

## *On the Structure of Sellars' Naturalism with a Normative Turn*

James R. O'Shea  
University College Dublin  
(jim.oshea@ucd.ie)

[Forthcoming in Willem deVries, ed., *Empiricism, Perceptual Knowledge, Normativity and Realism: Essays on the Anniversary of 'Empiricism and the Philosophy of Mind'*, Oxford University Press.]

In recent decades an increasing number of philosophers influenced by Wilfrid Sellars have stressed the importance of a distinction between the normatively structured 'logical space of reasons' on the one hand, and the proper domain of naturalistic causal explanations characteristic of modern natural science on the other. Three major works have been of particular significance in this respect: Richard Rorty's *Philosophy and the Mirror of Nature* (1979), John McDowell's *Mind and World* (1994), and Robert Brandom's *Making It Explicit* (1994). However, each of these authors also contends that the laudable aspects of Sellars' account of the space of reasons and his famous rejection of the 'myth of the given' must be detached from the regrettable *scientism* which they judge to be an unfortunate aspect of Sellars' own philosophy.

At the same time, on the other hand, an equally impressive group of philosophers ranging from the eliminative materialism of Paul Churchland to the various non-eliminativist scientific naturalisms characteristic of Daniel Dennett, Ruth Millikan, William Lycan, and Jay Rosenberg, have in their different ways been inspired rather than put off by Sellars' defense of a strongly scientific realist conception of reality, and in particular by his resulting investigations into how it is possible to reconcile that conception with our own experiential self-understanding as it appears within what Sellars called the 'manifest image of man-in-the-world'.<sup>1</sup>

What I explore in this paper is the difficult question of the nature of the relationship between the natural and the normative as it was conceived by Sellars himself. I shall argue that Sellars' own view represented an attempt to defend both the *irreducibility* of the normative space of reasons and yet, simultaneously and in another sense, its comprehensive *reducibility* from the perspective of an ideal scientific conception of the nature of reality and of the human being. Questions concerning the nature of the relationship between the 'natural' and the normative are of course among the most hotly

---

<sup>1</sup> The former philosophers (Rorty, et al.), who emphasize the quasi-Hegelian aspects of Sellars' 'space of reasons', have sometimes been characterized as the 'left wing' Sellarsians; while those philosophers who stress the importance of Sellars' naturalism and his strong scientific realism tend to be known as the 'right wing' Sellarsians. In this paper I argue that Sellars' view represented an attempt to show how the irreducibility of the normative 'space of reasons' is consistent, in another sense, with its thoroughgoing reducibility within what Sellars calls the 'scientific image of man-in-the-world'.

contested issues in contemporary philosophy. It may be that Sellars' own views on this particular issue still have something left to teach.

## Section I

### The normative 'space' of conceptual thinking:

in one sense *irreducible*, in another sense *reducible*

Many of the recent and most well-known discussions have quite properly emphasized Sellars' conception of *knowledge* as a normative standing in the logical space of reasons, as in the following famous passage from 'Empiricism and the Philosophy of Mind':

The essential point is that in characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says. (EPM VIII.36)

In a similar spirit, earlier in the same work, Sellars had claimed that

. . . the idea that epistemic facts can be analyzed without remainder – even 'in principle' – into non-epistemic facts, whether phenomenal or behavioural, public or private, with no matter how lavish a sprinkling of subjunctives and hypotheticals is, I believe, a radical mistake – a mistake of a piece with the so-called 'naturalistic fallacy' in ethics. (EPM I.5)

Furthermore, not only our epistemic states but more basically *all* our conceptual capacities generally were held by Sellars to be in some sense *irreducible in principle* to any description or explanation of those states in terms that refer only to non-conceptual processes and causal relations, however complex.

At the same time, however, in his overall views on truth and ontology Sellars defended what he called "the thesis of the primacy of the scientific image" (PSIM 38). As he put it in his famous (or infamous) *scientia mensura* dictum, "in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not" (EPM IX.42). In his ontology Sellars was both a thoroughgoing nominalist and a scientific naturalist from top to bottom. If we set aside certain important distinctions that have to be made in light of his novel view of 'sensa' and sensory consciousness,<sup>2</sup> Sellars' naturalism took the form of a comprehensive

physicalism or materialism.<sup>3</sup> As applied to the crucial case of the nature of the human being, Sellars makes it clear that in the final ontological reckoning or ideal synoptic vision of the world, persons are revealed to be complex “bundles” or “multiplicities” of micro-physical processes (cf. FMPP III.125: “The way would be open to a bundle theory of persons. A person would be a bundle of absolute processes...”; and cf. PH 101). On Sellars’ view, then, “the scientific image of man turns out to be that of a complex physical system” (PSIM 25).

The difficult interpretive issue, or so I believe, concerns just how Sellars conceived his account of the *irreducibility* of the norm-governed conceptual capacities of persons within the ‘logical space of reasons’ to be consistent with the ideal physicalist ontology and the strong *reducibility* claims that are embodied in his thesis of the primacy and completeness of the ideal scientific image of the human being. The following passage from ‘Philosophy and the Scientific Image of Man’ brings out the heart of this difficult issue, with reference to the essentially holistic nature of the normative space of reasons (the final sentence in particular is what I want to focus on):

... I want to highlight from the very beginning what might be called the paradox of man’s encounter with himself, the paradox consisting of the fact that man couldn’t be man until he encountered himself. It is this paradox which supports the last stand of Special Creation. Its central theme is the idea that anything which can properly be called conceptual thinking can occur only within a framework of conceptual thinking in terms of which it can be criticized, supported, refuted, in short, evaluated. To be able to think is to be able to measure one’s thoughts by standards of correctness, of relevance, of evidence. In this sense a diversified conceptual framework is a whole which, however sketchy, is prior to its parts, and cannot be construed as a coming together of parts which are already conceptual in character. The conclusion is difficult to avoid that the transition from pre-conceptual patterns of behaviour to conceptual thinking was a holistic one, a jump to a level of awareness which is irreducibly new, a jump which was the coming into being of man.

---

<sup>2</sup> The qualifications that have to be made in relation to Sellars’ views on sensory consciousness do not affect the present issue. Sellars’ ‘sensa’ are ultimately what he calls ‘physical<sub>1</sub>-but-not-physical<sub>2</sub>’ phenomena. Roughly, they are causally efficacious spatio-temporal processes (i.e., they are physical<sub>1</sub>) rather than being mere epiphenomena; however, Sellars’ proposes that our sensa do not obey the *mechanistic* physical laws that are adequate to describe non-living (i.e., physical<sub>2</sub>) matter. For further discussion see J. O’Shea (2007, in press), chapter six.

<sup>3</sup> It is important to bear in mind, though I shall not stress the point in this paper, that Sellars is a *non-reductive* physicalist in at least the following sense: he does not hold that such comparatively ‘higher-level’ sciences or ‘special sciences’ as neurophysiology, chemistry, or biology would be put out of business *as sciences* even by an ideally successful ontological reduction-by-identification of the *objects* of those sciences with systems of the sorts of objects treated in atomic physics. Our predictions and projections in terms of the empirical concepts and ‘stances’ (to borrow Dennett’s useful notion) of the higher level sciences and of the ‘manifest image or man-in-the-world’ will have enormous utility insofar as the relevant empirical generalizations are approximately true, well-founded phenomena. But Sellars’ strong scientific realist contention is that the nature and extent of this approximation to the truth of these ‘predecessor’ generalizations is ultimately adequately explained only by means of their ontological identification with their theoretical successor generalizations. For further discussion see J. O’Shea (2007, in press), chapter two.

There is a profound truth in this conception of a radical difference in level between man and his precursors. The attempt to understand this difference turns out to be part and parcel of the attempt to encompass in one view the two images of man-in-in-world which I have set out to describe. For, as we shall see, this difference in level appears as an irreducible discontinuity in the *manifest* image, but as, in a sense requiring careful analysis, a reducible difference in the *scientific* image. (PSIM 6)

Suppose we grant for present purposes that Sellars is right about the essentially normative and holistic nature of conceptual thinking as expressed in the first paragraph (these are of course controversial positions in their own right). And suppose we also set to one side the interesting remark on the holistic ‘jump’ that Sellars says it is ‘difficult to avoid’ concluding was involved in the evolutionary transition from pre-conceptual patterns of behaviour to conceptual thinking proper.<sup>4</sup> I want to focus instead on the distinction made in the second paragraph concerning the ‘difference in level’ between normative conceptual thinking and non-normative, non-conceptual patterns and processes.

Given what we have already seen, it is not difficult to understand why Sellars holds that this difference in level is conceived, on the one hand, as an “*irreducible discontinuity*” within the manifest image conception of ourselves as persons-in-the-world. But what does he mean by asserting that this ‘difference in level’ is, “in a sense requiring careful analysis, a reducible difference in the *scientific* image”? This is the difficult and (I suggest) important distinction that I want to reflect upon: namely, the general idea that the ‘difference in level’ between the normative space of conceptual thinking on the one hand, and naturalistically describable non-normative, non-conceptual processes on the other, is in one sense an irreducible difference, but in another sense a reducible difference. I shall refer to this in what follows as Sellars’ *irreducibility-cum-reducibility* distinction or position.

There are two general tendencies of interpretation in relation to the issues raised by Sellars’ irreducibility-cum-reducibility position that I want to reject as interpretations of what Sellars himself was up to. My suggestion will ultimately be that Sellars’ own view can capture what ought to be retained from each of these two tendencies.

Firstly, according to what we might call the ‘separating off’ interpretation, Sellars in emphasizing the irreducibility-cum-reducibility distinction is following a perennial line of thinking, one which can be traced from Plato’s *Phaedo* through Kant’s ‘phenomenal/noumenal’ distinction to Wittgenstein’s later philosophy, in distinguishing sharply between normative *reasons* and scientific *causes*. On this interpretation, to put it brusquely, once we occupy the naturalistic explanatory perspective of the scientific

---

<sup>4</sup> Sellars’ artful dodge here – his statement that *the conclusion is difficult to avoid that X* – leaves it as an exercise for the reader to figure out whether Sellars in fact believes after all that we should, or should not, come to the conclusion that X. I suspect that Sellars’ hesitation here is due to the fact that he does not take it to be incumbent upon himself as a philosopher to sort out in advance what is an ongoing matter for scientific inquiry: the question as to the ultimate evolutionary origins of natural languages.

image, the normative dimension of conceptual thinking *qua* conceptual thinking is simply no longer on our radar screen as a proper scientific explanandum.

There is certainly much in Sellars' philosophy that would fit an interpretation that 'separates off' normative reasons from scientifically lawful causes in this sense – in fact, there is all that goes into the 'irreducibility' side of Sellars' irreducibility-cum-reducibility position. However, it is clear from the passage above (i.e., PSIM 6) that the *reducibility* side of Sellars' position is supposed to apply to *the distinction of level itself*; that is, to the distinction between the normative-conceptual and the non-normative or pre-conceptual that constitutes the difference between 'man' and his 'precursors'. It is this distinction that Sellars sees (in a sense to be explained) as a *reducible difference of level* when viewed under the regulative ideal of the final scientific image of the human being. Sellars was clearly more ambitious in his synoptic and naturalistic explanatory aims than those who share the 'separating off' tendency of interpretation would wish to embrace. For Sellars, what is on the radar screens of *both* the manifest *and* the scientific images of 'man-in-the-world' is the nature of conceptual thinking *qua* conceptual thinking, along with the crucial distinction itself between normative reasons and naturalistic causes. This is what gives Sellars' philosophy a unifying explanatory boldness – whatever judgment one might ultimately make of its success or failure – that is lacking in those philosophers who in one way or another want us to rest content with separating off the naturalistic dimension of scientific causes from the normative dimensions of the logical space of reasons.

For broadly similar reasons I do not think that what might be called the 'eliminativist' tendency of interpretation in relation to Sellars' irreducibility-cum-reducibility position can adequately capture what he was up to. An eliminativist interpretation would quite properly focus on Sellars' strong scientific realist view that the sophisticated common sense ontology or manifest image of the perceptible world is in one sense ultimately *strictly speaking false*; that it is in principle if not yet fully in practice to be *replaced* by the explanatorily superior successor ontologies of the emerging scientific image of the world. Sellars' key distinction on this view might be taken to be characterizing the normative space of conceptual thinking as in some sense constituting an 'irreducible' but ultimately *false* conceptual framework; and this framework would be 'reducible' in the sense of being in principle *replaceable* by the ideal scientific picture of ourselves and the world.

However, this global eliminativist tendency of interpretation likewise seems to miss the mark. Sellars comments on his analogy of stereoscopic vision, which he uses to describe the philosopher's task of 'fusing' or integrating the two idealized global images or conceptions of the world into one coherent image, that "the very fact that I use the analogy of stereoscopic vision implies that as I see it the manifest image is not overwhelmed in the synthesis" (PSIM 9). He holds that "man is *essentially* that being which conceives of itself *in terms of the image which the perennial philosophy refines and endorses*" – that is, in terms of the manifest image (PSIM 8). That is the 'irreducible' or '*sui generis*' side of the story. On the other side of the distinction, however, Sellars does indeed want to hold that the ontology of persons as rational agents and conceptual thinkers within the space of reasons is in principle successfully accommodated *within* the comprehensively physicalist ontology of the ideal scientific image of the world. That is the 'reducibility' side of Sellars' position, and this is what needs to be clarified without reaching for either an eliminativist sledgehammer on the one hand, or a pluralist feather on the other.

## Section II

### Sellars' earlier version of the distinction:

#### 'logical irreducibility' cum 'causal reducibility'

Some insight can be gained on the nature of Sellars' central 'irreducibility-cum-reducibility' position by going back a decade from 'Philosophy and the Scientific Image of Man' to a predecessor distinction he had introduced in his 1953 paper, 'A Semantical Solution of the Mind-Body Problem'. Sellars begins that paper with the remark that "the mind-body problem" ... is notoriously a tangle in which all the major puzzles of philosophy can be found" (SSMB ¶1). He then proceeds to use the classic problem in moral philosophy concerning the relationship between 'Ought' and 'Is' to introduce a fundamental distinction between different senses of *reducibility*: namely, between what he calls *logical reducibility* (or irreducibility) and *causal reducibility* (or irreducibility). (Sellars' use of the term 'logical reducibility', given his wide use of the term 'logical', might also appropriately be called 'conceptual reducibility'.) As the article develops it becomes clear that for Sellars this distinction, and all that it involves, is supposed to help us understand the relationship between the intensional conceptual frameworks pertaining to mind, meaning, morals, and the modalities on the one hand, and the extensional ontology of an ideal scientific account of the human-being-in-the-world on the other. It involves ways of articulating the key 'irreducibility-cum-reducibility' distinction that were to remain central to Sellars' thinking throughout his career.

The main focus in this wide-ranging paper was on the question of whether 'mental acts' defined as exhibiting intentionality or aboutness "can be reduced to items which are not mental acts, . . . and if so, in exactly what sense of 'reduced'" (SSMB ¶5). Sellars' ultimate answer is that the mental or intentional is *logically irreducible* yet also *causally reducible* to complex patterns of behavior and brain processes describable within an ideal extensionalist scientific framework. Much of the article is devoted to exploring the nature of this particular way of making Sellars' 'irreducibility-cum-reducibility' distinction.

Let us take Sellars' analogy with the traditional Is/Ought problem first. (It should be noted that as Sellars' distinction will apply to the case of the Is/Ought problem, intentional mental states such as emotions, beliefs, and desires will be assumed to fall unproblematically on the naturalistic, psychological side of the divide. Subsequently Sellars will adjust the relevant distinction to tackle the mind-body problem concerning the nature of intentional psychological states themselves.)

Very briefly, then, as Sellars idealizes the classic Is/Ought dispute, the *ethical non-naturalists* were primarily concerned to defend the idea that 'ought' is *logically irreducible* to 'is' in that the meaning of 'ought' is indefinable or unanalyzable in descriptive terms. As his example of ethical non-naturalism in this sense Sellars uses ethical intuitionism (he mentions Ross and Prichard). By contrast, Sellars describes the ethical naturalist as concerned in the first instance to defend the idea that 'ought' is *causally reducible* to 'is'. The notion of causal reducibility itself, Sellars indicates, has to do with what one has to

appeal to in “a properly constructed causal explanation”. In the context of the Is/Ought problem, he puts it as follows:

If we use ‘ethical assertion’ in such a way that ‘Jones ought to pay his debt’ is an ethical assertion, but ‘Jones feels that he ought to pay his debt’ is not, then we can say that to claim that Ought is causally reducible to Is is to claim that one can give a causal explanation of the history of moral agents without making ethical assertions. (SSMB ¶6)

As we might put it in the material mode, to say that objective moral properties are ‘causally reducible’ to natural properties would be to say that one can give a fully adequate causal explanation of “the history of moral agents” without appealing to any objective moral properties themselves. And in contrast to both the ‘separating off’ and ‘eliminativist’ tendencies of interpretation, this would be to give an adequate causal explanation of the history of moral agents *qua* moral agents.

That the latter is supposed to be so is indicated by Sellars’ next move, in which he suggests that both the ethical non-naturalists and the ethical naturalists tended to assume that logical irreducibility and causal irreducibility necessarily go hand and hand. The ethical intuitionists, for example, who on Sellars’ story were concerned in the first instance to defend the logical irreducibility of moral assertions, felt compelled to argue that “the motive (cause) of conscientious action” must be mediated by ethical beliefs or intuitions that, so to speak, latch on to objective moral properties (SSMB ¶7). As Sellars describes the intuitionist view, “Human thinking on ethical matters is, as [the intuitionists] see it, ultimately grounded in and controlled by objective values and obligations. The existence of moral concepts and beliefs in the human mind [on their view] cannot be accounted for in purely naturalistic terms” (SSMB ¶7). That is, the intuitionists on Sellars’ story felt compelled to deny the causal reducibility of moral assertions in order to preserve their logical irreducibility to naturalistic descriptions. Causally explaining the ‘history of moral agents’, on their view, requires appeal to objective values somewhere in the explanans.

By contrast, the ethical naturalists, whom Sellars suggests were primarily concerned to emphasize the *causal reducibility* of ethical assertions to naturalistic terms, thereby also felt compelled to defend the *logical* reducibility or definability of ethical assertions in terms of such naturalistic descriptions. On both sides, Sellars suggests, logical reducibility or irreducibility was assumed to require causal reducibility or irreducibility.

For his own part, Sellars is of course staking out his own position, on both the ‘Is/Ought’ problem and the Mind-Body problem, as one of *logical irreducibility* yet *causal reducibility* (in an appropriately non-trivial explanatory sense of causal reducibility). Although his concern is not about moral philosophy *per se* in the ‘Semantical Solution’ paper,<sup>5</sup> Sellars does indicate that, on the one hand, he takes himself to be an ethical *non-naturalist* insofar as he defends the logical or conceptual *irreducibility* of assertions of ethical obligation. In this respect he suggests that ‘non-naturalism’ should be understood

---

<sup>5</sup> Sellars cites his article ‘Obligation and Motivation’ in *Readings in Ethical Theory*, Hospers and Sellars (eds), 1952.

in a broader sense than was assumed by the early intuitionists. As he puts it, “if one should use the term ‘Non-naturalism’ to cover any view, whether historically espoused or not, which holds that ethical terms have a cognitive meaning which is not definable in descriptive terms, then, no doubt, it is possible to be a Non-naturalist and yet accept the causal reducibility of Ought to Is” (SSMB ¶8). Thus Sellars is also able to agree with the *ethical naturalists* insofar as he thinks that the details of his account of ethical assertions can demonstrate that they are ‘causally reducible’, in a non-trivial sense, to naturalistic descriptions of particular patterns of socially acquired beliefs, motivations, and behavioral dispositions of moral agents. Not unreasonably he argued that the general structure of such a position would represent a middle way between the ethical naturalist and non-naturalist positions he was discussing.

For reasons soon to emerge, I believe that it is important for any interpretation of Sellars’ overall irreducibility-cum-reducibility position to take at least brief notice of his own views on the nature of normative ‘ought’s themselves, for example the moral ‘ought’. Briefly summarized, Sellars put forward an account of moral ‘ought’ statements in terms of what he called *community intentions* or ‘We shall’ intentions (an idea that was later developed in different ways by Michael Bratman and others). (The ‘shall’ terminology is Sellars’ simplifying technical idiom for intentions and – if the intended time of action is *now* – volitions.) Around this he built a view of morality that was broadly Kantian or universalizing in form, but was teleological and benevolence-based in substance. Here is one summary statement Sellars gave of his moral theory in a Letter to David Solomon:

[The] fundamental intention characterizing the moral point of view has the form, ‘We shall any of *us* do that which (in his/her circumstances) promotes (maximizes) *our* common good’. I have argued that such an intention can be construed as ‘categorically valid’ because sharing such an intention defines what it is to be members of a community.<sup>6</sup> (Letter to David Solomon, June 28, 1976, §15)

Sellars argues that on his view particular ought-judgments such as ‘Jones ought to pay his debts’, are objectively true or false (on the primary sense of ‘true’ as ‘correct semantic assertibility’ that Sellars defends; see *SM* IV.30). On his account, moral assertions are *intersubjectively impartial* relative to whatever community the ‘We shall’ can be regarded as operative over.

I am not concerned here either to analyze or to evaluate Sellars’ account of the moral ‘ought’. For my present purpose his key claim, on the one hand, is that whatever it is that cognitively significant moral assertions *say*, this is something that cannot be said without remaining within the ‘logical space’ of community intentions, so to speak. Moral assertions are held by him to be conceptually or ‘logically’ irreducible in that sense.

Underlying this account, on the other hand, Sellars put forward a naturalistic, causal explanation of the motivational efficacy of ‘shall’ intentions as socially acquired linguistic and psychological dispositions to follow up one’s ‘I shall do A’ intentions and

---

<sup>6</sup> Sellars’ correspondences are available on the ‘Problems from Wilfrid Sellars’ website maintained by Andrew Chrucky: <http://www.ditext.com/sellars/index.html>.

volitions, other things being equal, with the doing of A. Such a social-behavioral account of the origin, the content, and the motivational force of individual intentions, community intentions, and on that basis, of normative ‘ought’ statements, would not itself, as we saw him put it earlier, involve the *assertion* of any ‘ought’ statements (or the appeal to any objective moral properties). Yet this causal account is supposed to be more than just a trivial, subject-changing, or ‘separated off’ scientific explanation of various physical motions. For ideally such a *causal reduction* of normative ‘ought’s would explain specifically and exhaustively those particular psychological dispositions and complex patterns of behavior in which the practise of asserting and obeying intersubjectively valid moral assertions really consists. As I see it, this is the sense in which Sellars himself defends “a position which agrees with the Non-naturalist that Ought is logically irreducible to Is, and yet agrees with the Naturalist that Ought is causally reducible to Is” (SSMB ¶9).

It is important to recognize that Sellars in this sense took there to be available, in principle, a fully adequate naturalistic, ultimately extensionalist account of the nature and force of normative ‘ought’s themselves (and the same will hold, in this sense, for his views on intentionality and meaning as well; cf. SSIS 439). In the Preface to *Science and Metaphysics*, Sellars makes the following remark in this explanatory spirit, in anticipation of his account of normative ‘ought’s in the final chapter of that book:

. . . unless and until the ‘scientific realist’ can give an adequate explication of concepts pertaining to the recognition of norms and standards by rational beings his philosophy of mind must remain radically unfinished business. (*SM* p. x)

(And as we know, for Sellars, the philosophy of mind is “a tangle in which all the major puzzles of philosophy can be found” (SSMB ¶1).) Here it is clear that the adequate scientific or naturalistic explication of human beings’ rational recognition of norms is to be an account of such norms *qua* the norms that they are, and not merely an account of a separate scientific subject matter. Or consider the following pregnant passage from his earlier 1949 article, ‘Language, Rules and Behavior’:

The historically minded reader will observe that the concept of rule-regulated behavior developed in this paper is, in a certain sense, the translation into behavioristic terms of the Kantian concept of Practical Reason. Kant’s contention that the pure consciousness of moral law can be a factor in bringing about conduct in conformity with law, becomes the above conception of rule-regulated behavior. However, for Kant’s concept of Practical Reason as, so to speak, an intruder in the natural order, we substitute the view that the causal efficacy of the embodied core generalizations of rules is ultimately grounded on the Law of Effect, that is to say, the role of rewards and punishments in shaping behavior. The most serious barrier to an appreciation of Kant’s insights in this matter lies in the fact that most discussions in philosophical circles of the motivation of behavior stand to the scientific account (whatever its inadequacies) as the teleological conception of the adjustment of organisms to their environment stands to the evolutionary account. (LRB ¶18, fn. 3)

(In the final sentence Sellars is clearly referring to the *current* inadequacies of such scientific accounts.) In this respect Sellars regarded himself as outlining the basis for a

robustly causal-naturalistic or scientific account of the ultimate nature and force of normative rules. And it is in this non-trivial sense, I suggest, that he understood the normative in general to be *causally reducible*, though *conceptually irreducible*, to the scientific-natural.

I am aware that these last remarks are too sweeping both in themselves and as an interpretation of Sellars on the nature of normative ‘ought’s. But that Sellars was up to something of the kind just described is crucial for understanding his analogous views on the ‘logical irreducibility yet causal reducibility’ of intentionality, meaning, and conceptual thinking in general. For it is this earlier distinction, I suggest, that also underwrites Sellars’ broad claim in the passage from ‘Philosophy and the Scientific Image of Man’ from which we started (PSIM 6), concerning the *manifest image irreducibility* yet simultaneously the *scientific image reducibility* of the holistic framework of conceptual thinking (or the ‘space of reasons’). In the final section I will attempt to spell out in general terms how the overall structure sketched out above in the case of normative ‘ought’s was held by Sellars to apply to the case of meaning and conceptual thinking, along with a brief final remark on the nature of *persons* in the synoptic vision.

### Section III

#### **Mind, meaning, and persons in Sellars’ naturalism with a normative turn**

The basis for Sellars’ general approach to the mind-body problem, in a nutshell, is a normatively characterized conceptual or functional role semantics that is supposed to apply both across natural languages and by analogy to a theoretically posited ‘Mentalese’. The meaning of a linguistic term, on this view, is determined by its role within a wider pattern of ‘language entry’ responses to objects in perception (such as a •this apple is red•); in formal and material inference patterns (such as ‘if  $x$  is red, then  $x$  is colored’); and in ‘language exit’ transitions as described earlier in relation to ‘shall’-intentions (such as an •I’ll take the red one *non*• followed, *ceteris paribus*, by my taking the red one). The normative aspect of this view is that the relevant roles are held to be determined by communally shared implicit norms of usage or linguistic ‘ought-to-be’ rules. Sameness or similarity of meaning (and of thoughts), on this view, is sameness or similarity of normatively constrained functional role. Sellars’ so-called ‘Semantical Solution of the Mind-Body Problem’, in the paper we have been looking at, was in effect an early version of a conceptual role semantics and a functionalist philosophy of mind of this general kind, cleverly disguised behind a convoluted and uninviting discussion of what he tended to characterize as a possible ‘ideal scientific behaviorism’.

What the complex discussion in that early paper does shed light on, however, is what Sellars’ overall position of ‘logical irreducibility cum causal reducibility’ is supposed to amount to in the case of his views on meaning and conceptual content. The general distinction that he appeals to in this early article, and subsequently throughout his career, is between what is *asserted* by a statement – what the statement *says*, what it describes or mentions explicitly – as opposed to further information that is pragmatically *conveyed*, or

implied, or *presupposed* by the statement.<sup>7</sup> So consider the explicitly semantic statement made by one English speaker to another that

‘Es regnet’ (in German) means *it is raining*.

Or using Sellars’ dot-quoting device:

‘Es regnet’s (in German) are •it is raining’s.

On Sellars’ view what this explicitly semantic statement does, roughly speaking, is to call upon one’s antecedent knowledge of English in order to functionally classify the German ‘Es regnet’ as playing a relevantly similar rule-governed role as is played in *our* patterns of linguistic behavior, and in *our* perceptual responses, inferences, and actions, by ‘it is raining’s. Or as Sellars puts it in the ‘Semantical Solution’ paper:

... [A]lthough the use of semantical statements is a correct way to *convey information* about human behavior, semantical statements do not describe human behavior. Thus “‘Es regnet’ uttered by Jones means *it is raining*” does not *mention* biographical facts about the role [of] utterances of ‘es regnet’ in Jones’ struggles with his natural and social environment, even though it is a mode of speech properly designed to convey information of this kind. (SSMB ¶59)<sup>8</sup>

The English speaker’s background familiarity with the use of ‘it is raining’s enables her to take from the meaning statement the information that German speakers, for example, produce ‘Es regnet’s in various pattern-governed ways in relation to cancelled picnics, inferences concerning wet streets, meteorological reports, and so on.

The central idea is that, on the one hand, the latter behavioral and psychological patterns are what they are primarily as a result of the communally shared ‘ought-to-be’ norms or rules that have shaped them (together with whatever more basic representational structures the human animal comes equipped with; see Sellars MEV). On the other hand, however, causal explanations concerning the presupposed uniformities or patterns of linguistic behavior and inner processing themselves could in principle be given in entirely naturalistic, non-normative terms. The only real relations between mind and world, on this view, are the various resulting causal relations and patterns that have come to obtain as a result of the shared norms of linguistic behavior. As Sellars puts it in *Naturalism and Ontology*:

---

<sup>7</sup> In *Making It Explicit* Robert Brandom has devoted considerable attention to developing, within his own framework, this particular aspect of Sellars’ semantics.

<sup>8</sup> The passage as printed in SSMB appears to leave out both the ‘s’ and ‘means’ and also the word ‘of’ (or perhaps ‘plays’).

Thus, the fact that the uniformities (positive and negative) involved in language-entry, intralinguistic and language departure transitions of a language are governed by specific ought-to-be statements in its meta-linguistic stratum, and these in turn by ought-to-bes and ought-to-dos concerning explanatory coherence, constitutes the Janus-faced character of languagings as belonging to both the causal order and the order of reasons. This way of looking at conceptual activity transposes into more manageable terms traditional problems concerning the place of intentionality in nature. (*Naturalism and Ontology*, 1979 (NAO V.64))

And this, in broad terms, is also the sense in which the ‘framework of conceptual thinking’ and the ‘space of reasons’ are supposed to be *causally reducible*, in a non-trivial explanatory sense, within the ideal scientific image of the world.

Clearly the crucial link in this account is a certain conception of socially maintained linguistic norms and of *rule-following* behavior generally. In the article “Truth and ‘Correspondence’” in 1962 Sellars articulated an important meta-principle in this connection concerning the essential role of normative principles in shaping corresponding behavioral uniformities: namely, that the “Espousal of principles is reflected in uniformities of performance” (TC 216). This in effect falls out from the account of ‘ought’s, ‘shall’-intentions, and community intentions discussed earlier. As Sellars puts it in ‘Some Reflections on Language Games’ (1954):

Learning the use of normative expressions involves ... acquiring the tendency to make the transition from occupying the position ‘I ought now to do A’ to the doing of A. This motivating role of ‘ought’ in the first person present is essential to the ‘meaning’ of ‘ought’. (SRLG 350)

Sellars’ further comment upon this meta-principle makes clear connection with the key broader issue concerning *logical irreducibility cum causal reducibility*:

I am not claiming that to *follow* a principle, i.e. act on principle, is identical with exhibiting a uniformity of performance that accords with the principle. I think that any such idea is radically mistaken.<sup>9</sup> I am merely saying that the espousal of a principle or standard, *whatever else it involves*, is characterized by a uniformity of performance. And let it be emphasized that this uniformity, though not the principle of which it is the manifestation, is describable in matter-of-factual terms. (TC 216)

Here again is the idea that the normative principle itself is, on the one hand, conceptually irreducible to any ideal scientific explanation of it in causal-naturalistic terms. But as before, here again he also stresses that the patterns or ‘uniformities of performance’ themselves are in principle describable in purely naturalistic terms, and they are thus explainable *as* the particularly shaped patterns that they are. And we have also seen that

---

<sup>9</sup> Cp. the ‘radical mistake’ that is ‘of a piece with the naturalistic fallacy’ in the passage from EPM quoted at the outset.

on his view the normative principles or ‘ought’s do not themselves generate any ontologically problematic properties for the naturalist, given his particular account of ‘ought’s in terms of certain kinds of shared intentions and desired ends.

The upshot of this overall picture is that our normatively rule-governed linguistic practices both presuppose and systematically maintain a corresponding underlying structure of specific sorts of natural-causal connections between language (and mind) and the world. This norm/nature presuppositional structure, I believe, can be shown to hold across the board for Sellars’ views on the nature of meaning, intentionality, knowledge, and truth.<sup>10</sup> On his account, none of those normative phenomena themselves turn out to be *relations* between language or mind and the world. Yet the rule-governed practices in which those normative phenomena consist both presuppose and themselves generate, by their very nature, specific patterns of natural-causal relations and structures. For Sellars, it is the specific nature of the resulting *non-normative* causal relations and real mind-world isomorphisms that enable our cognitive systems – at least, at the bottom level and when we’ve got things right – to be *mirrors of nature* that correspond to empirical reality.

This is what I take to be the overarching strategy of what might be called Sellars’ ‘naturalism with a normative turn’. The strategy has essentially been one of exposing what *seem* on the surface to be certain puzzling ‘factualist’ or ‘ontological’ questions, perennially seen as requiring the appeal to various problematic primitive relations and quasi-relations to reality, to be in reality various complex questions concerning how our multifarious and projected *rule-governed practices* are related to the *natural-causal uniformities* which they both presuppose and shape. My suggestion has been that this is the general logical structure of Sellars’ attempt to account for the conceptual irreducibility of normative structures within an uncompromisingly scientific-naturalist ontology. What we have seen Sellars argue from the beginning of his career is that this same strategy must also be applied to normative discourse itself.<sup>11</sup>

---

<sup>10</sup> To take just one instance, consider Sellars’ views on the relationship between the normative ‘order of signification’ (or meaning) and the non-normative ‘order of picturing’ (mental/linguistic representations), as illustrated by his discussion of the case of possible android-robots in ‘Being and Being Known’ in 1960. As he sums up that account:

“In this sense we can say that isomorphism *in the real order* between the robot’s electronic system and its environment is a presupposition of isomorphism *in the order of signification* between robotese and the language we speak” (BBK ¶53).

Roughly put, that a complex ‘picturing’ relationship or structural isomorphism – one that is ideally describable in causal-naturalistic terms – has come to obtain between the robot’s inner representations and the objects and events in its environment, will be a presupposition of our semantic interpretation of the robot’s inner symbolic ‘language’ as *meaning* this or that. Working through Sellars’ papers with this basic ‘norm/nature’ presuppositional structure in mind reveals it to be the spinal chord of his overall philosophical system.

(Sellars’ discussion in BBK of the ‘intentional’ framework and the ‘engineering’ framework for interpreting the robot anticipates aspects of Dennett’s distinction between the ‘intentional stance’ on the one hand, and the ‘design stance’ and ‘physical stance’ on the other.)

<sup>11</sup> As he puts it toward the end of the ‘Semantical Solution’ paper:

“The situation is even clearer with respect to normative discourse. Whatever users of normative discourse may be *conveying* about themselves and their community when they use normative discourse, what they are *saying* cannot be said without using normative

This overall picture holds out the prospect of an integration of the normative with the scientific-natural that would enable us to preserve the insights of both those who emphasize the irreducibility of the logical space of reasons and those who work under the explanatory regulative ideal of an all-comprehensive scientific naturalist ontology. And this, after all, was the main goal of Sellars' original philosophical attempt to envision a synoptic, stereoscopic fusion of the manifest and scientific images of 'man-in-the-world' – a project which has subsequently splintered into the perspectives of his 'left wing' and 'right wing' admirers respectively.

Finally, a brief comment on the important question of the resulting place of *persons* within this account of Sellars' synoptic naturalism. What there ultimately *really* is, for Sellars, is so to speak what the ontology of the envisioned ideal scientific image finally says that there is. So in this sense persons, like everything else in nature, are ultimately complex patterns and sequences of micro-physical events (or 'absolute processes', on Sellars' ultimate account; see FMPP lecture III). How is that final ontological vision consistent with the irreducible conceptual unity of the person as a self-conscious, deliberative agent? On the one hand, as Sellars puts it, "the irreducibility of the personal is the irreducibility of the 'ought' to the 'is'" (PSIM 39). Yet, as we should by now expect, he also remarks in the same context that our task is to show "that categories pertaining to man as a *person* who finds himself confronted by standards (ethical, logical, etc.) ... can be reconciled with the idea that man is what science says he is" (PSIM 38).

On this key question as to the irreducible unity of the person, Sellars believed that it was Kant who had the key insight (assuming, of course, Sellars' proposed replacement of Kant's 'things in themselves' with the micro-ontology of the ideal scientific image). The following passage from Sellars' 'Phenomenalism' paper nicely situates the complex question of the ultimate nature of persons within the *irreducibility-cum-reducibility* structure that I have been attempting to clarify:

The heart of the matter is the fact that the irreducibility of the 'I' within the framework of first person discourse . . . is compatible with the thesis that persons can (in principle) be exhaustively described in terms which involve no reference to such an irreducible subject. For the description will *mention* rather than *use* the framework to which these logical subjects belong. Kant saw that the transcendental unity of apperception is a form of experience rather than a disclosure of ultimate reality. If persons are 'really' multiplicities of logical subjects <that is, swarms of micro-particles, etc.>, then unless these multiplicities used the conceptual framework of persons there would be no persons. But the idea that persons 'really are' such multiplicities does not require that concepts pertaining to persons be *analysable into* concepts pertaining to sets of logical subjects. Persons may 'really be' bundles, but the concept of a person is not the concept of a bundle. (PH 101)

---

discourse. The task of the philosopher cannot be to show how, in principle, what is said by normative discourse could be said without normative discourse, for the simple reason that this cannot be done. His task is rather to exhibit the complex relationships which exist between normative and other modes of discourse. It will be noticed that if one combines our assertion of the causal reducibility of Ought to Is, with our account of mentalistic discourse, the ethical naturalist gets everything he can reasonably hope for. Yet the fact remains that what is said by 'Jones ought to pay his debt' could not be said in even an ideal [extensionalist] PMese." (SSMB ¶66)

On Sellars' naturalism with a normative turn, then, the normative conceptual framework of persons, too, is 'logically irreducible' yet 'causally reducible' to the categorial ontology of the ideal scientific image. Making sense of the details of that difficult distinction must be central to any attempt to come to grips with Sellars' quest for a synoptic vision of our own ultimate place in the overall scheme of things.<sup>12</sup> I have also suggested that Sellars' own position may bear fruitfully on more recent and ongoing controversies concerning naturalism and the irreducibility of normative standards.

---

<sup>12</sup> For an alternative and probing account of the role of normativity within Sellars' final synoptic vision, see the final chapter of Willem deVries' *Wilfrid Sellars* (2005). As far as I can see, the different reconstruction I offer here is not inconsistent with the fundamentals of deVries' account. Within the framework sketched here, in O'Shea 2007 I am perhaps able to push the radical (in some respects Feyerabendian) nature of Sellars' scientific realist vision further than deVries might be willing to go, while nonetheless preserving in full the conceptual irreducibility of the 'manifest' framework of thinking, sensing, and practically active *persons*. In the book I emphasize Sellars' radical claim that in the ideal synoptic vision all of the *contents* of our perceivings, inferences, and volitions would be articulated in the language of the ideal scientific image itself. The support for these claims will be found in O'Shea 2007, chapters two, six, and seven.

## Bibliography

- Brandom, Robert (1994), *Making it Explicit: Reasoning, Representing, and Discursive Commitment*, Cambridge, MA and London: Harvard University Press.
- Churchland, Paul (1981), 'Eliminative Materialism and the Propositional Attitudes', *Journal of Philosophy* 78.
- Dennett, Daniel (1987), *The Intentional Stance*. Cambridge, MA: Bradford Books/MIT Press (n.b. page 341 on Sellars).
- deVries, Willem A. (2005), *Wilfrid Sellars*, Chesham, UK: Acumen Press.
- McDowell, John (1994), *Mind and World*, Cambridge, Mass.: Harvard University Press; reissued with a new introduction, 1996.
- Millikan, Ruth (2005), 'The Son and the Daughter: On Sellars, Brandom, and Millikan', ch. 4 of her *Language: A Biological Model*, Clarendon Press, Oxford.
- O'Shea, James (2007, in press), *Wilfrid Sellars*, in the *Key Contemporary Thinkers* series, Blackwell/Polity Press.
- Rorty, Richard (1979), *Philosophy and the Mirror of Nature*, Princeton, NJ: Princeton University Press.
- Rosenberg, Jay F. (1980), *One World and Our Knowledge of It*, Dordrecht, Holland: D. Reidel Publishing Co.

*Works Cited from Wilfrid Sellars* (using the standard abbreviations):

- BBK (1960) 'Being and Being Known', *Proceedings of the American Catholic Philosophical Association*: 28-49. Reprinted in *SPR*.
- EPM (1956) 'Empiricism and the Philosophy of Mind', originally in *Minnesota Studies in the Philosophy of Science*, Vol. I, edited by Herbert Feigl and Michael Scriven (University of Minnesota Press), 253-329. Reprinted in Sellars *SPR*.
- FMPP (1981) 'Foundations for a Metaphysics of Pure Process' (The Carus Lectures) *The Monist* 64: 3-90.
- LRB (1949) 'Language, Rules and Behavior', in *John Dewey: Philosopher of Science and Freedom*, edited by Sidney Hook (The Dial Press): 289-315. Reprinted in *PPPW*.
- MEV (1981) "Mental Events," *Philosophical Studies* 39: 325-45.
- NAO (1980) *Naturalism and Ontology* (Ridgeview Publishing Co.). The John Dewey Lectures for 1973-4. Reprinted with corrections in 1997.

- OM (1952) 'Obligation and Motivation', in *Readings in Ethical Theory*, edited by Wilfrid Sellars and John Hospers (Appleton-Century-Crofts): 511-17.
- PH (1963) 'Phenomenalism', in *SPR*: 60-105.
- PPPW (1980) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, edited and introduced by Jeffrey F. Sicha (Ridgeview Publishing Co.).
- PSIM (1962) 'Philosophy and the Scientific Image of Man', in *Frontiers of Science and Philosophy*, edited by Robert Colodny (University of Pittsburgh Press): 35-78. Reprinted in *SPR*.
- SM (1967) *Science and Metaphysics: Variations on Kantian Themes, The John Locke Lectures for 1965-66* (Routledge and Kegan Paul). Re-issued in 1992 by Ridgeview Publishing Company.
- SPR (1963) *Science, Perception and Reality* (Routledge and Kegan Paul). Re-issued by Ridgeview Publishing Company in 1991.
- SRLG (1954) 'Some Reflections on Language Games', *Philosophy of Science* 21: 204-28. Reprinted in *SPR*.
- SSIS (1971) 'Science, Sense Impressions, and Sensa: A Reply to Cornman' *Review of Metaphysics* 25: 391-447.
- SSMB (1953) 'A Semantical Solution of the Mind-Body Problem', *Methodos* 5: 45-82. Reprinted in *PPPW*.
- TC (1962) 'Truth and Correspondence', *Journal of Philosophy* 59: 29-56. Reprinted in *SPR*.