Error, Reference, and the First Horn of Hempel's Dilemma

Colin Klein
Department of Philosophy, M/C 267
University of Illinois at Chicago
601 S Morgan St.
Chicago, IL 60607
cvklein@uic.edu

September 21, 2008

It would be nice if our definition of 'physical' incorporated the distinctive content of physics. Attempts at such a definition quickly run into what's known as Hempel's dilemma. Briefly: when we talk about 'physics', we refer either to current physics or to some idealized version of physics. Current physics is likely wrong and so an unsuitable basis for a definition. 'Ideal physics' can't itself be cashed out except as the science which has completed an accurate survey of the physical; appeals to it to define the physical must therefore end up trivial or circular. So defining the physical in terms of physics looks like a non-starter.

Hempel's dilemma has attracted a lot of attention in recent years.[1] With the notable exception of Melnyk,[2] few try to grab the first horn. I will argue that current physics is a more suitable candidate for definition than many have supposed. I argue that the only *quick* route to the inadequacy of definitions based on current physics is through a problematic theory of reference. If we adopt a reasonable causal-historical theory of reference, definition in terms of contemporary physics can tolerate theoretical error. Standard presentations of Hempel's dilemma thus end up unreasonably pessimistic.

Before I begin, two brief notes about the scope of the project. First, I will not take care to distinguish between 'physical' and 'fundamentally physical,' because am not convinced there is a useful distinction along these lines.[3] If you are, consider what follows to be a story about the fundamentally physical, with the assumption that there is some way to link the fundamentally physical and the merely physical. (For example, rocks and toasters may count as physical by being entirely composed of the fundamentally physical).

Second, there are two projects we might undertake when we set out to define the physical. We might try to give an *intensionally* adequate definition: that

[1] A good introduction is Barbara Montero "The Body Problem". *Nous* 33:2 (1999), pp183–200. A collection of recent papers on the topic can found in a special issue of *Philosophical Studies* 131(2006).

[2] See Andrew Melnyk, "How to keep the 'Physical' in Physicalism" *The Journal of Philosophy* 94:12 (Dec 1997), pp 622–637.

[3] Discussion would take me too far afield, but briefly: the 'fundamentally physical' usually means particles and their properties. But not all physics is particle physics: Newton and Einstein gave laws that were meant to apply to everything, not just to the mereological simples.

is, a definition that captures something about the essence of the physical, or analyzes the meaning of the word ‘physical’, or gives necessary and sufficient conditions for being physical. An intensionally adequate definition would be exceedingly valuable. It would, for example, allow us to determine whether physicalism is true of other possible worlds, or to ask of non-actual alien properties whether *they* are physical.

I am pessimistic about our ability to give an intensionally adequate definition. I suspect that the best we can hope for are limited, purely negative definitions that are useful in specific contexts—Wilson’s ‘not mental’, say, might be enough for philosophy of mind even though it offers little in the way of positive characterization.[4] Further, there is a sense in which we have already given up on intensional adequacy when we turned to science for guidance: physics is the search for fundamental but contingent facts about the world, and I see no special reason to think that what it discovers would have to generalize to other possible worlds.

Instead, I will give an *extensionally* adequate definition of the physical. By this, I mean a definition that (if successful) picks out the actual set of physical entities. It need not be adequate if we port it to other worlds, and it need not give us any special insight onto the nature of the physical. Extensional adequacy is an easier target to hit; despite this, it is still important. If the physicalist cannot give even an extensionally adequate definition, then he cannot even pick out the set of things that do the explanatory heavy lifting in his account. An extensionally adequate definition would avoid that especially embarrassing outcome.

So much for preliminaries. Return now to the first horn of the dilemma. It says that a definition like

> ϕ: The properties and objects referred to by current physics

should end up extensionally inadequate. We should expect the set picked out by ϕ to include some non-physical things, or to lack some physical ones.

[4] Jessica Wilson, “On Characterizing the Physical”. *Philosophical Studies* 131(2006), pp 61-99.

We should expect this (the standard story goes) because current physics is likely *wrong* in some way or other. Epistemological humility demands that we concede as much. Given that, surely ϕ must be inadequate.

This is too hasty. There are two ways in which current physics might be wrong. First, it might be *inaccurate*: that is, it might assert or imply some false propositions about physical entities. Second, it might be *incomplete*: that is, it might fail to assert or imply propositions about physical entities that do exist. Each of these types of error comes in two forms. One form of the error is likely but unproblematic; the other form would be problematic but is unlikely. As a result, we have no reason to expect that ϕ is inadequate—or, at least, mere epistemological humility does not give us such a reason.

Begin with inaccuracy. Physics is not complete, and physicists are only human. Surely some false propositions have sneaked in to current theory. That alone does not show that ϕ is inadequate, however. For ϕ is adequate insofar as we refer to the physical, and we may be wrong about features of a theoretical entity while still managing to refer to it. Bohr was wrong about many features of the atom, but there is no reason to think that he was not talking about *atoms*, rather than failing to refer. This is not a point specific to theoretical terms, but a more general point about reference. I no doubt have false beliefs about, say, the Sears Tower. Perhaps my false beliefs are numerous and systematic. Nevertheless, I can successfully refer to the Sears Tower when I utter its name.

In the background of this suggestion is, of course, an appeal to a causal-historical theory of reference. What explains my successful reference to the Sears Tower is the fact that I have had the right sort of causal interaction with it—I've seen it, walked around it, and so on. This causal interaction in turn figures in my production of the name 'The Sears Tower', and that is how I can refer despite my false beliefs.

Similarly, reference to a theoretical entity demands less than being entirely correct about it. It requires only that causal interaction with the entity figures in the production of the term we use to name that entity. Bohr could talk about atoms precisely because he had the right sort of causal interactions with atoms. Kitcher has argued that theorists can be quite wrong about their domains while still managing to refer. If we conjoin a causal-historical theory of reference with a context-sensitive translation theory for scientific terms, we can preserve reference even in cases where the same term is used to refer to different entities in different contexts, or where the same term sometimes refers and sometimes fails to refer.[5]

So the mere fact that our theory may say some false things about the set picked out by ϕ does not threaten the supposition that ϕ picks out all and only the physical things. What would threaten the adequacy of ϕ would be errors of a more serious kind. Some theories contain parts that are so flawed that the terms used *fail* to refer. This sort of failure has occurred. Plausibly, theories of impetus were inaccurate in this more serious sense because there are many contexts in which we cannot trace a causal chain from utterances of 'impetus' back to any suitable entity.

Given a causal theory of reference, the adequacy of ϕ is threatened only if current physics is inaccurate in this more serious sense. There is no direct evidence that it is. Nor does humility demand that we think it is: claims that contemporary physics is *radically* mistaken would require serious evidence and argument, not a modest reflection on our epistemic capacities. Lacking such evidence, there is no reason to think that current physics is inaccurate enough to threaten ϕ.

Of course, one could insist on the inadequacy of ϕ by denying the causal-historical theory of reference for theoretical terms. A strong descriptivist theory of reference, for example, could turn small inaccuracies into referential failures. While the causal-historical theory of reference is philosophically controversial, two things that can be said in its favor in the present context.

First, the preceding has shown that *if* the causal-historical theory is a good theory of reference, then the standard version of the first horn goes by too quickly. The inadequacy of ϕ is not something we are forced to accept out of simple humility: if ϕ is threatened by inaccuracy, we will need a long and careful defense of the link between error and referential failure. Second, the strong descriptivism required to move from inaccuracy to failure would also

[5] Philip Kitcher "Theories, Theorists, and Theoretical Change", *The Philosophical Review* 87:4 (Oct 1978) pp519-547.

imply that widespread referential failure happens frequently—arguably, each time any substantial progress is made in science. That is, again, stronger than mere humility should force us to accept. In either case, epistemic humility alone does not create any problems.

So much for inaccuracy. What about incompleteness? If current physics simply remains silent about some of the physical objects, then ϕ will also be extensionally inadequate. And again, simple humility seems to force upon us the conclusion that physics is incomplete in *some* sense.

Again, careful attention to potential types of incompleteness shows that ϕ is more secure than it might first appear. Consider first how incompleteness works in other sciences. There are species of bacteria that exist, but that no human has come into contact with. Nevertheless, scientists have had the right sort of experimental interactions with enough of the distinctive properties of bacteria that they may legitimately say 'the bacteria' and include these undiscovered species. They are not limited to saying merely "anthrax, and botulism, and...", but can say something like "those *and* any other bacteria we have not yet seen". We can refer to a type as a whole without having acquaintance with all of the members of the type, or even to all relevant sub-types of that type.

Incompleteness is only problematic if it is more systematic. If the only disease-causing agents we knew of were bacteria, say, and we had never successfully referred (in theoretical contexts) to viruses—well, then there would be a problem. When we talked about the causes of disease, we would only capture bacterial diseases, as bacteria would be the only disease-causing agents we would have had had had theoretically relevant contact with. We would thus omit a whole type of thing, not just some tokens of a type with which we were already familiar.

Suppose that we update ϕ to reflect the fact that our current physics may lack terms for some of the physical types:

ϕ^* The properties and objects referred to by current physics, *and* any other member of any type successfully referred to by current physics.

There are two questions we can ask about ϕ^*. First, we can ask whether the added clause is *legitimate*—that is, whether the 'any other member...' portion is something that can be reasonably included in a definition. Second, we may ask whether ϕ^* is extensionally adequate. Like ϕ, it will not be threatened by inaccuracy, but it remains open whether it would still be threatened by incompleteness

There several reasons to think that ϕ^* is legitimate. First, note that the new clause is merely an extension of an abstraction process already in place in simpler definitions like ϕ. Current physics cannot enumerate each atom that exists: they are too numerous and too dispersed. But no one seriously argues that this prevents reference to the set of atoms. Physics can abstract away from the details of particular entities, and that abstraction permits reference to individual entities that we have not encountered and could not fully describe. Hence the 'any other member...' move is not obviously illegitimate; it is indeed already implicit in less controversial definitions.

Indeed, note that even a *intensionally* adequate definition wouldn't enumerate each physical type. That is, an intensionally adequate definition would ensure extensionally adequacy by ensuring that anything physical satisfied the defining description associated with it, *not* by listing the physical things. But this is all that ϕ^* does—save that instead of neutral description, it gives a description bound to the types already discovered by physics.

Finally, note that a causal-historical theory of reference can legitimate this process of type-abstraction. We refer to the set of all hydrogen atoms by having the right sort of causal interactions with some of the hydrogen atoms. The important kinds of causal interactions are those that involve some theoretically important properties of the entities interacted with. So, for example, what makes it the case that (some) of Priestley's utterances of 'dephlogisticated air' refer to oxygen is that he had an interaction with oxygen that revealed a theoretically important property of it—that of supporting ordinary combustion. This property is one that distinguishes it from the other elements. This interaction determines that he referred to oxygen. In referring to oxygen, he referred to oxygen as a *type*, not just to individual samples that he produced.

Similarly with more abstract physical kinds: as long as we've had the right sort of theoretical interactions with fermions (again, interactions where the theoretically distinctive properties of fermions are what figure in the production of the term 'fermion', whether or not we are right about what those properties are), then ϕ^* will capture all of the types of fermions, not just the ones we know about. If we've interacted with enough relevant features of fundamental particles, the term 'fundamental particle' may itself be enough to capture most of the fundamental entities.

So in sum, I see no reason to think that the final clause of ϕ^* is in any way illegitimate. A trickier question is whether it is *adequate*. For all I've said, it is still possible that physics is more seriously incomplete, and leaves out whole categories of things.

It would be peevish to insist that an objector produce some type of physical thing that he thinks physics leaves out. However, there is another sort of evidence that he could produce. He could show us phenomena that do not seem to be explicable by any *type* of thing that we are familiar with. I suggest that insofar as we should be queasy about ϕ^*, it is because we can point to some areas like this—the possibility that galaxies are composed of exotic sorts of matter, for example, might be one thing that would make $\phi*$ problematic.

Whether these cases *actually* threaten the adequacy of ϕ^*, however, is something that would still need serious discussion and debate. Again, mere humility does *not* require us to admit that current physics is incomplete in the threatening sense. Current physics may already have the resources to refer—sketchily, incompletely, unhelpfully—to all relevant kinds of physical thing. If so, there is still plenty for physicists to do—but ϕ^* would remain an extensionally adequate definition of the physical.

Current physics is surely wrong, and surely incomplete. That alone does not prevent us from using it to define the physical. Upon closer examination, the reasons for thinking that it might be problematic usually involve implausibly strong requirements on reference. A weaker, more realistic theory of reference tolerates error and ignorance; it may be tolerant enough to make the first horn of Hempel's dilemma easier to grasp.