuation?. *Philosophical Quarterly*, 48(193), 433–454.
- Sherrington, C. (1906). *The Integrative Action of the Nervous System*. Scribner.
- Sloman, A. (1989). On designing a visual system (toward a Gibsonian computational model of vision). Cognitive science research paper 146, School of Cognitive and Computing Sciences, University of Sussex, ([www.cogs.susx.ac.uk](http://www.cogs.susx.ac.uk)).
- Smith, B. (1996). *On the Origin of Objects*. MIT Press.
- Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11, 1–74.
- Sober, E. (1981). Evolutionary theory and the ontological status of properties. *Philosophical Studies*, 40, 140–176.
- Sober, E. (1982). Why logically equivalent predicates may pick out different properties. *American Philosophical Quarterly*, 19(2), 183–189.
- Sober, E. (1985). *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. MIT Press.
- Sober, E., & Lewontin, R. (1982). Artifact, cause and genic selection. *Philosophy of Science*, 49, 157–180.
- Soffe, S. (1993). Two distinct rhythmic motor patterns are driven by common premotor and motor neurons in a simple vertebrate spinal cord. *The Journal of Neuroscience*, 13(10), 4456–4469.
- Sonneborn, T. (1963). Does preformed cell structure play an essential role in cell heredity?. In Allen, J. (Ed.), *The Nature of Biological Diversity*, pp. 16–221. McGraw-Hill.
- Sonneborn, T. (1970). Gene action in development. *Proceedings of the Royal Society of London*, 176, 347–66.
- Stephenson, B. (1997). *Kepler's Physical Astronomy*. Princeton University Press.
- Sterelny, K., & Kitcher, P. (1988). The return of the gene. *The Journal of Philosophy*, 85(7), 339–361.
- Stich, S. (1981). Dennett on intentional systems. *Philosophical Topics*, 12, 38–62.
- Szentágothai, J. (1984). Downwards causation. *Annual Review of Neuroscience*, 7, 1–11.
- Tani, J., & Nolfi, S. (1998). Learning to perceive the world as articulated.. In Pfeifer et al. (Pfeifer et al., 1998).
- Tinbergen, N. (1951). *The Study of Instinct*. Oxford University Press.

- Tinbergen, N. (1965). Behaviour and natural selection. In Moore, J. (Ed.), *Ideas in Modern Biology*, Vol. 6 of *Proceedings of the International Congress of Zoology*, pp. 521–42.
- Tolman, E. (1932). *Purposive Behaviour in Animals and Men*. Appleton-Century-Crofts.
- Tootell, R., Reppas, J., Kwong, K., & Malach, R. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15(4), 3215–30.
- Torrance, J. (1985). Reproduction and development: a case for a ‘Darwinian’ mechanism in Marx’s theory of history. *Political Studies*, 33, 382–398.
- Turing, A. (1950). Can a machine think?. *Mind*, 59.
- Van Fraassen, B. (1980). *The Scientific Image*. The Clarendon Press.
- Van Gelder, T. (1992). What might cognition be if not computation ?. Cognitive Science Research Report 75, Indiana University.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
- Varmus, H. (1984). The molecular genetics of cellular oncogenes. *Annual Review of Genetics*, 18, 553–612.
- Veblen, T. (1899). The theory of the leisure class. In Lerner, M. (Ed.), *The Portable Veblen*. Viking Press. (Collection published 1948).
- Wagner, G., & Altenberg, L. (1996). Complex adaptations and the evolution of evolvability. *Evolution*, 50(3), 967–976.
- Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity-environment interactions. *Behavioral and Brain Sciences*, 13, 109–120.
- Wehner, R., Michel, B., & Antonsen, P. (1996). Visual navigation in insects: coupling of egocentric and geocentric information. *Journal of Experimental Biology*, 199, 129–140.
- Weismann, A. (1904). *The Evolution Theory*. E. Arnold.
- Wertsch, J. (Ed.). (1981). *The Concept of Activity in Soviet Psychology*. Sharpe, Armonk.
- Wheeler, M. (1994). For whom the bell tolls? The roles of representation and computation in the study of situated agents. Cognitive science research paper 320, School of Cognitive and Computing Sciences, University of Sussex, ([www.cogs.susx.ac.uk](http://www.cogs.susx.ac.uk)).
- Wheeler, M. (1998). Explaining the evolved: Homunculi, modules and internal representation.. In Husbands, & Meyer (Husbands & Meyer, 1998).
- Whittle, J. (1983). Litany and creed in the genetic analysis of development. In Goodwin, B., Holder, N., & Wylie, C. (Eds.), *Development and Evolution: Sixth Symposium of the British Society for Developmental Biology*, pp. 59–74. Cambridge University Press.
- Williams, G. (1966). *Adaptations and Natural Selection*. Princeton University Press.
- Williams, G., & Nesse, R. (1991). The dawn of Darwinian medicine. *Quarterly Review of Biology*, 66, 1–62.
- Wilson, D. (1976). Evolution on the level of communities. *Science*, 192, 1358–1360.

- Wilson, E. (1975). *Sociobiology: The New Synthesis*. Harvard University Press.
- Wimsatt, W. (1980). Reductionistic research strategies and their biases in the units of selection controversy. In Nickles, T. (Ed.), *Scientific Discoveries: Case Studies*. D.Reidel.
- Winograd, T., & Flores, F. (1986). *Understanding Computers and Cognition: A New Foundation for Design*. Ablex Norwood, NJ.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell.
- Wright, E., Levine, A., & Sober, E. (1992). *Reconstructing Marxism: Essays on Explanation and the Theory of History*. Verso.
- Wright, L. (1973). Functions. *Philosophical Review*, 82, 139–68.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16, 97–159.
- Wright, S. (1968). *Evolution and the Genetics of Populations Vol 1: Genetic and Biometric Foundations*. University of Wisconsin Press.
- Wu, R. (1978). DNA sequence analysis. *Annual Review of Biochemistry*, 47, 607–734.
- Zahavi, A. (1975). Mate selection — a selection for a handicap. *Journal of Theoretical Biology*, 53, 205–214.
- Zahavi, A. (1977). The cost of honesty (further remarks on the handicap principle). *Journal of Theoretical Biology*, 67, 603–5.