



## Review of Scott R. Sehon's *Teleological Realism*

Carol Slater  
Swanson Academic Center B-30  
Alma College  
614 W. Superior St. Alma,  
Michigan 48801, U.S.A.  
Copyright (c) Carol Slater 2007  
cslater@alma.edu

PSYCHE 13 (1), April 2007

**REVIEW OF:** Scott R. Sehon (2005). *Teleological Realism: Mind, Agency, and Explanation*. A Bradford Book/The MIT Press Cambridge MA.

Like the ring of fire around the Pacific, conceptual fracture between everyday acceptance of mentality and allegiance to the physical arouses uneasy attention. Theorists have dedicated impressive ingenuity to domestication of belief/desire psychology within a physical worldview; they have enthusiastically welcomed (or stoically contemplated) its demise in the wake of inevitable (or possible) falsification by future science. At least one philosopher has urged (if only briefly) that we cross our fingers when attributing intentional states. Rejecting assumptions common to these responses, Scott Sehon proposes that the claims of commonsense psychology (CSP) cannot and therefore need not be vindicated by inclusion among the truths of physical science (PS), nor can they be threatened by conflict with its findings. Facts about mind are in this sense “foundationless” but still “firm,” by virtue of making possible a view of each other as persons (p. 10). Central to Sehon’s brief is his proposal that CSP explanations of human behavior are teleological rather than causal, answering questions about the purposes of agents rather than the antecedents of their actions. Sehon takes up a wide range of related issues—Jerry Fodor’s brand of functionalism, Lynne Rudder Baker’s nonreductive “Practical Realism,” relations between global supervenience and entailment, and between causal explanation and Humean moral theory. Sehon does not, however, discuss either consciousness or intentionality beyond noting briefly that the first is very likely apt for reduction, the second, almost certainly not.

Sehon’s target is what he calls *strong naturalism*—the disjunctive claim that propositions of physical science (PS) can *either* subsume *or* contradict propositions of common sense psychology (CPS). Against strong naturalism, Sehon proposes that “common-sense facts about mind and [rational] agency” cannot be “squared” with a

scientific perspective (4). “If the question is how agents *fit* within the natural order, then my answer is that...agents do *not* fit within that order...The realm of reasons, values, agency, beliefs and desires is outside the realm of the natural sciences...” (230-231). Cast in Sehon’s preferred terms of art, CSP is logically autonomous with respect to PS, which cannot entail (any substantial part of) the propositions of CSP nor entail their contradictions. Unlike the murkiness of initially unformed intuitions about the nature of the mind-body problem, “The notion of entailment is given a precise meaning in the context of symbolic or formal logic....[It] is familiar from dozens of formal logic textbooks” (19).

What, then, blocks PS entailments that could either vindicate or falsify CSP? It might be supposed that brute disparity of predicates would do the trick. As Sehon points out,

CSP makes claims like “Joan wants a glass of wine.” The problem is that...*wanting*...is presumably not a predicate that will appear in the statements of physical science. This means that there is no way that the statements of physical science could formally entail “Joan wants a glass of wine.” In general, when two schemata or logical formulas do not share any predicate letters or sentence letters, there will be no entailment relationship between them...Thus, since there are predicates of CSP that do not appear in the statements of physical science, the facts of physical science alone cannot entail the claims of CSP. (28-29).

Since this observation applies equally to entailment of negations of CSP claims, it would appear that strong naturalism is precluded. Things are, however, not that simple. Although straightforward entailment between claims with different vocabularies is indeed impossible, Sehon argues that provision of “bridge laws,” similar to those involved in Nagelian reductions, allows entailment between PS and CSP. He cautions that the requisite auxiliary propositions could not be empirical (as Nagel appears to have proposed), but must be “claims that we [practitioners of CSP] accept about the mental irrespective of our commitment to the actual existence of mental states....[i.e.,] our own best reflections on the meaning or implications of the language of CSP...” (30). If I understand Sehon correctly, such “methodological observations” (what I would have called “coordinating definitions”) would allow entailments between PS and CSP. If this is the case, Sehon needs to give us another reason why the claims of PS and CSP are incommensurable. In Chapter 5, he does just this.

Sehon tells us that strong naturalism implies that “the propositions of CSP and physical science are, so to speak, playing the same game” (57).

If...CSP is to be subsumed within physical science, then CSP must be doing the same sort of thing as science, even if in a somewhat unsystematic and immature way....[If CSP is to contradict PS]...some central portion of CSP must make claims that conflict with those of physical science...[S]ome of its claims [must] at least purport to be scientific assertions (57).

For strong naturalism to be correct, therefore, CSP must be “a scientific or proto-scientific theory,” having “the same goals and...mov[ing] in the same conceptual space” as “paradigmatic sciences” such as physics, chemistry, and biology (57). Reminiscent of

Wittgenstein's critique of Frazer's *Golden Bough*, this claim is notably not argued for and may not strike readers as entirely obvious. Chapter 5 is devoted to showing that CSP does not meet these criteria: it exhibits both normativity and context sensitivity, and its proprietary terms fail to function as natural kind terms. (One might question whether Sehon may not be setting the bar here a bit high for "protoscience": historical and developmental forerunners of unproblematically scientific systems seem not infrequently to fall short of his standards but it seems odd to imagine that they were incommensurable with their successors.)

Sehon's concerns about normativity include the usual suspects.

- (1) "CSP explanations of human behavior typically imply that the agent had a normative reason [i.e., justification] to behave in the way she did. Causal explanations in the natural sciences do not carry any such implication" (59).
- (2) Unlike natural causal sequences, sequences of thoughts characterized as inferences are evaluable as right or wrong, reasonable or unreasonable (59–60).
- (3) The attribution of mental states is governed by principles of charity and thus by an assumption of rationality. "When we interpret a person with the resources of CSP, we assume that, roughly speaking, she believes what she *ought* to believe and desires what she *ought* to desire"; to attribute "a completely *irrational* set of beliefs to an agent defeats the purpose of belief attribution" (61-62, stress added).

Next, attributions of states such as understanding, belief and desire may be shown to *vary unsystematically with context*. Where Steve Stich has presented this as evidence that folk psychology is irredeemably *bad* theory, Sehon urges that it puts paid to CSP's being so much as in the same ballpark as the natural sciences (63-64).

Finally, Sehon urges that unlike terms proper to the paradigmatic sciences, mental state terms do not, on any of several understandings of 'natural kind term,' function as such.

On Sehon's account, then, CSP is not protoscience and therefore not susceptible to either vindication or falsification by PS. What is more, to the extent that it is the assumption that CSP *is* protoscience that has lent an air of inevitability to causal interpretation of psychological explanations, establishing that CSP is *not* protoscience opens space for "the cornerstone of teleological realism"—the claim that "action explanation is *not* reducible to causal terms" (58).

In Chapter 6, letting 'belief' go proxy for any of the propositional attitudes, Sehon presents a direct argument against causal psychological explanation, establishing each of (1) and (2) by subsidiary arguments.

- (1) Beliefs can only explain behavior if beliefs *are* (in some sense) brain states.
- (2) Beliefs are *not* brain states. Therefore,
- (3) Beliefs cannot explain behavior causally.

For (1), Sehon asks, “[I]f beliefs are *not* brain states, by virtue of *what* does one have the belief that *p*?” (76, stress added). He argues that the “most natural” alternative—Lynne Rudder Baker’s “pure dispositional” account of propositional attitudes—cannot provide adequate causal explanations of behavior, nor can a closely related dispositional account. Absent an alternative, he concludes that beliefs can only explain behavior if they are (in some sense) brain states.

For (2) Sehon presents a somewhat puzzling argument against the so-called Standard View, the view that beliefs are identical to (constituted or realized by) brain states. First, Sehon shows that instantiating a Quinean methodological maxim with two premises—*There are mental states* and *Science fails to find brain states with which mental states can be identified*—leads us to reject the Standard View. Next, he argues that falsity of the Standard View conditional on the empirical premise that science fails to find appropriate brain states allows us to reject it *tout court*. “If the Standard View is true at all, it must be true regardless of how the reification question is empirically answered” (88). If the Standard View is *not* true on the empirical assumption that reification fails, the Standard View is not true at all, thus giving us (2).

Sehon’s treatment is brisk and I may have failed to follow his reasoning. My best try, however, requires reading ‘Standard View’ in two different ways—one in order to establish its conditional falsity and another to establish that this, in turn, shows its unconditional falsity. Because (2) is a crucial step in establishing what Sehon takes to be a key claim, let me invite interested readers to come to their own conclusions.

A premise has high privilege—we are justified in resisting its falsification—either if it is based on observation or if its negation would require large scale alterations elsewhere (principle of conservatism).

Methodological Maxim (10).

- (i) If *X* and *Z* jointly imply  $\sim Y$  and
  - (ii) if *X* and *Y* are highly privileged and
  - (iii) *Z* is not highly privileged and
  - (iv) *X* alone does not imply  $\sim Y$ ,
- then reject *Z* (86, reformatted for clarity).

Sehon instantiates the Maxim as follows:

Let *X* be the proposition that reification fails, that empirical science does not find brain states that are plausible candidates for identification with mental tokens.

Let *Y* be the claim that there are mental states.

Let *Z* be the Standard View (86).

He then argues that all four conditions of the antecedent are satisfied, permitting us to reject the Standard View.

The failure of reification (*X*) and the Standard View (*Z*) must jointly imply that there are no mental states ( $\sim Y$ ).

This is the case if ‘the Standard View’ (*Z*) is read as *If there are mental states, then they are identical to brain states* (p. 84). Call this reading Standard View 1.

(i) is satisfied.

*X*, the claim that science fails to find the relevant brain states, is presumably observational, and therefore has high priority.

*Y*, the claim that there are mental states, is a key component of everyday practice and therefore also has high priority.

(ii) is satisfied.

As a philosophical claim, *Z*—if there are mental states, then they are identical to brain states—does *not* have high privilege. According to Sehon, “[B]arring an extremely compelling rationale, it is philosophical hubris to think that philosophy can dictate great changes in scientific or common-sense practice....Despite its intuitive appeal to many philosophers, the Standard View must be seen as a claim of low privilege unless it can be given a compelling rationale” (87-88).

(iii) is satisfied.

Finally, by itself, *X*, the claim that science fails to find brain states corresponding to mental states, does not imply  $\sim Y$ , that there are no mental states.

(iv) is satisfied.

If we accept the methodological Maxim and Standard View 1, we do indeed find ourselves obliged to reject *Z*, which is to say, we reject Standard View 1.

How about Sehon’s next claim? Call it Truth Regardless.

(TR): “If the Standard View is true at all, it must be true regardless of how the reification question is empirically answered” (88).

If we are still talking about Standard View 1, (TR) is by no means immediately obvious. What, precisely, entitles us to discharge the conditional premise *X*?

It may be helpful here to note that in arguing for the low privilege of the Standard View, Sehon says, “[T]he Standard View is a philosophical claim about [1] the commitments of common sense and [2] the nature of mental states” (88). There are, in fact, two importantly distinct proposals here.

What we have called Standard View 1 is a claim about the nature of mental states: if there are any, they are brain states. If the Standard View is also a claim about what common sense is committed to, it would seem to be the claim that common sense is committed to the claim that if there are any mental states, they are brain states.

Let us call this second claim Standard View 2.

Standard View 2: It is a “commitment of the folk” with respect to CSP that if there are any mental states they are identical to brain states.

*Standard View 2 is thus the claim that Standard View 1 is a conceptual truth of CSP.* On this interpretation of ‘Standard View,’ (TR) is obviously true. By definition, if *p* is a conceptual truth, then *p* cannot be falsified by any possible way things might be; if *p* can be falsified conditional on an empirical premise, *p* is not a conceptual truth. Standard View 2, which claims that Standard View 1 is a conceptual truth of CSP, would indeed be

false if it turned out that Standard View 1 could be falsified conditional on an empirical premise.

Clearly, Standard View 1 and Standard View 2 are not logically equivalent claims and cannot be swapped *salve veritate*. Just as Standard View 1 cannot be substituted for Standard View 2 with regard to establishing falsity *tout court* in the face of conditional empirical falsification of Standard View 1, so Standard View 2 cannot be substituted for Standard View 1 in (i) of the Maxim: empirical failure of reification is neither here nor there with regard to the truth of a claim about the folk's commitment to identification of beliefs with brain states. Reading 'Standard View' in two different ways is a problem.

Problematic or not, Sehon believes that he has a convincing argument that mental states are not brain states and is thus entitled to conclude that beliefs (or other propositional attitudes) cannot provide causal explanations of behavior. For good measure, Chapter 7 is directed against attempts by Alfred Mele, John Bishop and Christopher Peacocke to reduce CSP's distinctively teleological explanations to the causal terms of natural science. Their collective failure, Sehon tells us, "is highly problematic for the strong naturalist view of mind and agency" (93). (Reprise: (1) Scientific explanations are typically causal; (2) If CSP explanations are not causal, CSP is not in the same ballpark as science; (3) If CSP is not in the same ballpark as science, neither CSP's claims nor their negations can be entailed by PS. Thus, if CSP explanations are not causal strong naturalism fails.)

Sehon now turns to "the key move in the attempt to show the logical independence of CSP and physical science" (135), a positive characterization of psychological explanation as *teleological*. On this view, locutions such as 'Mary went to the kitchen because she wanted some wine' do not explain Mary's kitchen-going behavior by identifying its causal antecedents but, rather, identify Mary's purpose, the goal toward which Mary's kitchen-going behavior was directed. In Chapter 9, Sehon presents what he calls "the epistemology of teleological explanation" (137)—the systematic principles in terms of which practitioners of CSP make and justify such identifications.

Identification of goals is guided by two broad principles, the first shared with theorizing in physical science:

(S) Given two theories, it is unreasonable to believe the one that leaves significantly more unexplained mysteries" (138).

When theorizing about agents, (S) is augmented with the Davidsonian

(R) Given two theories of an agent, it is unreasonable to believe the one according to which the agent is significantly less rational" (139).

(R), in turn, is refined and qualified.

In trying to find a teleological explanation according to which an agent  $\phi$ ed in order to  $\phi$ , we...do the best we can in jointly satisfying instructions (I<sub>1</sub>) and (I<sub>2</sub>).

(I<sub>1</sub>) Find a  $\phi$  such that  $\phi$ ing is optimally appropriate for  $\phi$ ing, given a viable theory of the agent's intentional states and circumstances.

- (I<sub>2</sub>) Find a  $\varphi$  such that  $\varphi$ ing is the most valuable state of affairs toward which  $\phi$ ing could be directed, given a viable theory of the agent's intentional states and circumstances (146, 147).

Readers with a longish memory may see this descriptive work as an armchair version of what sociologists once called 'ethnomethodology'—the making explicit of everyday grounds for everyday knowledge claims. It could also be regarded, in a more contemporary context, as theorizing about an important aspect of social cognition. Insofar as the goal seems to be justification of the practices of teleological explanation as well as their description, one might discern a family resemblance to good old-fashioned rational reconstruction of science. ("To a large extent, the concerns and questions of CSP differ from those of the natural sciences, but this does not mean that anything goes. CSP is constrained by its own internal principles, in addition to the principles of logic and simplicity that also constrain the natural sciences" (231)). A reader might therefore wonder how much of his inquiry Sehon takes to involve "largely conceptual question[s] that can be addressed in a relatively a priori fashion" (9) and where, if anywhere, empirical methods might be useful.

Justification of a teleological explanation, in turn, involves assessing the truth of relevant counterfactuals. Sehon notes that although this is also the case with causal explanations, the relevant counterfactuals will differ: "[T]he fact that teleological explanations support a form of counterfactual not supported by causal explanation strongly suggests that teleological explanation is not reducible to causal explanation" (159-160).

Finally, in something of a coda, Sehon assures us that the solution he proposes neither lusters us with surplus "mysteries," in violation of Occam's canon, nor asks us to countenance spooky supernatural properties, in violation of materialist scruples. "That there is a realm of facts that employs concepts having no application to most of the rest of the natural world...does not mean that we are beyond the physical laws of nature... We are every bit as much a part of nature as are the inanimate elements that surround us. But there are also facts about what we value, what we think, and what we do, and these facts have no counterparts when the subject is a rock or a tree" (231).

This review cannot do justice to Sehon's detailed presentation of a provocatively nonstandard position. Anyone interested in philosophy of action will find it worthwhile to work through his arguments with care and go back to the sources he considers. Sehon writes clearly, gracefully and with unmistakable passion—I found this book touching as well as challenging. MIT Press has done a notably handsome job of production and a respectable if not flawless job of copy-editing. It would, I think, be the better part of caution to fly *Teleological Realism* past a few students before taking it on in a seminar.