REITH LECTURES 1984: Minds, Brains and Science

John Searle

Lecture 5: A Changing Reality

TRANSMISSION: 5 December 1984 – Radio 4

In this lecture I want to discuss one of the most vexing intellectual problems of the present era: why have the methods of the natural sciences not given us the kind of payoff in the study of human behaviour that they have in subjects like physics and chemistry? And what sort of ‘social’ or ‘behavioural’ sciences can we reasonably expect anyhow? I’m going to suggest that there are certain radical differences between human behaviour and the phenomena studied in the natural sciences. And I believe that these differences account both for the failures and the successes that the methods of the natural sciences have had in the human sciences.

At the beginning I want to call your attention to an important difference between the form of commonsense explanations of human behaviour and the standard form of scientific explanations. According to the standard theory of scientific explanation, explaining a phenomenon consists in showing how its occurrence follows from certain scientific laws. These laws are universal generalisations about how things happen. For example, if you are given a statement of the relevant laws describing the behaviour of a falling body, and you know where it started from, you can actually deduce what will happen. Similarly, if you want to explain a law, you can deduce the law from some higher-level law. On this account explanation and prediction are perfectly symmetrical. You predict by deducing what will happen; you explain by deducing what has happened.

Now, whatever merit this type of explanation may have in the natural sciences, one of the things I want to emphasise in this lecture is that it is quite worthless to us in explaining human behaviour. And this is not because we lack laws for explaining individual examples of human behaviour, but rather it’s because even if we had such laws, they would still be useless to us. And I think I can easily get you to see that by asking you to imagine what it would be like if we actually had a ‘law’, that is, a universal generalisation concerning some aspect of behaviour.

Suppose that in the last election you voted for the Tories, and suppose that you voted for the Tories because you thought they would do more to solve the problem of inflation than any of the other parties. Now, let’s suppose that’s just a plain fact about why you voted for the Tories, as it is an equally plain fact you did vote for the Tories. Suppose, furthermore, that some political sociologists come up with an absolutely exceptionless universal generalisation about people who exactly fit your description—your socio-economic status, your income level, your education, your other interests, and so on.

Suppose this absolutely exceptionless generalisation is to the effect that people like you invariably vote for the Tories. Now, I want to ask: which explains why you voted
for the Tories? Is it the reason you sincerely accept? Or is it the universal generalisation?

I believe we would never accept the generalisation as the explanation of our own behaviour. The generalisation states a regularity. And knowledge of such a regularity may be useful for prediction, but it doesn’t explain anything about individual cases of human behaviour. If a generalisation were found, it itself would require explanation of the son we were after in the first place. And where human behaviour is concerned, the sort of explanation we normally seek is one that specifies the mental states—the beliefs, fears, hopes, desires, and so on—that function causally in the production of the behaviour.

Well, let’s return to our original question: why do we not seem to have laws of social sciences in the sense that we have laws of the natural sciences? Now, there are several standard answers to that question. Some philosophers point out that we don’t have a science of behaviour for the same reason that we don’t have a science of furniture. We couldn’t have such a science because there aren’t any physical features that all furniture—chairs, tables, rugs, and other items—have in common that would enable them to fall under a common set of laws of furniture. And besides, they argue, we don’t really need such a science because anything we want to explain—for example, why are wooden tables solid, or why does iron lawn furniture rust?—can already be explained by existing sciences. Similarly, there aren’t any features that all human behaviours have in common. And besides, particular things we wish to explain can be explained by physics, and physiology, and all the rest of the existing sciences.

In a related argument some philosophers point out that perhaps our concepts for describing ourselves and other human beings don’t match the concepts of such basic sciences as physics and chemistry in the right way. Perhaps, they suggest, human science is like a science of the weather. We have a science of the weather, meteorology, but it’s not a strict science because the things that interest us about the weather don’t match the natural categories we have for physics. So weather concepts such as ‘bright spots over the Midlands’ or ‘partly cloudy in London’ are not systematically related to the concepts of physics- A powerful expression of this sort of view is in Jerry Fodor’s work. He suggests that special sciences like geology or meteorology are about features of the world that can be realised in physics in a variety of ways and that this loose connection between the social sciences and the more basic science of physics is also characteristic of the social sciences.

Just as mountains and storms can be realised in different sorts of microphysical structures, so money, for example, can be physically realised as gold, or silver, or printed paper. And such disjunctive connections between the higher order phenomena and the lower order do indeed allow us to have rich sciences, but they do not allow for strict laws, because the form of the loose connections will permit of laws that have exceptions.

Another argument for the view that we cannot have strict laws connecting the mental and the physical is in Donald Davidson’s claim that the concepts of rationality, consistency and coherence are partly constitutive of our notion of mental phenomena; and these notions simply don’t relate systematically to the notions of physics. As Davidson says, they have no ‘echo’ in physics. But a difficulty with this view is that
there are lots of sciences which contain constitutive notions that similarly have no echo in physics but are none the less pretty solid sciences. Biology, for example, requires the concept of organism, and ‘organism’ has no echo in physics, but biology does not thereby cease to be a hard science.

Another view, widely held, is that the complex interrelations of our mental states prevent us from ever getting a systematic set of laws connecting them to neurophysiological states. But, once again, this argument is inconclusive. Suppose, for example, that Noam Chomsky is right in thinking that each of us has a complex set of rules of universal grammar programmed into our brains at birth. Well, there’s nothing about the complexity or interdependence of the rules of universal grammar that prevents them from being systematically realised in the neurophysiology of the brain. Interdependence and complexity by themselves are not a sufficient argument against the possibility of strict psycho-physical laws.

I find all of these accounts suggestive, but do not believe that they adequately capture the really radical differences between the mental and the physical sciences. The relation between sociology and economics on the one hand, and physics on the other, is really quite unlike the relation, for example, of meteorology, geology and biology and other special sciences to physics, and we need to try to state exactly how. Now, ideally I’d like to be able to give you a step-by-step argument to show the limitations on the possibilities of strict social sciences and yet at the same time show the real nature and power of these disciplines. I think we need to abandon once and for all the idea that the social sciences are like physics before Newton and that what we are waiting for is a set of Newtonian laws of the human mind.

Well, first one might ask: what exactly is the problem supposed to be? One might say, ‘Surely social and psychological phenomena are as real as anything else. So why can’t there be laws of their behaviour? Why should there be laws of the behaviour of molecules but not laws of the behaviour of societies?’

One way to disprove a thesis is to imagine that it’s true and then show how that supposition is somehow absurd. So let’s suppose that we actually had laws of society and laws of history that would enable us to predict when there would be wars and revolutions.

The real problem is this: whatever else wars and revolutions are, they involve lots of molecule movements. But that has the consequence that any strict law about wars and revolutions would have to match perfectly with the laws about molecule movements. In order for a revolution to start on such and such a day, the relevant molecules would have to be blowing in the right direction.

But if that’s so, then the laws that predict the revolution will have to make the same predictions at the level of revolutions and their participants that the laws of molecule movements make at the level of the physical particles. So now our original question can be reformulated: why can’t the laws at the higher level, the level of revolutions, match perfectly with the laws at the lower level, the level of particles?
Well, to see why they can’t, let’s examine some cases where there really is a perfect match between the higher order laws and the lower order laws, and then we can see how these cases differ from the social cases.

One of the all-time successes in reducing the laws at one level to those of a lower level is the reduction of the gas laws—Boyle’s law and Charles’s law—to the laws of statistical mechanics. How does the reduction work?

The gas laws concern the relations between pressure, temperature and volume of gases. They predict, for example, that if you increase the temperature of a gas in a cylinder, you will increase the pressure on the walls of the cylinder. The laws of statistical mechanics concern the behaviour of masses of small particles. They predict, for example, that if you increase the rate of movement of the particles in a gas, more of the particles will hit the walls of the cylinder and will hit them harder. Now, the reason you get a perfect match between these two sets of laws is that the explanation of temperature, pressure and volume can be given entirely in terms of the behaviour of the particles. Increasing the temperature of the gas increases the velocity of the particles, and increasing the number and velocity of the particles hitting the cylinder walls increases the pressure. So it follows that an increase in temperature will produce an increase in pressure.

Now, suppose for the sake of argument that it wasn’t like that. Suppose there was no explanation of pressure and temperature in terms of the behaviour of more fundamental particles. Then any laws at the level of pressure and temperature would be simply miraculous. Because it would be miraculous that the way that pressure and temperature were going on coincided exactly with the way that the particles were going on; if there was no systematic relation between the behaviour of the system at the level of pressure and temperature, and the behaviour of the system at the level of the particles. This example is a very simple case. So let’s take a slightly more complex example.

It is a law of ‘nutrition science’ that caloric intake equals caloric output, plus or minus fat deposit. That’s not a very fancy law perhaps, but it’s pretty realistic none the less. It has the consequence known to most of us that if you eat a lot and don’t exercise enough, you get fat. Now, this law, unlike the gas laws, is not grounded in any simple way in the behaviour of the particles. The grounding isn’t simple— because, for example, there’s a rather complex series of processes by which food is converted into fat deposits in living organisms. None the less, there’s still a grounding—however complex—of this law in terms of the behaviour of more fundamental particles. Other things being equal, when you eat a lot, the molecules will be blowing in exactly the right direction to make you fat.

Well, now, given these examples, we can argue further towards the conclusion that there will be no laws of wars and revolutions in a way that there are laws of gases and of nutrition. The phenomena in the world that we pick out with concepts like war and revolution, marriage, money and property, are not grounded systematically in the behaviour of elements at the more basic level in a way that the phenomena that we pick out with concepts like fat deposit and pressure are grounded systematically in the behaviour of elements at the more basic level.
But now the question arises: if the social and psychological phenomena aren’t grounded in this way, why aren’t they? Why couldn’t they be? Granted that they are not grounded, why not? That is, wars and revolutions, like everything else, consist of molecule movements. So, why can’t such social phenomena as wars and revolutions be systematically related to molecule movements in the same way that the relations between caloric inputs and fat deposits are systematic?

Well, to see why this can’t be so we have to ask what features social phenomena have that enable us to bind them into categories. What are the fundamental principles on which we categorise psychological and social phenomena? One crucial feature is this: for a large number of social and psychological phenomena the concept that names the phenomenon is itself a constituent of the phenomenon. So in order for something to count as a marriage ceremony or a trade union, or property or money or even a war or revolution, people involved in these activities have to have certain appropriate thoughts. And in general they have to think, that’s what it is. So, for example, in order to get married or buy property, you and other people have to think that that’s what you are doing.

This feature is crucial to social phenomena. But there’s nothing like it in the biological and physical sciences. Something can be a tree or a plant, or some person can have tuberculosis even if no one thinks, here’s a tree or a plant or this is a case of tuberculosis, and even if no one thinks about it at all.

But many of the terms that describe social phenomena have to enter into their very constitution. And this has the further result that such terms have a peculiar kind of self-referentiality. The concept ‘money’ refers to whatever people use as and think of as money. ‘Promise’ refers to whatever people intend as and regard as promises. I’m not saying that in order to have the institution of money people have to have that very word or some exact synonym in their vocabulary. Rather, they must have certain thoughts and attitudes about something in order that it count as money, and these thoughts and attitudes are part of the very definition of money.

Now, there’s another crucial consequence of this feature, The defining principles of such social phenomena set no physical limits whatever on what can count as the physical realisation of them. And this means that there can’t be any systematic connections between the physical and the social or mental properties of the phenomenon. The social features in question are determined in part by the attitudes we take towards them. The attitudes we take towards them are not constrained by the physical features of the phenomena in question. Therefore, there can’t be any matching of the mental level and the level of the physics of the sort that would be necessary to make strict laws of the social sciences possible.

So the main step in the argument for a radical discontinuity between the social sciences and the natural sciences depends on the mental character of social phenomena. And it’s this feature which all those analogies I mentioned earlier—that is, between meteorology, biology and geology—neglect. The radical discontinuity between the social disciplines on the one hand and the natural sciences on the other derives from the role of the mind in the social disciplines.
Consider Fodor’s claim that social laws will have exceptions, since the phenomena at the social level map loosely, or disjunctively, on to the physical phenomena. Once again, this doesn’t account for the radical discontinuities I’ve been calling attention to. Even if this sort of disjunction had been true up to a certain point, it’s always open to the next person to add to it in indefinitely many ways.

Suppose money had always taken a limited range of physical forms—gold, silver and printed paper, for example. Still, it’s open to the next person or society to treat something else as money. And indeed the physical realisation doesn’t matter to the properties of money as long as the physical realisation enables the stuff to be used as a medium of exchange.

‘Well,’ somebody might object to me, ‘in order to have a rigorous social science we don’t need a strict match between properties of things in the world; all we need is a strict match between the psychological properties and features of the brain. The real grounding of economics and sociology in the physical world is not in the properties of objects we find around us, it is in the physical properties of the brain. So even if thinking that something is money is essential to its being money, still thinking that it is money may well be, and indeed on your account is, a process in the brain. So, in order to show that there can’t be any strict laws of social sciences you have to show that there can’t be any strict correlations between mental states and brain state types, and you haven’t shown that.’

To see why there can’t be such correlations, let’s examine some areas where it seems likely that we will get a strict neuropsychology: that is, strict laws correlating mental phenomena and neurophysiological phenomena. Consider pain, for example: it seems reasonable to suppose that neurophysiological causes of pains, at least in human beings, are quite limited and specific. Indeed, we discussed some of them in an earlier lecture. There seems to be no obstacle in principle to having a perfect neurophysiology of pain. But what about, say, vision?

Once again, it’s hard to see any obstacle in principle to getting an adequate neurophysiology of vision. We might even get to the point where we could describe perfectly the neurophysiological conditions for having certain sorts of visual experiences. The experience of seeing that something is red, for instance.

But now here comes the hard part: though we might get systematic correlations between neurophysiology and pain, or neurophysiology and the visual experience of red, we couldn’t get similar accounts of the neurophysiology of seeing that something was money. Why not?

Granted that every time you see that there’s some money in front of you some neurophysiological process goes on, what’s to prevent it from being the same process every time?

Well, from the fact that money can have an indefinite range of physical forms; it follows that it can have an indefinite range of stimulus effects on our nervous systems. But since it can have an indefinite range of stimulus patterns on our visual systems, it would once again be a miracle if they all produced exactly the same neurophysiological effect on the brain. And what goes for seeing that something is
money goes even more forcefully for believing that it’s money. It would be nothing short of miraculous if every time someone believed that he was short of money, in whatever language and culture he had this belief in; it had exactly the same type of neurophysiological realisation. And that’s simply because the range of possible neurophysiological stimuli that could produce that very belief is infinite. Paradoxically, the way that the mental infects the physical prevents there ever being a strict science of the mental.

Notice that in cases where we do not have this sort of interaction between the social and the physical phenomena this obstacle to having strict social sciences is not present. Consider the example I mentioned earlier of Chomsky’s hypothesis of universal grammar.

Suppose each of us had innately programmed in our brains the rules of universal grammar. Since these rules are in the brain at birth and independent of any relations the organism has with the environment, there’s nothing in my argument to prevent there being strict psychophysical laws connecting these rules and features of the brain, however interrelated and complicated the rules might be.

Well, I promised to try to give you at least the sketch of a step-by-step argument. Let’s see how far I got in keeping the promise. Let’s set the argument out as a series of steps.

1. For there to be laws of the social sciences in the sense of which there are laws of physics, there must be some systematic correlation between phenomena identified in social and psychological terms and phenomena identified in physical terms. It can be as complex as the way that weather phenomena are connected with the phenomena of physics, but there has to be some systematic correlation. In the contemporary jargon, there have to be some bridge principles between the higher and the lower levels.

2. Social phenomena are in large part defined in terms of the psychological attitudes that people take. What counts as money, or a promise or a marriage, is in large part a matter of what people think of as money, a promise or a marriage.

3. This has the consequence that these categories are physically open-ended. There is, strictly speaking, no physical limit to what we can regard as, or stipulate to he, money, or a promise or a marriage ceremony.

4. That implies that there can’t be any bridge principles between the social and the physical features of the world: that is, between phenomena described in social terms and the same phenomena described in physical terms. We can’t even have the sort of loose disjunctive principles we have for the weather or digestion.

5. Furthermore, it’s impossible to get the right kind of bridge principles between phenomena described in mental terms and phenomena described in neurophysiological terms: that is, between the brain and the mind. And this is because there is an indefinite range of stimulus conditions for any given social concept. This enormous range prevents concepts that are not built into us from being realised in a way that systematically correlates mental and physical features.
I want to conclude this lecture by describing what seems to me the true character of the social sciences.

The social sciences in general are about various aspects of intentionality. Economics is about the production and distribution of goods and services. Notice that the working economist can simply take intentionality for granted. He assumes that entrepreneurs are trying to make money and that consumers would prefer to be better off rather than worse off. And the ‘laws’ of economics then state systematic fallouts or consequences of such assumptions. Given certain assumptions, the economist can deduce that rational entrepreneurs ‘ill sell where their marginal cost equals their marginal revenue.

Now, notice that the law doesn’t predict that the businessman asks himself: ‘Am I selling where marginal cost equals my marginal revenue?’ No, the law doesn’t state the content of individual intentionality. Rather, it works out the consequences of such intentionality. The theory of the firm in microeconomics works out the consequences of certain assumptions about the desires and possibilities of consumers and businesses engaging in buying, producing and selling. Macroeconomics works out the consequences of such assumptions for whole nations and societies.

Since economics is grounded not in systematic facts about physical properties such as molecular structure, in the way that, for example, chemistry is grounded in systematic facts about molecular structure, but rather, economics is grounded in facts about human intentionality, about desires, practices, states of technology and states of knowledge, it follows that economics cannot be free of history or context. Economics as a science presupposes certain historical facts about people and societies that are not themselves part of economics. And when those facts change, economics as a subject has to change.

For example, until recently the Philips curve, a formula relating a series of factors in industrial societies seemed to give an accurate description of economic realities in those societies. Well, lately, it hasn’t worked so well. I think most economists believe that this is because it didn’t accurately describe reality. But they might consider: perhaps it did accurately describe reality as it was at that time. Ever after the oil crises and various other events of the Seventies, reality changed. Economics is a systematic formalised science, but it’s not independent of context or free of history. It is grounded in human practices, but those practices are not themselves timeless, eternal or inevitable.

Well, let’s now turn to linguistics. The standard contemporary aim of linguistics is to state the various rules—phonological, syntactic and semantic—that relate sounds and meanings in the various natural languages. An ideally complete science of linguistics would give the complete set of rules for every natural human language. I’m not sure that this is the right goal for linguistics, or even that it is a goal that’s possible of attainment, but for the present purposes the important thing to note is that it is once again an applied science of intentionality. It is not in the least like chemistry or geology. It is concerned with specifying those historically determined intentional contents in the minds of speakers of the various languages that are actually responsible for human linguistic competence. As with economics, the glue that binds linguistics together is human intentionality.
The upshot of this fifth lecture can now be stated quite simply. The radical discontinuity between the social and the natural sciences doesn’t come from the fact that there’s only a disjunctive connection of social and physical phenomena. It doesn’t even come from the fact that social disciplines have constitutive concepts which have no echo in physics or even from the great complexity of social life. Many disciplines, such as geology, biology and meteorology, have these features, but that does not prevent them from being systematic natural sciences. No, the radical discontinuity derives from the intrinsically mental character of social and psychological phenomena.

The fact that the social sciences are powered by the mind is the source of their weakness vis-à-vis the natural sciences. But it’s also precisely the source of their strength as social sciences. What we want from the social sciences and what we get from the social sciences at their best are theories of pure and applied intentionality.