

# Causal perspectivalism\*

Huw Price<sup>†</sup>

March 6, 2005

## I Foreign metaphysics and the benefits of travel

As objects go, foreigners are a pretty respectable bunch. They're not figments of our collective imagination, or social constructions, or useful fictions. They're not mind-dependent, and they don't disappear when we don't keep an eye on them. Our 'folk theory' about foreigners isn't subject to some global error, and the term 'foreigner' certainly manages to refer. Some of our beliefs about foreigners are mistaken, no doubt, but only by failing to accord, case-by-case, with the objective reality to which they are certainly answerable. There are many facts still to be discovered about foreigners, such as their precise distribution in space and time. Moreover, these are matters for scientific study. And so on. In a nutshell, foreigners are as real as we are.

Yet think of the discovery each of us made when, minds broadened by travel, we realised that foreigners themselves use the very same concept, but apply it to us! What we learnt (at that unsettling moment) was that the distinction between them and us, foreigners and compatriots, isn't as objective as we'd assumed. It is a distinction drawn 'from a perspective'—that of a *local* speech community, embedded in a tribal population. There are objective divisions in the world, of course, but not the asymmetric distinction that each side sees from where it stands. God sees us as Afghanis, Zimbabweans, and many things in between, perhaps, but not as locals and foreigners.<sup>1</sup> So, the reality of foreigners notwithstanding, there's a sense in which foreignness is a less objective matter than we used to think.

Perspectivity of this kind raises important philosophical issues. Some are general issues: How is the relevant notion of perspective best characterised? Is it one phenomena or

---

\*This is a draft of a paper to appear in Huw Price and Richard Corry, eds., *Causation, Physics and the Constitution of Reality: Russell's Republic Revisited*, Oxford: Oxford University Press., 2005.

<sup>†</sup>Centre for Time, Department of Philosophy, Main Quad A14, University of Sydney, NSW 2006, Australia. E-mail: huw@mail.usyd.edu.au.

<sup>1</sup>I'm thinking here of the philosophical 'view from nowhere' kind of god, of course. Some gods have regrettably perspectival viewpoints.

several? And how far does it extend—how many of our conceptual categories are perspectival, in whatever sense or senses the notion turns out to encompass? Others are local issues: For some particular concept or group of concepts of philosophical interest, is that concept or group of concepts perspectival, in any interesting sense? And if so, how, and how can we tell? In this paper, I'm interested in one of these local issues: roughly, the question whether our *causal* concepts are perspectival.

I began with the local–foreigner distinction with three points in mind. First, it is a striking and familiar example of the general phenomenon I'm interested in, and illustrates well that noticing perspective is often a matter of learning to see things 'from another point of view'—from the viewpoint of someone who uses the same concept as we do but applies it differently, in virtue of a difference between their circumstances and ours. Sometimes, of course, it can be very hard to make this imaginative shift, especially when the alternative standpoint simply doesn't occur within our own linguistic community—when 'we' all occupy the same viewpoint, in the relevant respect. (Call this *homogeneous* perspective.) I'll argue that causation is like this. For basic physical reasons, all humans share a homogeneous perspective, in the relevant respect.

The second useful lesson is the one I started with. Perspectivity doesn't automatically lead to simple-minded forms of antirealism or subjectivism. It may lead to more sophisticated forms of the same thing, but in virtue of being more sophisticated, these won't challenge the 'obvious' truths of common sense in the same way.

Finally, the example has a useful structural similarity to the case of causation. Why? Because the local–foreigner distinction is strongly asymmetric, a fact variously revealed in the different ways we behave towards foreigners and locals. The example thus provides a case of an *apparent* asymmetry in reality (apparent, anyway, from a sufficiently naive viewpoint) that has turned out to be merely perspectival. I want to argue that in an analogous way, a perspectival view of causation is that it makes better sense than alternatives of both the *asymmetry* of causation, and its *temporal orientation*—that is, of the intuitive difference between cause and effect, and the fact that at least in general, causes precede their effects in time.

Although the local–foreigner distinction is asymmetric, it isn't temporal, and another useful way of approaching the issue of causal perspectivity is via more general questions about the conceptual manifestations of our temporal situation—the temporal aspects of the ways in which we are constructed, situated, embedded and oriented in time. In some respects, this is familiar philosophical territory. Tense, and the distinctions between past, present and future, are widely regarded as products of our own temporal perspective. On this view, clearly, there's an analogy between 'now' and 'then', on the one hand, and 'us' and 'them', on the other. True, this comparison is somewhat controversial. Some writers, such as presentists, maintain that there is an objective *now* in a sense in which there is not an objective *us*. And in any case, the two distinctions are clearly disanalogous in another respect: our location with respect to whatever fence we mark by 'us' and 'them', compatriot

and foreigner, while perhaps not immutable, is very much more constant than the temporal location we mark by ‘now’.

Not all parts of the temporal territory are so inconstant, however. I’m especially interested in the conceptual manifestations of our asymmetric *orientation* in time. In my view, some of our modal notions, including causation itself, have a perspectival character closely linked to the ‘oriented’, or temporally-asymmetric aspects of our constitution, as entities or structures located in spacetime. Clearly, these aspects of our temporal situation are not things that are likely to change, either for each of us individually from time to time, or from one of us to another, across the community. In this respect, then, we humans all share the same temporal perspective (even though, as we’ll see, physics might allow other possibilities). So any conceptual perspectivity grounded on our temporal orientation is likely to be homogeneous, if anything is; and therefore, presumably, very hard to see.

At this point, travel can no longer help us. We need alternative resorts, alternative metaphors. One of the most powerful is Kant’s comparison of the revelation (as he saw it) that there is an anthropocentric ingredient in some of our conceptual categories to the Copernican discovery of the perspectival character of the sun’s apparent motion around the earth. Transposed to the present project, Kant’s analogy makes several important points: that perspective can be hard to see; that its discovery may challenge deeply-entrenched intuitions about the nature of reality; and—perhaps most importantly of all, in the present context, in virtue of the role of causation and related concepts in science—that scientific virtue may lie on the perspectivalist’s side.

As the case of *foreigner* made clear, unmasking the perspectival character of a concept does not lead to simple-minded antirealism—we may continue to use the concept, and even to affirm, in a variety of ways, the objectivity of the subject-matter concerned, despite our new understanding of what is involved (of where we ‘stand’) in doing so. Nevertheless, there is a tendency to think that perspectivity is incompatible with good science, in the sense that science always aims for the perspective-free standpoint, the view from nowhere. In my view, it is important to see that science itself might challenge this philosophical conception of science. For suppose we came to accept a perspectival genealogy for causation and related notions (perhaps for counterfactual reasoning, for example). In essence, this would be a scientific account of a particular aspect of human linguistic and cognitive practice, explaining its origins in terms of certain characteristics of ourselves, as structures embedded in time in a particular way. A corollary would be that uses of these very concepts in science—including, indeed, in this very explanation—would themselves be held to reflect the same embedded perspective. Thus some aspects of current scientific practice would be revealed *by science* to be practices that only ‘make sense’ from this embedded perspective—so that if, *per impossibile*, we could step outside this perspective, these aspects of science would cease to be relevant to us.

Would this be a *reductio* of the perspectival account of causation? Or a fundamental challenge to science? Neither, in my view. On the contrary, it would be continuous with a

great scientific tradition, a tradition in which science deflates the metaphysical pretensions of its practitioners, by revealing yet further ways in which they are unlike gods. For Kant, Copernicus was the giant in this tradition. For us, Darwin towers beside him. The vertiginous lesson we need continually to relearn, on these lofty shoulders, is how insignificant we are, from the world's point of view; how idiosyncratic the standpoint from which we attempt to make sense of it. But however unsettling we ourselves find the blows that science thus delivers to our metaphysical self-image, science itself has not only survived, but thrived, on this diet of self-imposed self-deprecation. I see no reason to think that the present case will be any different, if some of science's own core categories and activities turn out to be perspectival in a newly-recognised way; a way that depends on the peculiar standpoint that science's own practitioners occupy in time.

But would the relevant aspects of current science then stand revealed as bad science? Or would it be the philosopher's 'view from nowhere' conception of science that would have been shown to be mistaken—an inaccurate conception of what science is, or could be? Again, neither, in my view. The perspectivity of (some aspects of) current scientific practice turns out to be entirely appropriate, given its role in the lives of creatures in our situation. In that sense, it is not 'bad science', and doesn't need to be reformed or eliminated. In appreciating this perspectivity, however, we get a new insight into the nature of the non-perspectival world, which 'looks like this', from our particular point of view. So there's good news, too, for 'detached' science.

I'll return briefly to these issues at the end of the paper. In the main, however, I'm going to focus on the first-order question about the character of the causal relation itself. Is the distinction between cause and effect like the distinction between us and them—a perspectival projection onto a non-perspectival reality? Or is it better understood as non-perspectival from the start? I want to make a case for the perspectival view.

My argument relies heavily on analogy. I take the optimistic view that once the key elements relevant to the issue of the status of the causal asymmetry are laid out in a clear way, the attractions of the perspectival view are easy to see. The role of the analogies is to help to achieve the required clarity: 'Look', I want to be able to say, 'The options in the causal case are just like *that*.' With this aim in mind, I'll begin with some simple examples of perspectival and non-perspectival concepts, in use in particular arenas, and then offer a slightly more elaborate example, which has all the structure I take to be needed to provide a useful comparison with the causal case. The early sections of the paper will be devoted to developing this argument for the perspectival character of causal asymmetry.<sup>2</sup>

The latter part of the paper tries to take the analysis a stage further, proposing a kind of diagnosis of the perspectival character of causation. By identifying some key elements

---

<sup>2</sup>As I'll explain, the focus on this asymmetry plays a dual role: it provides an important explanatory puzzle, to which the perspectival view offers the best solution (or so I'll argue); and it supports a powerful 'existence proof', based on symmetry considerations, for the possibility of an alternative causal perspective that reverses the asymmetry.

in those aspects of our epistemic and practical ‘architecture’ that seem essentially associated with causal thinking, I’ll offer an abstract characterisation of what might be called the *causal viewpoint*: a distinctive mix of knowledge, ignorance and practical ability that a creature must apparently exemplify, if it is to be capable of employing causal concepts. My project is thus a kind of naturalised Kantianism about causation. It aims to understand causal notions by investigating the genealogy and preconditions of causal thinking; by asking what general architecture our ancestors must have come to instantiate, in order to view the world in causal terms.

## 2 Perspective on the field of play

Imagine a soccer field, viewed from a goalkeeper’s perspective. There’s a near end of the field and a far end of the field, and the difference makes a great deal of practical difference to the goalkeeper’s role in the game: typically, goalkeepers have much more work to do when the ball is at the near end than when it is at the far end. Clearly, however, this difference is perspectival. The opposing goalkeeper has the same conceptual categories, but applies them with precisely the opposite orientation to the physical arena that separates him from his opponent—and neither, obviously, is objectively right or wrong. (The referee, in the middle of the pitch, can’t settle the issue—from his perspective, the near–far distinction, as applied by the goalkeepers, is simply inapplicable.<sup>3</sup>)

Contrast this to the case in which, literally, the two goalkeepers are not playing on a level field. Suppose that the pitch is higher at one end than the other. This asymmetry, too, makes a practical difference, in various ways. Each goalkeeper needs to adjust his kicks, for example, to allow for the effect of the slope. But here, clearly, the situation is not symmetric: one goalkeeper needs to kick away from his own goal more forcefully than he would if the field were level; the other needs to kick away less forcefully than he would if the field were level. This lack of symmetry reflects the fact that the high end–low end distinction is objective, or non-perspectival—in contrast, therefore, to the near end–far end distinction. (The referee is certainly able to rule on which is the higher end of the pitch.)

In these spatial examples there are four kinds of factor in play:

- (i) The role that certain concepts—*near* and *far*, *high* and *low*, for example—play in some practical game or activity, and the reasoning associated with that activity.
- (ii) Some spatial asymmetries in the application of these concepts.

---

<sup>3</sup>The referee can adopt a kind of second-hand perspectival vocabulary, of course, referring to the two ends of the pitch as ‘the near end for goalkeeper A’ and ‘the near end for goalkeeper B’, for example. Although parasitic on the perspectival concepts employed by the two goalkeepers, however, these descriptions are not themselves perspectival.

- (iii) Some spatial asymmetries in the location or other attributes of the users of these concepts.
- (iv) Some physical attributes of the environment, such as the slope or altitude of the pitch, which are of (possible) relevance to the activity in question, and which may also be distributed asymmetrically in space.

Within these simple parameters, we have already seen how natural it is to regard the spatial asymmetries of some concepts (e.g., *near* and *far*) as reflecting those of the users of the concepts; those of other concepts (e.g., *high* and *low*) as reflecting those of external features of the environment.

The relevance of these examples is that they have much of the structure associated with the case of causal asymmetry. As I'll explain, we can distinguish an analogous set of four factors in the causal case—the main difference being that in that case, of course, time replaces space. My next example is intended to make the analogy even more explicit. It introduces an asymmetric binary relation, defined on pairs of spatial points, which is intended to provide a useful analogue of the cause–effect relation itself. The example is designed so that the spatial relation in question is best understood in a perspectival way; thereby (or at least this is the intention) helping to make it clear how the same might be true of the causal relation.

### 3 The forest of forking paths

Imagine a network of paths, on a forested north-facing hillside. Imagine creatures who map this terrain, and choose trajectories to travel through it. These trajectories depend on the availability of paths, but not only on that, for there is a further major constraint. These creatures begin their journeys (or, as we might say, their lives) at the top of the slope. From that point, things go downhill; for, like rolling stones, these little travellers rely on gravity to make their journey.

In thinking about actual and possible trajectories, these creatures employ a binary relation, defined on pairs of spatial points. They say that point B is *accessible* from point A if there is a possible trajectory that would lead them from A to B. The notion of trajectory is thus directed and transitive, and turns out (let us suppose) to be asymmetric: if A is accessible from B, then B is not accessible from A. Suppose also that at least in most cases in which B is accessible from A, B is to the north of A. We now have several obvious candidates for what it is for B to be accessible from A.

- 3.1 REDUCTION TO SPATIAL DIRECTION. B is accessible from A iff B is connected to A by a suitable path and lies to the north of A.
- 3.2 REDUCTION TO SLOPE. B is accessible from A iff B is connected to A by a suitable path and B is at a lower altitude than A.

- 3.3 REDUCTION TO TYPICAL SLOPE. B is accessible from A iff B is connected to A by a suitable path and lies in the direction from A in which the terrain is *typically* lower.<sup>4</sup>

In addition, we have a perspectival option:

- 3.4 REDUCTION TO THE TRAVELLER'S PERSPECTIVE. B is accessible from A iff a typical traveller arriving at A could proceed to B (at least in principle, perhaps).

Which of these alternatives makes best sense of the use that the creatures in question make of the notion of accessibility?

Let's begin with 3.1. The first issue is whether it gets the extension of the concept right. Is it in fact true, in practice, that a point B is accessible from A only when B is to the north of A? Or do some trajectories run east–west, or even southwards, at least for short distances? If so, then these are counterexamples to 3.1.

Suppose for the moment that 3.1 passes this first test. Let's now think about the question, 'What is it that *explains* the fact that a point B is accessible from A only when B is to the north of A?' In the light of 3.1, the appropriate answer is that this is simply what it *means* for B to be accessible from A. But the puzzle now bursts out somewhere else, of course. If this is what it means for B to be accessible from A, why is it that our creatures can only travel in an 'accessible' direction? In other words, why is accessibility relevant to their practical and cognitive behaviour, in the way that we have assumed that it is?

We're inventing the story as we go along, of course, and so we could simply stipulate that this is how it is. These creatures are simply keyed to rolling north, and that's that. This 'primitivist' strategy seems likely to be unappealing in any realistic case, however, because we don't expect space to play this direct explanatory role, and (what's perhaps the same coin) we can't see any mechanism which could produce such behaviour, absent any other kind of physical property with suitably asymmetric spatial distribution.

In any realistic case, then, we expect that the fact that our creatures travel north will be explained by some spatial asymmetry in their own physical constitution, or that of their environment, or both. The prevailing slope might well provide such a feature, but in that case it seems more accurate to say that accessibility tracks slope, and that it is the further fact that the slope is towards the north, that explains the fact that accessibility is towards the north. 3.1 thus defers to 3.2 or 3.3.

The merits of 3.2 and 3.3 can be assessed by looking for counterexamples. Concerning 3.2, for example, we want to know whether there are cases in which B is accessible from

---

<sup>4</sup>There are two importantly different ways to read this proposal. According to one, accessibility is a fundamentally asymmetric relation, the asymmetry resting on the fact that the relation is constituted, in part, by some relation to the prevailing slope. According to the other, accessibility is essentially a symmetric relation, to which an asymmetric labelling is affixed by convention—a convention that links in practice to the direction of the prevailing slope, though any available 'signpost' would do as well. It is the former reading we have in mind here. The latter could be made explicit, but would then play a role in what follows very similar to that of 3.1, being vulnerable to very similar objections. I ignore it, henceforth, for ease of exposition.

A, without being at a lower altitude than A. Perhaps there are small rises on the generally downhill slope, and some trajectories rise over these hills, before continuing downwards. If so, then this counts against 3.2

Concerning 3.3, similarly, we want to know whether there are cases in which B is accessible from A even though B lies with respect to A in the direction in which the terrain is typically higher, rather than lower—i.e., as it happens, south rather than north of A; or, less dramatically, but also in tension with 3.3, whether there are cases in which B is accessible from A even though B is merely east or west of A. Either kind of case counts against 3.3.

Let's add some detail to the story, to settle the issue. Suppose, as above, that our creatures are rolling stones—freewheeling but intelligent rocks, who rely on gravity and momentum for their journey through this forest of forking paths. It is now easy to imagine how, even though the prevailing direction of travel must be downhill (and therefore, in the assumed environment, to the north), small variations of two kinds are possible. Our rocks can roll uphill for short distances, thanks to their momentum. And they might sometimes roll southwards for short distances, either downhill, when small variations in the prevailing slope permit it; or uphill, again, when they are able to convert kinetic to potential energy. (Perhaps they achieve this trick by first tackling a small uphill rise, and then doubling back—imagine the options available to a snowboarder!)

It now seems clear that 3.4 is closest to the truth. Accessibility is partly 'relative to the perspective of the rolling stone'. It depends not only on the external environment, but also on properties of the stone: its speed, mass and direction of travel (i.e., its momentum). There are various ways in which this might be formalised. Momentum might figure in a truth-condition, an assertibility condition, or perhaps other alternatives. This level of detail is unnecessary, however. The crucial point is simply that any satisfactory elaboration will be perspectival (or 'lithocentric'), in taking accessibility to depend on contingent attributes of the stones in question, as well as on the properties of the environment.

Accessibility is intended to be a prospective notion. Our intelligent rocks are planners, who think in advance about where their journeys might take them. What they want to know, when they consider whether B is accessible from A, is whether, if their journey takes them to A—if they arrive at A in typical fashion—they can expect or hope to reach B. In virtue of the general orientation of their hillside, and their own reliance on its slope for their locomotion, a typical way of arriving at A is from the south. Absent any special circumstances in which they are able to progress in a curving path back towards the south, B will be accessible from A only if it lies to the north, or at least not to the south.

Accessibility thus has a prevailing northerly direction, in virtue of the typical motion assumed in thinking about new cases. The explanation of the typical motion appeals to the typical slope, and this in turn explains why accessibility tends to be aligned with the downhill slope. But if the creatures have no trouble, in theory or in practice, in handling occasional small sections of level or even uphill path, then this correlation between ac-



cessibility and slope will often break down at a local level (with or without concomitant exceptions to accessibility's prevailing general northerly orientation). It may even break down in dramatic ways, at least in theory, if our rolling stones, idealising beyond their real physical limitations, can imagine arriving with lots of momentum at the base of substantial uphill sections of path. In those cases, despite the dramatic departure from the normal downhill gradient, the uphill sections will seem accessible, at least in principle.

Imagine that some of the more far-sighted rocks, deducing on theoretical grounds that no downhill slope can continue for ever, realise that their own hillside will eventually level out. Perhaps the terrain stays flat after that, or perhaps their world is more symmetric, and their own downhill slope is matched by an uphill slope on the far side of a valley. In either case, unlikely though it might be that any actual rock should roll that far, a sufficiently idealised notion of accessibility might well extend to such regions. If so, the accessibility 'arrow' would still be taken to point predominantly to the north, even in these regions which lack the usual gradient in that direction.

Note the contrast between two ways in which the prevailing slope can be relevant. For the objectivist options, 3.2 and 3.3, slope comprises part of what is effectively a truth-condition for ascriptions of accessibility. For the perspectival option, it plays a very different role: it is something like a necessary physical precondition for the existence of creatures equipped to *occupy* the perspective in question. Roughly, their perspective depends on the fact that they are in motion, and (in the circumstances as described, at least) their motion depends on the slope. There is a big difference between these two ways of characterising the relevance of the prevailing slope; even though it only shows up, roughly speaking, where the slope gives out. If slope is part of a truth or content condition, then where the slope gives out, so too does the extension of the concept in question. If it is a precondition for occupation of a perspective, then even the most level terrain may be viewed asymmetrically, provided only that there is enough slope, somewhere, to produce the kind of asymmetric creature who can occupy that perspective.

Thus, as our thoughtful little rocks consider their intuitions about how the notion of accessibility is properly applied, in the light of an understanding of their own nature and circumstances, they are already in a position to see that the notion is perspectival; that its application depends on an important spatial asymmetry in their own condition. Moreover, it seems to me, they are able to reach this conclusion without having the further thought that there might be creatures in which a similar spatial asymmetry had a different orientation—creatures who would hence apply the same notion with a different extension.

Nevertheless, this further thought, when it is achieved, acts as a virtual trump card. Realising that they might live on one side of a valley, our creatures suddenly recognise that there might be rocks on the opposite side of the valley ('rock\*s', perhaps) who would see things in precisely the way that they themselves do, but with the opposite spatial orientation. At this point, our creatures cross the kind of conceptual frontier that we ourselves crossed, when we realised that other people see us as foreigners—and see the most powerful

reason for regarding accessibility as perspectival.

True, a proponent of 3.2 or 3.3 might argue that this case shows simply that the direction of accessibility varies from place to place—just as that of *downhill* does, for example. But we’ve already seen why this isn’t a serious rival to the perspectival proposal. It doesn’t explain as well as the perspectival view the postulated extension of the notion of accessibility to cases lacking the actual or typical slope in question (or, indeed, why slope matters in the way that it does to the practical activities of the creatures in question). And in any case, if the perspectivalist’s opponent is the kind of realist who wants to regard the arrow of accessibility as a deep structural feature of reality, these reductionist alternatives provide little real comfort. So long as the symmetries of the relevant physics permit reversal of whatever asymmetry is taken to provide the reductive ground of accessibility, its arrow becomes at best contingently a global constant; and risks extinction altogether, in regions without asymmetry in question, such as the floor of the valley that separates rocks from rock\*s. This ‘reversibility objection’ remains very powerful, then, even for an opponent blind to the advantages of perspectivalism.

#### 4 Applying these lessons to the causal asymmetry

In the previous example, we had four kinds of factor in play: a certain concept, *accessibility*, playing a central role in a practical activity; a marked spatial asymmetry in the application of this concept; a spatial asymmetry in the physical characteristics of the users of the concept in question; and some spatially-asymmetric attributes of the environment, of relevance to the practical activity in question. As we saw, the key issue was whether the spatial asymmetry of the relation of accessibility should be regarded as analytic (3.1), as reducible to the environmental asymmetry either directly (3.2) or on average (3.3), or as grounded on the ‘intrinsic’ asymmetry of the users of the concept (3.4). The example was chosen so that the perspectival option made best sense of the use of the concept in question. And the focus on spatial asymmetry played a dual role: it provided an important *explanandum*, to which the perspectival option offers the best *explanans*; but also the basis of a kind of existence proof, based on symmetry considerations, for an alternative perspective that reverses the asymmetry in question.

In the causal case, similarly, we have four factors to consider:

- (i) The concepts *cause* and *effect*, with important conceptual links to certain practical activities, viz., intentional action and deliberation. In particular, an action and its intended outcome are held to be related as cause and effect: means and end are cause and effect.
- (ii) Some temporal asymmetries in the application of these concepts, such as the fact that causes typically *precede* their effects.

- (iii) Some time-asymmetric characteristics of the users of the concepts in question, such as the fact that they typically deliberate about *future* actions, on the basis, at least in part, of information received in the *past*.
- (iv) Some temporal asymmetries in the environment, such as the prevailing thermodynamic asymmetry.<sup>5</sup>

Again, the issue is how best to explain the temporal asymmetry in the application of the concepts in question, and to relate it satisfactorily to the ‘internal’ and environmental asymmetries of (iii) and (iv). And again, there seem to be four main options:

- 4.1 REDUCTION TO TEMPORAL DIRECTION. B is an effect of A iff B is causally connected to A and occurs later than A.
- 4.2 REDUCTION TO THERMODYNAMIC GRADIENT. B is an effect of A iff B is causally connected to A and B is at a higher entropy than A.
- 4.3 REDUCTION TO TYPICAL THERMODYNAMIC GRADIENT. B is an effect of A iff B is causally connected to A and B lies in the temporal direction from A in which entropy *typically* increases.<sup>6</sup>
- 4.4 REDUCTION TO THE AGENT’S PERSPECTIVE. B is an effect of A iff doing A is a means of bringing about B, from an agent’s perspective—roughly, if controlling A is a means of controlling B.

Here’s a sketch of a case for preferring 4.4, along the lines of the case we made earlier for preferring 3.4. Again, let’s begin with 4.1. The first issue is whether it gets the extension of the concept right. Is it in fact true, in practice, that an event B is an effect of an event A only when B occurs later than A? Or is causation sometimes simultaneous, or even backwards? If so, then these are counterexamples to 4.1.<sup>7</sup>

Suppose for the moment that 4.1 passes this first test. Let’s now think about the question, ‘What is it that *explains* the fact that an effect always occurs later than its cause?’ In the light of 4.1, the appropriate answer is that this is simply a consequence of what

---

<sup>5</sup>For present purposes I’ll assume that the thermodynamic asymmetry, broadly construed, is the only time-asymmetric feature of the physical world that might plausibly be held to be relevant at this point. Thus I rule out, for example, the possibility that the asymmetry of causation might have something to do with the small T-symmetry violations known in microphysics. This assumption doesn’t seem to me to be controversial, but in any case, much of the argument below would easily transpose to a proposal of the latter kind.

<sup>6</sup>As in the case of 3.3, there are two possible readings of this proposal. I ignore the reading according to which the thermodynamic asymmetry is merely a convenient temporal ‘signpost’, by means of which to label the two ends of a time-symmetric relation.

<sup>7</sup>This is a familiar objection to Humean conventionalism about the direction of causation, according to which it is merely a terminological matter that causes precede their effects.

it *means* for one event to be an effect of another. But if this is true, why, then, is the distinction between cause and effect relevant to human action, in the way noted in (i)? Why are the *ends* of our actions typically *effects* of the *means* by which we achieve them?

If there is to be a satisfactory answer to this question, in the light of 4.1, it must appeal to the principle that the means–end relation itself has the appropriate temporal orientation: ends must occur later than their corresponding means. But this in itself, now, is something that calls for explanation—and it is no use simply appealing again to linguistic convention. It can't be convention all the way down. There is a genuine temporal asymmetry in our deliberative practice—viz., roughly, the fact that our deliberations are future-directed but not past-directed—and this calls for explanation. In the end, then, this temporal asymmetry needs to be explained in terms of some temporal asymmetry in our own physical constitution, or in our environment, or some combination of the two.

As in the spatial case, it is possible, at least in theory, that the relevant environmental asymmetry might be a primitive asymmetry of time itself, on which our deliberative behaviour somehow depends. In other words, it might be held that there is an intrinsic direction to time—an intrinsic distinction between past and future—independent of, or at least more basic than, any other physical time-asymmetry; and that it is somehow a necessary fact that deliberation is sensitive to this fundamental directionality, again in a manner unmediated by other physical time-asymmetries.

This primitivist view will perhaps attract more support than the corresponding view in the spatial case. Certainly, some philosophers profess to believe that time has some such intrinsic directionality—and at least in some cases, seem to think that their grounds for believing it are direct, or phenomenological, of a kind which is insensitive to the distribution and possible reversibility of the known physical time-asymmetries (such as that of thermodynamics). For present purposes, I simply want to note the theoretical cost of this view. It requires a deep link between the mental, on the one hand, and some deep and fundamental time-asymmetric aspect of physical reality, on the other—without the time-asymmetry concerned being manifest at intermediate levels!

Similar remarks would apply to attempts to appeal to a primitive asymmetry of causation itself, as a brute metaphysical fact. Again, it would be mysterious how such a fact could have the relevance it needs to have for deliberation, if the connection isn't somehow mediated by asymmetries in us or our environment. Otherwise, once again, we need some peculiar primitive link between minds and fundamental metaphysics.

If we set these primitivist proposals aside, then the time-asymmetry of deliberation needs to be explained in terms of some other temporal asymmetry, in our own physical constitution, in our environment, or in some combination of the two. The thermodynamic time-asymmetry might provide such a feature, for example, but in this case it seems appropriate to say that the direction of deliberation follows that of increasing entropy, and

that it is the fact that entropy increases towards (what we call)<sup>8</sup> the future that explains the fact that we deliberate ‘in that direction’. Thus 4.1 defers to 4.2 or 4.3.

On the face of it, just as in the spatial case, the plausibility of 4.2 and 4.3 can be decided by looking for counterexamples to one or other suggestion. Concerning 4.2, for example, we want to know whether there are cases in which B is an effect of A, without being at a higher entropy than A. If so, then this counts against 4.2

Concerning 4.3, similarly, we want to know whether there are cases in which B is an effect of A even though B is oriented with respect to A in the direction in which entropy is typically lower, rather than higher—i.e., earlier rather than later than A; or, less dramatically, but also in tension with 4.3, whether there are cases in which B is an effect of A even though B is merely simultaneous with A. Either kind of case counts against 4.3.

My own view is that we have no trouble making sense of simultaneous causation; that we have little trouble in making sense of ‘entropy-reducing’<sup>9</sup> and ‘backward’ causation<sup>10</sup>; and, as I’ll argue in a moment, that we certainly have no trouble in making sense of ordinary, directed causation in cases in which—in virtue of the manifest time-symmetry of the physics in question—it is simply implausible that typical slope is important (except in so far as it supports our perspective). If I’m right, then as for the rolling stones, only the perspectival view gets the intuitive extension of the concepts in question more or less right. According to the perspectival view, causal reasoning too depends on the standpoint of creatures engaged in a certain kind of journey. In this case, it is a journey into an uncertain future, a journey in which—from the epistemic standpoint of the creatures themselves<sup>11</sup>—their choices determine what path they take through a tree of forking possibilities.

I want to develop one example, in order to illustrate the very striking advantages of the perspectival view, and also to motivate an hypothesis about the genealogy of a crucial element of the asymmetry of our causal reasoning.

---

<sup>8</sup>The qualification is needed because ‘past’ and ‘future’ may well themselves be perspectival notions. More on this below.

<sup>9</sup>Here’s a sketch of an argument to this effect. Consider a typical entropy-increasing process, such as gas molecules escaping from a bottle. Gradually reduce the number of particles. Presumably this doesn’t change the essential causal structure of the case, and only changes by degree the extent to which it is an entropy-increasing process (however this is characterised). But reduce the number far enough, and you approach a case we could produce in reverse in a laboratory. In the reverse case—so our intuitions tell us—the causal arrow would still run from past to future; and hence, now, in the direction in which the entropy of the local system in question *decreases*.

<sup>10</sup>I’ll mention some examples in a moment.

<sup>11</sup>This qualification is indispensable, in my view, and marks the second major respect in which causal reasoning is perspectival. More on this below.

## 5 Stargate Doughnut

Imagine a photon,  $p$ , which spends billions of years in intergalactic space as it travels from one distant galaxy,  $G_{past}$ , to another,  $G_{future}$ . At some point in between, at a time  $t$ , it passes through the central aperture in a tiny doughnut-shaped object, which happens to be spinning on a transverse axis, somewhere in deep space. As this doughnut spins, it periodically occludes the path the photon takes from  $G_{past}$  to  $G_{future}$ , and hence acts as what we could call a ‘stargate’. At  $t$ , however, the gate is open, and the path is unobstructed.

Consider the following counterfactuals:

**Proposition 1** If Stargate Doughnut had been closed at time  $t$ , the photon  $p$  would not have been absorbed at  $G_{future}$ .

**Proposition 2** If Stargate Doughnut had been closed at time  $t$ , the photon  $p$  would not have been emitted at  $G_{past}$ .

We take Proposition 1 to be true, and Proposition 2 to be false. Accordingly, we take the orientation of the stargate to be a cause of the *later* position of the photon, but not the *earlier* position of the photon. What is the source of this time-asymmetry?<sup>12</sup>

On the face of it, the example seems completely independent of any thermodynamic details of the systems in question. It is far from clear that any sense can be made of the notion of the entropy of a discrete microscopic system of this kind. Even if it could, however, the example doesn’t depend on the existence of a thermodynamic gradient. After all, we could quite well imagine the same kind of situation, in a universe in thermodynamic equilibrium. (We’d need to replace galaxies with some photon source at the same temperature as everything else, but this makes no significant difference to the example.)

In some sense, of course, our intuitions about the case are sensitive to temporal direction. What else could we rely on in distinguishing Proposition 1 from Proposition 2, after all, given their apparent symmetry in other respects? But as we saw, however, it is no use trying to rely simply on temporal direction, to explain the difference between the two cases. If we say it is merely a conventional matter—a matter of the meaning of ‘cause’ and ‘effect’—that causes occur before their effects, this will certainly imply that Proposition 2 (or a causal variant of it) must be false; but it doesn’t explain why that is relevant to our decision behaviour, in the way that it is. (We think that we could use the stargate to influence future photon positions but not past photon positions, for example.)

---

<sup>12</sup>There is another intuitive asymmetry lurking here, related to that of Kripkean necessity of origins: we are somewhat more inclined to allow that it is the *same* photon we consider in Proposition 1 than in Proposition 2, viz., the one that actually passes through the stargate at  $t$ . I don’t have space to explore the origins of this second asymmetry here, but I regard it as a very plausible target for the perspectival approach, with likely connections to the causal and counterfactual cases.

The perspectival proposal is that what we bring to the case, in imagination, is the typical perspective *we* have as deliberating agents—the perspective we bring to the situation, quite unconsciously, when we think about *manipulating* the stargate. This deliberative perspective displays a very marked temporal bias, of course. Roughly, it treats the past as fixed, and only the future (or some subset of the future) as under our control. According to the perspectival view, it is this asymmetric perspective *on our part* that grounds the intuitive asymmetry between Proposition 1 and Proposition 2.

It might be objected that in introducing ‘us’ in this way, we implicitly ‘import’ the thermodynamic gradient, on which our existence depends—thus contradicting my claim that the case is insensitive to the thermodynamic details. However, there is no such conflict, as the spatial analogy developed earlier makes clear. In that case, the prevailing slope was relevant in the sense that without it, there could be no rolling stones of the kind described, and hence no creatures from whose perspective the notion of accessibility would make sense. As we saw, however, this is very different from reduction of the relation of accessibility to any fact about the prevailing slope.<sup>13</sup>

In the rolling rocks case, I suggested that in making judgements about accessibility, the creatures consider the issue of the accessibility of a point B from a point A, under the assumption that they arrive at A in a typical manner—i.e., in their case, from the south, with a certain amount of momentum. We noted that momentum is easily varied in imagination, as it were, supporting a well-defined sense of ‘accessible in principle’. Accessibility assessed in this manner thus comes to transcend the capacities of any actual creature.

What is the corresponding exercise in the causal case? In my view, we get an excellent explanation of our intuitions about the stargate example if we postulate, again, that what underlies those intuitions is an exercise of imagination: we imagine *intervening*, to change the orientation of the doughnut, and assess the likely outcomes of such interventions. Roughly, anything that changes or ‘wiggles’ in these imagined cases, when the orientation of the doughnut is ‘wiggled’, is regarded as an *effect* of that orientation.

Of course, it is hardly news that some such notion of intervention plays a crucial role in our understanding of causation. Manipulability approaches to causation have a long history, and the importance of a notion of intervention to an understanding of causal reasoning has been displayed in a powerful formal way in recent years by writers such as Glymour, Pearl and Woodward.<sup>14</sup> By and large, however, these authors think of their project in objectivist terms. In other words, they think of the causal structure of the world—including interventions themselves, to the extent that they comprise a crucial element of that structure—as something that exists independently of human agents. But there is a deep tension in such a viewpoint, in my view, stemming from the fact that intervention is

---

<sup>13</sup>Compare: we need brains to be in a position to talk about anything at all, but it doesn’t follow that in talking about anything at all, we are talking *about* our brains.

<sup>14</sup>See especially Spirtes, Glymour and Scheines (1993), Meek and Glymour (1994), Hausman (1998), Pearl (2000), and Woodward (1997, 2000, 2003). There is also an excellent introduction in Woodward (2001).

a deeply perspectival notion. Intervention acts as a kind of Trojan Horse against objectivist approaches: when (rightly) accorded a central role, it thwarts the metaphysical ambitions of the theories in question.

Why is intervention a Trojan Horse? Because of the in-built bias I mentioned ago—the bias manifest, *in our case*, in the fact that when we imagine intervening, we carve up the relevant aspects of reality, on broadly temporal lines, into a fixed or ‘given’ past and an open or mutable future. This carve-up, I claim, is perspectival: it reflects contingent features of our own circumstances, in such a way that other thinkers, differently ‘situated’ in the relevant respects, would carve matters up in a different way. In their hands, then, the same conceptual framework acquires a different extension, in the manner characteristic of perspectival concepts.

We can’t simply take for granted that the carve-up in question is perspectival, of course; the point needs to be argued. So far, the stargate example is simply an intuition pump. It makes it clear that we retain strongly time-asymmetric intuitions about the possibilities for intervention, and hence about causal direction itself, even where the ‘objective’ physics of the situation is manifestly symmetric. I’m going to reinforce this intuition pumping in two ways. First, in §6, I’ll present what seems to me the most powerful argument in favour of the perspectival view. As in the rolling rocks case, it exploits symmetry considerations, to argue for the possibility of creatures with an alternative perspective on the same objective reality—roughly, creatures whose carve-up holds fixed what we hold open, and vice versa. (As I’ll explain, the relevant sense of ‘possibility’ is rather thin: the argument is really a further, very forceful, intuition pump.)

The second kind of reinforcement, in §7, is both more concrete and more general. I’ll aim for an abstract characterisation of deliberation—a characterisation that abstracts away from various contingencies about the kinds of agents we ourselves happen to be—in order to identify the contingency in question in the most general possible terms. This will demonstrate, I think, that the notion of intervention is ineliminably perspectival, and also provide a deeper understanding of the origins of our intuitions about the stargate case. I want to show that our intuitions about the case are just what they ought to be, if they are indeed manifestations of the viewpoint of an agent in the general sense, who (to put it somewhat figuratively) happens to be embedded in time in a particular way.

## 6 The reversibility argument: causation in a Gold universe

In arguing that the perspectival view makes better sense than the alternatives of our intuitions about the direction of causation, I’ve so far avoided placing any weight on the suggestion that there might be creatures who use causal concepts, but whose perspective does not coincide with ours. As I have presented it, then, the argument would be unaffected by the discovery that there were good physical reasons to exclude this possibility.



As we'll see, however, the dialectic is far from even-handed at this point. The reversibility argument, exploiting physical time-symmetries to make a case for the possibility of alternative causal perspectives, doesn't need to wait for physics to convince us that these perspectives are actually occupied. As in the rolling stones case, in my view, the reversibility argument therefore provides something close to a trump card for perspectivalism. Yet it is a card that remains strangely neglected, and the role of the following brief excursion into physics is to bring it into play.<sup>15</sup>

As I have already emphasised, the main candidate for a physical asymmetry that seems likely to be associated with the causal asymmetry, whether by the reductive or perspectival routes, is the asymmetry associated with the second law of thermodynamics—the general tendency of entropy always to increase. The behaviour the second law describes is strongly time-asymmetric, of course. As we might put it, entropy increases in one temporal direction ('towards the future'), but decreases in the other ('towards the past'). Since the late nineteenth century, physicists have puzzled about the origins of this time-asymmetry, especially in the light of the apparent time-symmetry of the underlying laws of physics.

One simple but crucial insight is that entropy would not be increasing in our region, if it were not for the fact that entropy was low, at some point in the past. After all, if entropy had had its maximum possible value in the past, the effect of the second law would simply have been to keep it at that maximum—there would be no general entropy increase. In the last forty years, modern cosmology has thrown some remarkable light on the nature of this low entropy 'boundary condition' in the past. It now appears that matter was extremely smoothly distributed, about 100,000 years after the Big Bang. In a system in which the dominant force is the attractive force of gravity, this is a highly ordered or low entropy state: the 'natural', equilibrium-seeking behaviour of gravitating matter is to agglomerate into large, inhomogeneous clumps; and the tendency of matter to do this, coupled with the homogeneous initial state, drives the production of galaxies and stars.

The upshot seems to be that the conditions required for evolution of creatures such as ourselves depend on this cosmological low entropy boundary condition in the past. Moreover, it seems plausible that it is no accident that we are aligned in the way that we are with respect to the entropy gradient provided by the existence of this condition—that any creature is bound to 'remember' in the direction in which entropy decreases, and hence to regard that direction as the 'past'.

In fact, the suggestion that 'past' and 'future' are perspectival in this way was made already by Boltzmann in the 1890s—though against the background of a rather different proposal about why entropy is low in what we call the past. Following a proposal by his assistant, Schuetz, Boltzmann suggested that we might live in the kind of chance fluctuation from equilibrium which, although extremely rare, would nevertheless be inevitable

---

<sup>15</sup>For a more extended introduction to the topic, and recommendations for further reading, see Price (2004a).

eventually, in a sufficiently ancient universe. Such fluctuations would be two-sided dips in the entropy curve, of course, and Boltzmann suggested that it is no accident that we seem to live on an ‘uphill’ rather than a ‘downhill’ side of the dip—i.e., a side on which entropy is increasing, rather than decreasing. The sense of uphill or downhill depends on our sense of the direction of time, and that, Boltzmann suggested, is perspectival.

Thus Boltzmann’s proposal already introduces the possibility of creatures whose temporal perspective (and hence also, presumably, causal perspective) is reversed relative to ours. However, this is not the most congenial form of this idea, because Boltzmann’s hypothesis has some unwelcome consequences. It turns out that if Boltzmann’s statistics are our guide, it is much easier for random fluctuations of this kind to produce fake records and memories, than the real events of which they purport to be records.<sup>16</sup> As a result, Boltzmann’s hypothesis about the origins of the low entropy past implies that, almost certainly, all our ‘records’ and ‘memories’ are misleading. So the proposal is hard to take seriously.

However, the possibility of creatures whose temporal perspective is reversed relative to ours emerges again in a new form, in the modern cosmological alternative to Boltzmann’s account of the origins of the low entropy past. One of the early contributors was the cosmologist, Thomas Gold. Among other things, Gold (1962) proposed that there might be a deep connection between the second law of thermodynamics and the (then) newly-discovered expansion of the universe. Roughly, Gold’s idea was that the expansion of the universe creates new possibilities for matter (rather like a piston withdrawing in a cylinder of gas, which makes new positions accessible to the molecules of gas within the cylinder). Gold proposed that the increase of entropy simply reflected the tendency of matter to take up these new possibilities, just as the gas expands behind the retreating piston. Moreover, Gold appreciated the symmetric implications of this proposal: it implies that if the universe were eventually to recollapse to a ‘Big Crunch’, entropy would need to decrease. After all, a recollapse is simply an expansion viewed in reverse. Hence if expansion implies entropy *increase*, then recollapse implies entropy *decrease*.

Gold’s suggestion seems to be flawed, however, in a way that comes to light if we make a *modus tollens* of his *modus ponens*. It doesn’t seem to be the case that recollapse automatically leads to decreasing entropy. Why should the gravitational clumps in recollapsing matter gradually disperse, after all, rather than simply becoming even larger? But if recollapse doesn’t imply that entropy decreases, then expansion doesn’t imply that entropy

---

<sup>16</sup>To give an example I’ve used elsewhere, suppose that God wants to leaf through possible worlds, until he finds the complete works of Shakespeare, in all their contemporary editions. According to Boltzmann’s own statistical measure, it is vastly more likely that God will hit upon a world in which the texts occur as a spontaneous fluctuation of modern molecules, than that He’ll find them produced by Shakespeare himself. This is simply because entropy is much higher now than it was in the sixteenth century. According to Boltzmann, probability increases exponentially with entropy. So the higher-entropy twenty-first century—‘Shakespearean’ texts and all—is much more likely than lower-entropy sixteenth century: almost all possible worlds that include the former don’t include the latter. In Boltzmann’s terms, then, it is extremely unlikely that Shakespeare and his contemporaries ever existed. And the same goes for the rest of history!

increases (again, it is a simple symmetry argument).

So something else is required to explain why entropy increases: a low entropy ‘boundary condition’ in the past, of the kind we’ve already described (and cosmology has apparently discovered). But the role of this condition means that one aspect of Gold’s proposal is in a sense reopened. Would entropy decrease, if the universe recollapses? Well, it depends. It would do so if there is a ‘boundary condition’ near the Big Crunch of the same kind as the one we’ve discovered near the Big Bang—a condition that ensures that entropy is low in that region.

In this sense, then, Gold’s symmetric cosmological model—the ‘Gold universe’, as it has come to be called—remains a live possibility, at least until we have a better sense of how to account for the low entropy condition in the past. And at this point, symmetry comes to Gold’s aid. We can’t dismiss the possibility of a Gold universe on the grounds that it would require an incredibly unlikely ‘fine-tuning’ to make entropy decrease. For, at least in the absence of any time-asymmetry in the underlying physics, the fine-tuning required is the same in either temporal direction. It doesn’t exclude a low entropy past, so it can’t exclude a low entropy future.

At present, then, it remains a live empirical possibility that the universe contains regions in which the thermodynamic gradient is reversed. In such regions, it seems likely that intelligent creatures would have a time-sense reversed relative to ours. There isn’t an objective matter which side—us or them—gets it right about the direction of time, or about the direction of causation. They deliberate with respect to what we regard as earlier events. For them, the causal arrow runs directly counter to the way it runs for us.

This possibility seems to me to make a very strong case for the perspectivity of the causal asymmetry. Moreover, the strength of the case seems to depend very little on how seriously we need to take the hypothesis that our own universe might be a Gold universe. The hypothesis simply gives us an easy way to imagine the possibility that there might be creatures, elsewhere in the actual universe, whose time-sense is the mirror-image of ours. Suppose we grant that if there were such creatures, of whatever origins, then two things would follow: (i) they would think that the causal arrow is oriented in the direction that we would call future-to-past; and (ii) their perspective would be as valid as ours. Then we have all it takes to establish that causal direction is perspectival *for us*—whether they exist or not!

Point (ii) might conceivably be challenged. It might be claimed that if there were such time-reversed creatures, their reversed causal perspective would simply ‘get it wrong’. But what could it be that these creatures are thus supposed to be wrong *about*, exactly, if not something grounded in the physics (which is symmetric, by assumption)? This response seems committed to a version of causal primitivism, in other words, with the attendant disadvantages. Whatever this primitive causal relation is supposed to be, how could we

possibly know which side—us, or our time-reversed cousins—is right about its direction?<sup>17</sup> And in any case, why should it matter? Why should such a relation have any particular connection to deliberation (unlike what we might call ‘quasi-causation’, which would be the perspectival substitute, grounded on deliberation from the start)?

Of course, one might respond to the appeal to the Gold universe by trying to defend a non-perspectival alternative, such as 4.3 or 4.4. Such a view would imply, as a thoroughly objective matter, that causation simply changes direction in a Gold universe, in regions in which the entropic arrow reverses. As we’ve already seen, however, these objectivist approaches make poor sense of our intuitions about a range of cases—the stargate case, for example, and apparent cases of entropy-reversing and backward causation—and certainly do less well than the perspectival view in the task of explaining the links between causation and deliberation. And in any case, to echo a point I made in the rolling rocks case, such objectivist approaches seem to concede a large part of the game, in this context, by conceding that the direction of causation is at best contingently a global constant.

The force of the appeal to the Gold universe lies in the fact that it exploits a symmetry associated with time, in order to take us in thought (so to speak) where travel cannot take us in fact. Beginning with the conditions on which our own perspective depends, a simple temporal reflection yields conditions that support a conflicting perspective on the same reality. But although this is a powerful way of exhibiting the perspectival character of causal thinking, it doesn’t get to the heart of the matter. *Why* does temporal reflection make such a big difference to our causal intuitions? I’ve already suggested that the answer lies in the links between causal reasoning and deliberation. To support this suggestion, I now want to make some general remarks about what is involved in deliberating, or ‘being an agent’. Among other things, I want to try to identify some more basic sources of the perspectivity of the causal viewpoint—sources that seem only contingently linked to our temporal perspective, in the way that the reversibility argument exploits so successfully.

## 7 The architecture of deliberation

In considering the nature of the agent’s perspective, we should distinguish two projects. The first aims for an abstract characterisation of the structure, or functional architecture, of deliberation in general—what is essential to anything that deserves to count as an agent. The second considers the temporal aspects of that architecture, in so far as it is instantiated in us. Thus we need to distinguish ‘essential’ features of deliberation from contingent facts

---

<sup>17</sup>Another possibility, suggested to me by some recent work by Tim Maudlin, would be to argue that these time-reversed ‘creatures’ could not really be minds or agents at all. In raising this possibility, however, Maudlin appeals to the intuition that such time-reversed states would not enjoy the causal relations to one another that the corresponding states in our brains enjoy. In the present context, then, this proposal would have to turn on the kind of primitivism just mentioned. It couldn’t provide a non-question-begging argument *against* the symmetric view.

about the temporal characteristics of deliberation in our case. In principle there might be non-human agents, who—while instantiating a broadly similar functional architecture to us, *qua* deliberators—do so in a way that involves a different relation from ours to time and perhaps to space.<sup>18</sup>

Let's begin, then, with some general remarks about the structure of deliberation. In any deliberative process, presumably, there must be a range of things that the deliberator in question takes to be matters for deliberation: in other words, the alternatives among which she takes herself to be deliberating. For formal convenience, let's regard these alternatives as a class of propositions, denoted by *OPTIONS*. These are the propositions the agent takes herself to have the option of 'deciding to make true', in other words. It will be helpful to subdivide this class into *DIRECT OPTIONS*, comprising those matters over which an agent takes herself to have immediate control, and *INDIRECT OPTIONS*, comprising those ends she takes herself to be able to accomplish indirectly, by an appropriate choice from her *DIRECT OPTIONS*. And let the *FIXTURES* denote everything else—all matters of fact that are not held to be a matter of choice in the deliberation in question.

*FIXTURES* will contain a subset, *KNOWNs*, comprising those facts the deliberator takes herself to *know* at the time of deliberation, and also a larger subset, *KNOWABLEs*, comprising matters she regards as either known or knowable, at least in principle, before she makes her choice.<sup>19</sup> Why must *KNOWNs* and *KNOWABLEs* be subsets of *FIXTUREs*? Because it seems incoherent to treat something both as an input available to the deliberative process, at least in principle, and as something that can be decided by that process. Control trumps a claim to knowledge: I can't take myself to know that *P*, in circumstances in which I take myself to be able to decide whether *P*, in advance of that very decision.<sup>20</sup>

So far, then, we have a very simple template, characterising a deliberator's view of the world. In terms of this template, acting, or intervening, is a matter of fixing something not already fixed—of moving something from *OPTIONS* to *FIXTUREs*, as it were. This deliberative template is linked to, but importantly distinct from, an even simpler epistemic template, that divides the world into *KNOWNs*, *KNOWABLEs*, and everything else. The link turns on the fact that *KNOWNs* and *KNOWABLEs* seem to provide a (perhaps *the*) major constraint on what goes into *FIXTUREs*. As we've just seen, what is accessible to us cannot be

---

<sup>18</sup> Compare D. C. Williams's (1951) suggestion that there might be consciousnesses spread out across space, their successive mental states related like the palings in a picket fence.

<sup>19</sup> Note that 'knowable' is ambiguous in this context. I can believe that something not known could be known, but in a way which would then deprive me of the relevant choice. For example, I believe I can decide where my garden furniture will be placed tomorrow; but I also acknowledge that I could come to know the truth about this matter, prior to any decision on my part, say if I were to discover that someone had concreted the furniture in place, in its present location. The sense of knowability I have in mind here is not this one, however, but that of knowability without loss of control in the matter in question.

<sup>20</sup> Choice is a contextual defeater for a claim of knowledge, as we might put it. Note that the bilking argument against backward causation relies on this conflict, in the case in which the outcomes of an action may lie in the past.

something we can take ourselves to control. Yet, as we also noted, there is much flexibility: something knowable and hence in **FIXTURES** under some circumstances, may nevertheless be regarded as controllable and hence in **OPTIONS** in other circumstances.

In practice, for us, this seems especially true of states of affairs in the future. We plan under certain assumptions about what the future will be like, which we take as **KNOWNs**—e.g., normally, that the sun will rise tomorrow. But this seems to be very context-sensitive: if we want to consider an action that involves eliminating the sun, we won't take the fact that it will rise tomorrow as a given—its rising will be in **OPTIONS**, not **FIXTURES**.

Let's think some more about the temporal characteristics of **OPTIONS** and **FIXTURES**. We've already seen that the division between the two does not in general line up in a tidy way with the future–past distinction, because much of the future will typically be regarded as **FIXTURES**, for the purposes of any particular deliberation. But is what is in **OPTIONS** always in the future? Here we need to be careful, I think. Deliberation seems to presuppose a personal or subjective time for the deliberating agent herself, and in that time, the choices that comprise the relevant **DIRECT OPTIONS** certainly lie in the personal future—indeed, perhaps this is what 'future' should *mean*, in this context. But this doesn't seem to imply that the material manifestations of these choices need also occur later in time than the process of deliberation itself. We can make some sense of the thought that deliberation might take place outside the temporal arena of the material world altogether—think of deliberation by a god, outside spacetime herself, but able to fix matters of fact by fiat, anywhere within spacetime.

All the same, we human deliberators seem to be firmly embedded in time, in a way that ensures that for us, the 'internal' or personal time of deliberation does line up with the external time of the material world; and hence guarantees that the immediate material manifestations of our choices occur *after* the deliberations of which they are the outcomes. For us, then, the immediate choices that we regard as **DIRECT OPTIONS** always do lie in the 'external' future, as well as in our personal future.

Our **INDIRECT OPTIONS** would have to lie entirely in the future, as well, if we took it that what lies in the past goes into **FIXTURES** by default. As a fact about how we ordinarily think about these matters, this does seem to be the case, or close to the case. Indeed, it seems to me plausible to regard this assumption—call it the *fixed past principle*, or **FPP**—as something like a piece of naive physics. Certainly, this suggestion about the presuppositions of our causal reasoning provides an immediate explanation of our intuitions about the stargate case—in effect, it is as if we assume we could know the position of the photon, before it reaches the stargate.

Calling **FPP** a piece of naive physics does not prejudge the issue as to whether it is perspectival. We are familiar with the fact that some parts of naive physics turn out to be more perspectival than others.<sup>21</sup> All the same, it is easy for considerations of the status

---

<sup>21</sup>Colour more so than up and down, perhaps, and that more so than weight.

of **FPP** to be pulled in two different directions, with different apparent implications for its objectivity. There is no real conflict, in my view, but it is important not to let either consideration obscure the other.

On the one hand (favouring objectivity), we seem to be able to make sense of ways our world might turn out to be in which **FPP** would clearly fail. With a little conceptual effort, we seem to be able to make sense of backward causation: cases in which the *indirect* outcomes of some of our actions lie in the past. Thus we seem to be able to make sense of versions of Newcomb problems in which one-boxing can be made to seem rational, because we can take ourselves, in effect, to *choose* what is already in the box. And more realistically, it has been suggested (e.g., by me, in Price 1996) that quantum mechanics might deserve to be interpreted in a way which gives us some kind of control over the past. Getting one's head around these suggestions certainly requires a bit of effort, but it seems doable, at least for most of us; which suggests, as has often been noted, that backward causation is not ruled out on analytic grounds. To that extent, then, **FPP** seems flexible, at least at the edges, in a way that depends on how the world turns out to be.

On the other hand (favouring perspectivity), this flexibility does not extend to what we might regard as the core of **FPP**: the assumption that the past is *typically* fixed, whereas the future is not. However much we acknowledge that in a Gold universe our time-reversed cousins would see things differently, we can't imagine our own perspective shifting to align with theirs. Yet as the comparison makes clear, inflexibility does not imply objectivity. The core of **FPP** remains perspectival. It is inflexible because we can't change our viewpoint, not because there is no viewpoint involved.

Thus I suggest that to the extent that our default deliberative template embodies **FPP**—the principle that nothing in the past can lie in our **OPTIONS**, **DIRECT** or **INDIRECT**—this is a product of two kinds of contingency. The more basic contingency concerns our own 'situation' in the world. It is characterisable, albeit indirectly, as the respect in which we differ from our Gold universe cousins; and in being grounded on this contingency, **FPP** is perspectival. The less basic contingency concerns the world itself, and is characterisable as the absence, at least in familiar arenas, of the kinds of phenomena which would be interpreted—by asymmetric agents such as ourselves—as allowing some control of the past. As I've already noted, the hypothesis that **FPP** provides our default deliberative template offers a ready explanation of our intuitions about the stargate case—and this could be further confirmed, I think, by exhibiting the malleability of those intuitions in the light of proposals about backward causation.

The view that **FPP** is perspectival does not imply that the fact that we hold the past and not the future fixed is 'up to us', or conventional in that sense. On the contrary, it seems to be determined by factors we certainly cannot change. To a large extent, I've suggested, it seems to be a consequence of our epistemic template. We regard the past as fixed because we regard it as knowable, at least in principle. This is clearly an idealisation, but one with some basis in our physical constitution. As information-gathering systems, we have

epistemic access to things in (what we call) the past; but not, or at least not directly, to things in (what we call) the future.

Plausibly, this fact about our constitution is intimately related to the thermodynamic asymmetry, at least in the sense that such information-gathering structures could not exist at all, in the absence of an entropy gradient. Although the details remain obscure, I think we can be confident that the folk physics reflected in the temporal asymmetry of our epistemic and deliberative templates does originate in *de facto* asymmetries in our own temporal orientation, as physical structures embedded in time.

These latter asymmetries are not perspectival, of course—they are as objective as the fact that our rocks rolled downhill in a northerly direction. But these objective asymmetries *in us* underlie a perspectival asymmetry in our causal concepts, in much the way that the objective asymmetry of the rocks underlies the perspectival asymmetry of accessibility. And as in that case, the perspectivity of the concepts is revealed in the fact that they continue to be applied asymmetrically, even in instances that lack any relevant intrinsic asymmetry in themselves. That was the point of the stargate example, in §5.

## 8 Contingency, ignorance and manipulability

At the end of §5, I noted the strong emphasis on intervention in much recent work on causation. In interventionist terms, the issue of the temporal directedness of causation comes down to the question as to why dependent ‘wiggles’ typically occur *after* rather than *before* the wiggle on which they depend. We now have two arguments to show that this is perspectival (and thus, as I claimed, that interventions are a Trojan Horse for causal objectivists).

The first argument exploited links to temporal factors, in order, by means of a simple symmetry argument, to make sense of the idea of agents with a different perspective—agents with the opposite temporal orientation, at the other end of a Gold universe. I’ve offered the following diagnosis. In posing the question, ‘What else wiggles, if we wiggle *this?*’, we normally simply presuppose that the past does not change, as a given of the enquiry in question. The constraint comes from the needs of the deliberative standpoint, as instantiated in creatures whose typical epistemic access is to things in the past. Idealised in a natural way, I’ve suggested, this epistemic imbalance requires us to treat the past as fixed, for deliberative purposes. *For us*, then, the notion of a wiggle thus becomes strongly time-asymmetric, in the way revealed by our intuitions about the stargate case.<sup>22</sup>

---

<sup>22</sup>Objection: Can’t we simply *observe* that wiggling a variable doesn’t change anything earlier than the time of the wiggle? If so, how can it be a perspectival matter? Reply: It isn’t so, for the issue in question always concerns a counterfactual—what would have been the case if we had wiggled, or wouldn’t have been the case if we hadn’t wiggled. In assessing these counterfactuals, we don’t rely simply on observation, for we never observe the counterfactual case.



The second argument puts these considerations into a more general framework. From a purely logical point of view, there's no need to constrain wiggling in this temporal way. I've already suggested that we can make sense of the idea of an atemporal god, able to wiggle the material world in a much less temporally-constrained manner; and in principle, such an example could be modified to 'fix' any set of facts whatsoever. But *something* has to be held fixed, for otherwise the question, 'What changes, if we change this?' has a trivial answer: 'Everything!'

So the strong temporal asymmetry of the notion of intervention—and hence, apparently, of our causal thinking in general—stems not merely from the fact that we are agents, but also from a further contingency concerning our temporal circumstances: above all, the strong temporal bias of our epistemic access to our environment. Our causal intuitions are a product both of general aspects of the architecture of deliberation, and of specific facts about the way in which that architecture is implemented in our own case.

I now want to show that there are two further respects in which the causal viewpoint is necessarily perspectival, also brought to light by an understanding of general features of the architecture of deliberation. Again, both factors have a lot to do with the epistemic constraints on deliberation, though they differ from the constraint we have already encountered in that it is ignorance, not knowledge itself, that does the crucial work.

The first new factor stems from a constraint on deliberation that may seem trivial. As we've already observed, the agent's perspective presupposes lack of prior knowledge of the outcome of the choice in question. The *OPTIONS* and the *KNOWABLES* are necessarily disjoint, as we put it earlier. In terms of interventions, this amounts to the observation that to ask in a sensible way, 'What else changes, if this changes?', there must be things we leave open, as well as things we hold fixed. So there is a contingency to the standpoint of the would-be interventionist that depends on what she is ignorant about, and hence can leave open, as well as one that depends on what she assumes known, and hence holds fixed. As I'll argue below, this apparently trivial fact turns out to have important implications for the plausibility of strong forms of realism about causation.

The second new constraint stems from the fact that interventions are supposed to be 'free variables', independent of anything except their effects. As is often noted in the recent interventionist literature, this leads to a puzzle, in cases in which the (physical or biological) agent of an intervention is part of the same closed system as the object 'wiggled'. Interpreted in an objectivist manner, then, interventionism makes the metaphysics of causation hostage to the possibility that there may be no causation, literally speaking, because there are no genuinely open systems.

The perspectival solution to this puzzle rests on an observation due to F. P. Ramsey. Ramsey is famous as a pioneer of pragmatic subjectivism about probability. In one of his last papers, he extends this subjectivist viewpoint to laws and causation. He links the asymmetry of cause and effect explicitly to the perspective we have as agents, saying that 'the general difference of cause and effect' seems to arise 'from the situation when we are

deliberating.’ (1978: 146) He then goes on to identify what he seems to take to be the crux of the agent’s perspective, viz., the fact that from an agent’s point of view, contemplated actions are always considered to be *sui generis*, uncaused by external factors. As he puts it, ‘my present action is an ultimate and the only ultimate contingency.’ (1978: 146)

I’ve argued elsewhere (Price 1992b) that this amounts to the view that an agent thinks of her own actions as probabilistically independent of everything except their effects—as not themselves determined by anything ‘further back’. This is where causal chains *begin*, as it were, from the agent’s own perspective. And this should be read in reverse, I think. We should explain the genealogy of the notions of cause and effect by noting that we apply the terms, initially, on the following basis: we say that B is an effect of A, when we think that *doing* A would be a way of ensuring B.<sup>23</sup> This is entirely in keeping with the interventionist insight, of course.<sup>24</sup>

I’ve also argued that this approach provides the most promising basis for a probabilistic theory of causality. Among its virtues is the fact that it avoids the problem of spurious causes: correlations due to common causes don’t translate into probabilistic dependencies from the agent’s point of view, because the presence of the common cause is incompatible with the assumption of *sui generis* origins. The argument turns on a defence of evidential decision theory against Newcomb-style objections. Indeed, I think the viability of the approach in general depends on this defence, for it is this that ensures that the probabilities in question need only be evidential, and hence not dependent on a prior modal notion (as invoked in causal decision theories). What needs to be shown is that correlations between prior causal states and actions do not translate into evidential dependencies from the agent’s perspective; and the crucial point is that in the means–end context any such subjective dependency would itself be a causal factor, so that the principle of total evidence would immediately undermine the judgement on which it was based. (For the details, see Price 1986, 1991, 1992b.) This argument shows that Ramsey’s suggestion is coherent in purely evidential terms, I think. As such, it is then available to ground our causal concepts in the way that Ramsey suggests.

For the moment, the relevant point is that Ramsey’s ‘contingency’ is entirely a product of the perspective of a deliberating agent. Its source lies not in the world, but in a certain kind of *ignorance* on the agent’s part—roughly, ignorance of the causes of her own actions.<sup>25</sup> Again, then, it is an epistemic constraint on deliberation and intervention.

In the most general terms, then, the argument for the perspectival character of causal concepts that goes something like this. Causation has a conceptual tie to intervention, and hence to deliberation. But the possibility of deliberation is epistemically constrained, in several ways—it depends on both knowledge and ignorance, and these epistemic factors are

---

<sup>23</sup>Or increasing the probability of B, in a more general version.

<sup>24</sup>Though without ambitions to objectivist metaphysics.

<sup>25</sup>*Pace* E. F. Schumacher, then, it is ignorance, not knowledge, that makes us free, at least in this case. Knowledge of the causes of our own actions would make us mere spectators.

contingencies, whose limits may well vary from agent to agent. Hence causal judgements are correspondingly perspectival: they are necessarily ‘situated’, relative to some implicit boundaries to the knowledge and ignorance of the agent concerned.

We’ve now distinguished three epistemic constraints, which all contribute to the deliberative perspective. The first was a constraint imposed ‘from below’—a lower bound on an agent’s knowledge, providing the main constraint on FIXTURES (and apparently based, in our own case, on our epistemic access to matters in the past). The two new factors are constraints imposed ‘from above’—upper bounds on an agent’s knowledge, deriving from its necessary ignorance, as it deliberates, both of what it is actually going to do, and of some of the immediate precursors of its decision. Again, these epistemic constraints are more fundamental than their temporally-based manifestations in our own case. As before, we can imagine a divine creature, able to intervene at will at arbitrary points of a space-time arena. Such a creature might have an atemporal notion of causation (the ‘effects’ of its interventions showing up in various directions, throughout the manifold). But it must share our limitations in one respect, if it is to think of itself as deliberating at all. It must be sufficiently ignorant for the notion of choice to make sense, by its own lights.<sup>26</sup>

It is worth noting that the ignorance required by the two new epistemic constraints seems deeply protected by the structure of the deliberative process. Deliberation always allows a kind of feedback loop, whereby the conclusions at any intermediate stage can feed into the deliberative process itself. This feedback can easily be made self-undermining, in the face of any claim to knowledge either of the causes of the contemplated action, or of the nature of the action itself. As cognitive structures embodying this kind of feedback, there is thus no danger that we are suddenly going to prove too clever for our own good, and have to abandon the deliberative perspective altogether. But the ignorance is there all the same, no matter how ineradicable. And our sense of ability, or control, depends on these epistemic limitations.

## 9 Locating causal perspectivalism (I): comparison with chance

This conclusion has a surprising corollary: far from being omnipotent, an omniscient creature *could not deliberate at all*. Recall the playing field example from §2—there, we noted that the referee doesn’t see the pitch as the goalkeepers do, asymmetrically divided into the near end and the far end. Similarly, I maintain, an omniscient creature does not see the world in causal terms—if science aims for the god’s-eye point of view, then Russell was right, and science has no place for causation. (More on this in a moment.)

This might be thought a *reductio* of the perspectival view. But is it really absurd?

---

<sup>26</sup>As we noted earlier, the notion of choice seems to presupposes a personal time in which choice takes place. So such an agent cannot be entirely atemporal, even if it occupies a different time dimension than the one in which its god-like interventions manifest themselves in our world.

On the contrary, I think, for it has a familiar and comparatively orthodox cousin in the contemporary philosophical landscape. Many writers maintain that if the physical world is deterministic, there are no non-trivial chances, or objective probabilities. Thus in a deterministic world all probability is epistemic, on this view, and we need it only because we are ignorant. Laplacean gods, omniscient about the complete present state of the world, and hence able to infer the rest, would have no use for probabilistic notions. My conclusion is that something similar is true of causation. Once we appreciate that agency depends on ignorance, we see that causation becomes epistemic in a similar way. Again, it is a way of thinking about the world that we need because we are not gods.

Indeed, the two cases go hand-in-hand. In the deterministic case, Laplacean gods omniscient about the complete present state of the world cannot be deliberators, because they have inferential access to their own interventions. So they have no more use for intervention-grounded causation than for probability. While in the indeterministic case, a stronger omniscience that gave a creature direct access to any part of spacetime would render redundant objective chance, as much as intervention-grounded causation. Chance, too, thus has a perspective-making tie to ignorance—a fact revealed, I think, as it is for causation, by the strong time-asymmetry of ordinary notions of chance and propensity.<sup>27</sup>

## 10 Locating causal perspectivalism (II): effective strategies

Another useful way to position causal perspectivalism in the philosophical landscape is to relate it to a well-known discussion by Nancy Cartwright, which explores similar territory in pursuit of different quarry. In ‘Causal Laws and Effective Strategies’, Cartwright begins with the familiar distinction between laws of association and causal laws. Noting Russell’s argument that, as she puts it, ‘laws of association are all the laws there are, and that causal principles cannot be derived from the causally symmetric laws of association’, she goes on to argue ‘in support of Russell’s second claim, but against the first.’ (1979: 419) So Cartwright agrees with Russell that ‘[c]ausal laws cannot be reduced to laws of association’, but maintains that ‘they cannot be done away with’. (1979: 419)

Cartwright’s argument for the latter conclusion is that causal laws are needed to ground an ‘objective’ distinction between effective and ineffective strategies. She illustrates this distinction with some ‘uncontroversial examples’. Thus:

[W]hat is, and what is not, a good strategy ... is an objective fact. ... Building the canal in Nicaragua, the French discovered that spraying oil on the swamps is a good strategy for stopping the spread of malaria, whereas burying contaminated blankets is useless. What they discovered was true, independent of their theories, of their desire to control malaria, or of the cost of doing so. (1979: 420)

---

<sup>27</sup>Rightly taken to be for grounds for suspicion about the ontological credentials of propensities, in my view, by Carl Hoefer (2004).

However, Cartwright argues, the ‘objectivity of strategies requires the objectivity of causal laws’. In other words,

causal laws cannot be done away with, for they are needed to ground the distinction between effective strategies and ineffective ones. . . . [T]he difference between the two depends on the causal laws of our universe, and on nothing weaker. (1979: 420)

Thus Cartwright contends that our primary reason for believing in the objectivity of causation is that it is needed to draw a distinction we need when we deliberate.

For my part, I’ve agreed that there is a deep conceptual link between causation and deliberation. However, I’ve argued that this suggests that causation is perspectival, because the deliberator’s viewpoint is necessarily ‘partial’, or incomplete. So the needs of deliberators are not a good guide to the nature of the world, as seen by God.<sup>28</sup> But what, then, of the distinction with which Cartwright began, between effective and ineffective strategies? Am I committed to denying that this distinction is objective? To maintaining that it is an anthropocentric matter that oiling swamps controls malaria, but burying blankets doesn’t?

Recall the lessons of the case with which we began. From the French perspective, it was certainly an objective matter that there were foreigners in Nicaragua. It was independent of French preferences in the matter, for example, or of any theories or projects conceived in Paris. Nevertheless, the distinction the French drew between themselves and ‘*les étrangers*’ was perspectival—a distinction drawn from the French viewpoint. The fact that the presence of foreigners was an objective matter, from that viewpoint, does not imply that there wasn’t a viewpoint involved.

Similarly in the case of causation, I maintain. From the homogeneous deliberative perspective that we humans all share, it is an objective fact that oiling Nicaraguan swamps is an effective strategy for reducing malaria. Indeed, it is an objective fact that when the swamps were oiled, that oiling *caused* a reduction in malaria. But again, the objectivity of these matters, from this viewpoint, does not imply that there is no viewpoint involved. It doesn’t imply that there is a viewpoint involved, either, of course. To investigate that matter, we need to do the hard work we’ve been doing in this paper. In particular, we need to think about the distinctive preconditions of causal judgements, of the use of causal concepts, in search of the tell-tale contingency and variability that is a mark of perspectivity.

I’ve argued that that investigation reveals that causal judgements are indeed perspectival. And it seems to me that properly understood, Cartwright’s own argument provides additional support for this view. Why? Because if we accept with Cartwright that causal laws are not reducible to laws of association, we face an issue analogous to one we raised at several points for various reductivist and primitivist proposals: viz., that of explaining *why*

---

<sup>28</sup>At least not directly, though of course we can still ask how the world has to be, for creatures of like us to have a need for these perspectival concepts. And we can still ask the question whose analogue in the Copernican case is this one: What kind of solar system looks like this, from the perspective of observers on the surface of a rotating planet in orbit around a sun? More on this below.

causation matters in deliberation in the way that it does. The argument goes something like this. If we make causation an objective matter, then either its relevance to us goes via physics, or it is somehow 'direct'. The latter option is mysterious, for reasons we've already canvassed in the case of various other primitivist proposals. The former option is in tension with the plausible Humean view that physics can be cashed in terms of laws of association.

True, the realist will respond that the point of Cartwright's argument is precisely that it shows that this Humean view of physics is mistaken—that the physical world contains primitive causal facts of some kind, not reducible to matters of association. But this reply simply avoids one horn of the dilemma at the cost of the other. Causal facts become mysterious primitives once more, knowable only by their relevance to decision. This dilemma is the real 'hard problem' for causal realism, in my view.<sup>29</sup> It is avoided, however, by a view that *begins* with deliberation, and sees causal judgements as projections and idealisations of judgements made from the agent's perspective.

Realists are likely to argue that the first horn of the dilemma is more attractive than I've claimed. One popular strategy, comparable to David Lewis's strategy in the case of chance, is to compare the access that this approach gives us to causation to the kind of access we have to theoretical entities in science. In each case, what we know about the object in question is that it is whatever 'plays a certain theoretical role'—in the case of causation, a role that includes, *inter alia* and centrally, a particular significance for decision.

Because it accords such a central role to the role of causal notions in practice, this approach is in one sense very close to the kind of pragmatic perspectivalism that I favour. How close depends on the prospects for the kind of reference-fixing machinery that is supposed to make it more realistic. And the prospects are rather poor, in my view. Briefly, if the crucial notion of theoretical role is cashed in causal terms, as it is in Lewis's account of reference-fixing elsewhere in science, then the approach is viciously circular in the case of causation itself—in this sense, causation isn't just another theoretical term.<sup>30</sup> If it is cashed in (non-causally-grounded) semantic terms, then, as Stephen Stich (1998) points out in an analogous context, the metaphysical enterprise becomes hostage to a theory of reference that is nowhere in range, above or beyond the horizon. And to what end? The interesting part of the story, the part this approach presumably shares with pragmatists such as me, is a story about how creatures like us come to use causal concepts. Moreover, as we noted earlier with reference to reductionist responses to the reversibility arguments, the arguments that support perspectivalism show that if and when reference does get fixed, in some interesting non-minimal sense, it will be in a way which reveals a relativity to the circumstances of the speaker—which implies that we pick out something different by 'cause' than do our time-reversed cousins in a Gold universe. In all, then, a great deal of work, for no significant benefit—or so it seems to me.

---

<sup>29</sup>To borrow some terminology from Hitchcock (2005), who borrows it in turn from David Chalmers.

<sup>30</sup>See Price (2001), §7, for an elaboration of this point.

The key point is that by *beginning* with deliberative practice, a perspectival view resolves the mystery that Cartwright's paper brings close to the surface, viz., the question why causation should be relevant to deliberation in the first place.<sup>31</sup> Moreover, the 'cost' of perspectivalism, in so far as there is one, is nicely circumscribed by Cartwright's discussion. For if Cartwright is right in thinking that the primary argument for causal realism turns on the apparent objectivity of the distinction between effective and ineffective strategies, then the main task facing perspectivalism is to find some way around that obstacle—to find some way to reconcile perspectivalism with our intuitions about the objectivity of effective strategies. And that, I've suggested, isn't all that hard to do. These are precisely the intuitions we should expect, given the homogeneous nature of the causal perspective.

## II Russell's republic revisited

I've agreed with Cartwright that causal laws cannot be done away with, because they are needed in deliberation. But for me, this reflects the peculiarities of the deliberative standpoint, rather than a perspective-independent reality. So for me, the conclusion that causal laws are indispensable comes with a qualification: *as long as we continue to deliberate*. In practice, of course, this qualification makes little difference. There's no risk that we humans will cease to be deliberators. But there's still a formal question about the implications of my view to the issue between Cartwright and Russell. Cartwright contends that Russell is wrong to think that causal laws can be eliminated from science. Where does my view stand on this issue?

The answer depends on how deeply embedded in science the causal viewpoint turns out to be, in two senses: whether it is dispensable, and if not, whether it encompasses the whole of science, or just an aspect of science. Concerning the former, it is easy to see how the causal viewpoint might be ineliminable. If explanation is an indispensable element of science, and explanation is essentially causal, then to adopt a scientific stance is necessarily to view one's subject-matter in causal terms. And if causal explanation is not only *a* goal but *the* goal of science, then we have the stronger conclusion. Causal perspectivalism would imply that the scientific viewpoint is *wholly* as well as *ineliminably* 'embedded'.

As I said earlier, this would not amount to a *reductio* of the perspectivalist view. It would do so if a perspectivalist claimed non-perspectival authority for his own pronounce-

---

<sup>31</sup>Lewisian realists do just as well in this respect, of course, for they begin in the same place. But they must avoid the 'dormative virtue' temptation, of thinking that their additional metaphysics buys them some deeper account of why causation matters to decision. As I have said, I think it is doubtful whether it buys them anything useful. Certainly knowledge of the causal facts can do no work in explaining or justifying our judgements about effective and ineffective strategies. Once the link between causal facts and facts about effective strategies becomes analytic, we don't have knowledge of the former until we've decided the issue about the latter. The epistemology of causation becomes parasitic on the epistemology of rational decision, rather than vice versa.

ments, for the perspectivalist project is itself scientific—it treats the use of causal concepts as an aspect of human behaviour and psychology, and sets out to provide an explanation, in the same scientific spirit in which we might investigate the emotions, say, or vision. Such an explanation will presumably *use* causal concepts (and related elements of reasoning, such as counterfactuals), but there seems to be no vicious circularity here (no more so, in my view, than in the fact that we use language in theorising about the origins of language). There would only be an inconsistency if the perspectivalist claimed a standpoint more detached than that of science in general, but such a claim seems entirely unmotivated, by the perspectivalist's lights.

Accordingly, the view that all of science is interestingly perspectival would not be inconsistent, so far as I can see. All the same, I don't think that *causal* perspectivalism has this consequence. Not only is there a legitimate place for a non-causal viewpoint in science, in my view, but such a viewpoint seems implicit in the perspectivalist project itself, considered as a scientific enquiry. After all, think about the kind of explanation the perspectivalist is looking for: viz., an understanding of how we humans came to employ causal concepts, in order to prosper in environments in which causation is not part of the pre-existing furniture.

As in §9, the case of probability provides a useful analogy. Think of the project of explaining the practical utility of acquiring epistemic notions of probability, for creatures in a deterministic (and therefore, on most accounts, chance-free) environment. In both cases, as I've argued elsewhere (Price 2004b), a natural way to approach the issue is in terms of what I called 'meta-models'—simplified models of idealised agents in idealised environments, intended to make sense of the question as to what difference it makes to creatures when they model their own environments in causal or probabilistic terms. This approach brings with it the idea of a non-causal or non-chancy reality, within which causal or probabilistic reasoning evolves.

In the present context, then, the point I want to stress is that to treat causal concepts as perspectival, in this naturalistic spirit, is inevitably to theorise about the non-causal world to which we apply these concepts—the bare Humean world, in the midst of which we embedded creatures come to think in causal ways. I've already noted the Kantian character of this project. In these terms, the bare Humean world plays the role of the non-causal 'thing in itself'. But unlike Kant's own *Ding an sich*, it is a realm to which empirical science—here, the naturalised Kantian enterprise of perspectivalism itself—gives us some kind of access.<sup>32</sup> Indeed, as I noted above, Kant's own Copernican analogy provides a model for the investigation in question: we ask ourselves, 'What kind of reality would look like *this*, from the peculiar standpoint we humans happen to occupy?'

In endorsing this enquiry, then, I side with Russell, against Cartwright, in favour of

---

<sup>32</sup>Of course, this would be compatible with the conclusion that there is some deeper sense in which the *Ding an sich* is off-limits to science, for more general reasons.



Humean science. But there's another important sense in which I disagree with Russell (though without thereby agreeing with Cartwright). By contemporary lights, it is natural to read Russell as a causal eliminativist. By these lights, the discovery that causation is a harmful relic of a bygone age, as Russell puts it, would be comparable to the discovery that there is no phlogiston, or luminiferous ether. But these are misleading comparisons for the implications of causal perspectivalism, in my view.

Why? Simply because there are two different ways to discover that something isn't as real as we thought. There's the kind of discovery we made with respect to phlogiston, ether, unicorns, leprechauns, and the like; and there's the kind of discovery we made with respect to foreigners. We didn't discover that foreigners don't exist, but merely that the concept *foreigner* is perspectival. And that's where causation belongs, in my view, along with folk favourites such as up and down, night and day, and the rising sun itself.<sup>33</sup>

Thus although I'm less realist about causality than Cartwright, I'm also a less revolutionary anti-realist than Russell himself. I don't want to eliminate causation altogether from science, but merely to put it in its proper place, as a category that we bring to the world—a projection of the deliberative standpoint. Causal reasoning needn't be bad science, on my view. On the contrary, it's often an indispensable construct for coping with the situation we find ourselves in, as enquirers and especially as agents. It is bad science to fail to appreciate these facts, but not bad science to continue to use causal notions, where appropriate, having done so. Some perspectives simply cannot be transcended.

By offering a modest, pragmatic, agent-centered view of causation, perspectivalism thus provides an the irenic third way between Russell and Cartwright. It forments revolution, but a quiet revolution, in the spirit of Kant's Copernican revolution, that avoids the mysteries of 'monarchist' metaphysics without the anarchic nihilism of causal eliminativism. It dethrones causation, certainly, but saves it, for all ordinary purposes, by revealing its human face—truly, then, a republican compromise.<sup>34</sup>

## Bibliography

- Cartwright, N. 1979. 'Causal Laws and Effective Strategies', *Noûs*, 13, 419–437.  
Gold, T. 1962. 'The Arrow of Time', *American Journal of Physics*, 30, 403–410.  
Hausman, D. 1998. *Causal Asymmetries*, Cambridge: Cambridge University Press.  
Hitchcock, C. 2005. 'What Russell Got Right', to appear in H. Price and R. Corry, eds.,

---

<sup>33</sup>The difference between causation and these examples is mainly that it is much easier to change perspective by moving around the planet, or moving off it, than it is to alter our temporal and epistemic perspective. That's why causation *seems* more objective, and is more indispensable, in practice. But the difference is one of degree.

<sup>34</sup>I am indebted to many people for comments and discussion of this material, among them David Braddon-Mitchell, Amit Hagar, Chris Hitchcock, Carl Hoefer, Jenann Ismael, Doug Kutach, Peter Menzies, Brad Weslake and Jim Woodward, and many participants in conferences in Sydney in July, 2003, and Venice in May 2004. I am also grateful to the Australian Research Council, for research support.

- Causation, Physics and the Constitution of Reality: Russell's Republic Revisited*, Oxford: Oxford University Press.
- Hofer, C. 2004. 'Time and Chance Propensities', typescript.
- Meek, C. and Glymour, C. 1994. 'Conditioning and Intervening', *British Journal for the Philosophy of Science*, 45, 1001–1021.
- Pearl, J. 2000. *Causality*, New York: Cambridge University Press, New York.
- Price, H. 1986. 'Against Causal Decision Theory', *Synthese*, 67, 195–212.
- 1991. 'Agency and Probabilistic Causality', *British Journal for the Philosophy of Science*, 42, 15–76.
- 1992a. 'Agency and Causal Asymmetry', *Mind*, 101, 501–520.
- 1992b. 'The Direction of Causation: Ramsey's Ultimate Contingency', in D. Hull, M. Forbes and K. Okruhlik (Eds.), *PSA 1992: Volume 2*, East Lansing, MI: Philosophy of Science Association, 253–67.
- 1996. *Time's Arrow and Archimedes' Point*, New York: Oxford University Press.
- 2001. 'Causation in the Special Sciences: the Case for Pragmatism', in D. Costantini, M. C. Galavotti and P. Suppes, eds., *Stochastic Causality*, CSLI Publications, 103–120.
- 2004a. 'On the Origins of the Arrow of Time: Why There is Still a Puzzle About the Low Entropy Past', in C. Hitchcock, ed., *Contemporary Debates in the Philosophy of Science*, Oxford: Basil Blackwell, 219–239.
- 2004b. 'Models and Modals', in Donald Gillies, ed., *Laws and Models in Science*, London: King's College Publications, 49–69.
- Ramsey, F. P. 1978. 'General Propositions and Causality', in D. H. Mellor, ed., *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, London: Routledge and Kegan Paul, 133–151.
- Russell, B. 1913. 'On the Notion of Cause', *Proceedings of the Aristotelian Society*, 13, 1–26. reprinted in *Mysticism and Logic*, Doubleday, 1953, 171–196.
- Spirtes, P., Glymour, C. and Scheines, R. 1993. *Causation, Prediction and Search*, New York: Springer-Verlag.
- Stich, S. 1998. *Deconstructing the Mind*, New York: Oxford University Press.
- Williams, D. C. 1951. 'The Myth of Passage', *Journal of Philosophy*, 48, 457–472.
- Woodward, J. 1997. 'Explanation, Invariance, and Intervention', *PSA 1996: Volume 2*, East Lansing, MI: Philosophy of Science Association, S26–41.
- 2000. 'Explanation and Invariance in the Special Sciences', *British Journal for the Philosophy of Science*, 51, 197–254.
- 2001. 'Causation and Manipulability', in *The Stanford Encyclopedia of Philosophy* (Fall 2001 Edition), Edward N. Zalta, ed., URL = <<http://plato.stanford.edu/archives/fall2001/entries/causation-mani/>>.
- 2003. *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.