

Wilfrid Sellars (1912–1989)

JAY F. ROSENBERG

Life and work

Had Wilfrid Stalker Sellars never written an original philosophical word, his accomplishments as an editor would likely be sufficient to earn him a place of honor in the history of postwar analytic philosophy. In 1950, he and Herbert Feigl co-founded *Philosophical Studies*, the first scholarly journal explicitly devoted to analytic philosophy, which they edited jointly until 1971 and Sellars then edited alone for three more years. A year earlier, Feigl and Sellars had already published a seminal anthology, *Readings in Philosophical Analysis*; *Readings in Ethical Theory*, co-edited by Sellars and John Hospers, followed three years later. The “philosophical analysis” represented in these volumes had been transplanted from its origins and early development at Cambridge and Oxford and enriched by generous cross-fertilization from the logical empiricism of an expatriate Vienna Circle, most notably by the work of Rudolf Carnap, and indigenous strains of pragmatism, critical realism, and evolutionary naturalism. From such seeds, the “analytic” style of philosophizing and its agenda of problems grew to dominate American academic philosophy, definitively changing its intellectual landscape. The “logico-linguistic turn” became the new methodological center of philosophical inquiry, and regional philosophies of logic, language, mind, and science first joined and then gradually began to supplant more broadly-conceived traditional metaphysical and epistemological studies, while normative ethical inquiries gave ground to issues in metaethics and moral psychology.

But Wilfrid Sellars in fact wrote many an original and important philosophical word, and so not only helped to stimulate the growth of postwar analytic philosophy, but also became one of its most distinguished and influential practitioners. His academic trajectory took him from studies at the University of Michigan and, as a Rhodes Scholar, at Oxford University, to faculty positions at the University of Iowa, the University of Minnesota, where he served as Chair during the mid-1950s, and Yale University, before he became University Professor of Philosophy and Research Professor of the Philosophy of Science at the University of Pittsburgh, a position which he held from 1963 until his death in 1989. His intellectual trajectory meanwhile carried him from an early period, during which he worked out his fundamental philosophical ideas in over two dozen dialectically-challenging essays, through a middle period characterized by the

development and exposition of a systematic philosophical vision of remarkable scope and depth, into a late period of consolidation, refinement, and deepening of mature theses and insights that were simultaneously coming to be more fully appreciated and explicitly appropriated by a new philosophical generation.

The critique of givenness

Sellars's revolutionary 1956 essay "Empiricism and the Philosophy of Mind," immediately acknowledged as a contemporary classic, marks the beginning of his exceptionally productive and influential middle period. (This can be treated, somewhat arbitrarily, as culminating in 1972 with the publication of his 1970 American Philosophical Association Eastern Division Presidential Address on the Kantian text, "this I or he or it (the thing) which thinks.") The central theme of "Empiricism and the Philosophy of Mind" is a thoroughgoing and general critique of what Sellars famously called the "myth of the given," a perennial and polymorphic philosophical motif manifested *inter alia* in the idea, characteristic of classical sense-datum theory, that empirical knowledge rests on a foundation of "immediate awarenesses" and on the assumption that the "privacy" of the mental and one's "privileged access" to one's own mental states are primitive features of experience, logically and epistemologically prior to all intersubjective concepts pertaining to inner episodes.

Sellars criticized sense-datum and other traditional epistemic foundationalisms for failing properly to distinguish non-conceptual states of sensing from conceptually structured perceptual takings. Perceiving always involves taking something sensorily present *to be* this or that, and so, as Kant recognized, has a judgmental form which mobilizes and applies, correctly or incorrectly, descriptive and classificatory concepts. Perception properly so-called is consequently a normative business, and the ability to engage in it requires more than reliable differential dispositions to respond to sensory stimuli. "The essential point," Sellars wrote, "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" (1963a: 169). A perceptual judgment may consequently be *direct* or *immediate* in the sense of being a unmediated response to stimulation, i.e. not itself inferred from other judgments, but the *epistemic authority* of such judgments depends upon their being appropriately caught up in the intersubjective game of having and giving reasons, and so cannot be independent of their inferential relationships to other judgments. Sellars thus notoriously advocated a strong *epistemic internalism*, according to which "observational knowledge of any particular fact . . . presupposes that one knows general facts of the form *X is a reliable symptom of Y*" (1963a: 168).

Epistemic authority

Consistent with his strong internalism, Sellars interpreted even first-person epistemic authority with respect to the sensory aspects of one's own experience as built on and presupposing an intersubjective status for sensory concepts *per se*. Correlatively, he decisively rejected the idea that sensory consciousness supplies a form of empirical knowledge that (1) is immediate (i.e. non-inferential); (2) presupposes no knowledge of other

matters of fact, particular or general; and (3) constitutes the ultimate court of appeals for all factual claims (1963a: 164). Although a person can justifiably believe an empirical truth without having inferred it from other propositions, no empirical beliefs are self-justifying, self-warranting, or self-authenticating. Instead, Sellars argued, “to say that someone directly knows that-p is to say that his right to the conviction that-p essentially involves the fact that the [thought] that-p occurred to the knower in a specific way” (p. 188).

The epistemic authority of a non-inferential perceptual belief can be traced to the fact that, in the course of learning perceptual language, the believer has not only acquired propensities for the reliable use of the relevant concepts in perceptual situations but also has come to know what is involved in learning to use perceptual sentences reliably in perceptual contexts. Thus, assuming that he has mastered the use of the relevant words in suitable perceptual situations, a person who candidly and spontaneously thinks-out-loud “Lo! Here is a red apple” – Sellars’s customary model of a perceptual taking – is justified in reasoning:

I just thought-out-loud “Lo! Here is a red apple” (no countervailing conditions obtain); so, there is good reason to believe that there is a red apple in front of me. (1975d: 341–2)

This “trans-level” reasoning, as Sellars called it, does not have the original perceptual judgment as its conclusion, but is rather an inference from the character and context of the original non-inferential experience to the existence of a good reason for accepting it as veridical.

Central to Sellars’s thoroughgoing epistemic internalism, indeed, is his conviction that the reasonableness of accepting even *first principles* is a matter of the availability of good arguments warranting their acceptance. What is definitive of first principles (FP) is the unavailability of sound arguments in which they are derived as conclusions from still-more-basic premises, that is, arguments of the form:

(A1) $P_1, P_2, \dots, P_n \vdash \text{FP}$.

Here, too, Sellars appeals to the notion of a “trans-level” justificatory inference, pointing out that the absence of good arguments of the form (A1) is entirely compatible with the existence of sound arguments of the form

(A2) $P_1, P_2, \dots, P_m \vdash$ It is reasonable to accept FP.

Sellars interpreted the conclusion of (A2) as a claim to the effect that a particular course of epistemic *conduct*, accepting the principle FP, can be supported by adequate reasons. This suggested the in-principle availability of yet another argument, specifically a piece of sound *practical* reasoning, whose conclusion expresses an intention to engage in such conduct:

(A3) I shall achieve desirable epistemic end E.
 Achieving E implies accepting principles of kind K.
The principle FP is of kind K.
 Therefore, I shall accept FP.

On Sellars’s view, such patterns of practical reasoning also govern the warranted acceptance of lawlike generalizations and theoretical systems. Adopting a systematic

theoretical framework is ultimately justified by appeal to the epistemic end of “being able to give non-trivial explanatory accounts of established laws” (1975c: 384), and adopting nomological claims that project the observed statistical frequency (including a frequency = 1) of some property in an open class to further unobserved finite samples of the class, is ultimately justified by appeal to the epistemic end of “being able to draw inferences concerning the composition with respect to a given property Y of unexamined finite samples . . . of a kind X, in a way which also provides an explanatory account of the composition with respect to Y of the total examined sample, K, of X” (1975c: 392). What is crucial is that these ends concern

the realizing of a logically necessary condition of being in the framework of explanation and prediction, i.e. being able to draw inferences concerning the unknown and give explanatory accounts of the known. (1975c: 397)

Sellars argued, in short, that inductive reasoning does not need to be vindicated, that is, shown to be truth-preserving, but is rather itself fundamentally a form of vindicative reasoning, justifying our engaging in determinate epistemic conducts. Its ends-in-view must consequently be of a sort that can be known to be realized or obtain. Unlike such Reichenbachian ends as being in possession of limit-frequency statements which are within a certain degree of approximation of the truth (where such limits exist), the ends of being in possession of explanatory laws and principles envisioned by Sellars satisfied that constraint.

Self-knowledge

Sellars famously engaged the Cartesian picture of direct and incorrigible *self*-knowledge with his “myth of Jones,” a story set in a fictional community whose members speak a hypothetical sophisticated “Rylean language.” The fundamental descriptive vocabulary of this language pertains to public spatiotemporal objects, and while it includes logical operators, subjunctive conditionals, and even the fundamental resources of semantic discourse (enabling its speakers to say of their peers’ utterances that they mean this or that, stand in various logical relations to one another, and are true or false), it nevertheless lacks any resources for speaking of inner episodes, whether thoughts or experiences.

In this community, then, a genius, Jones, develops a theory according to which overt utterances are but the culmination of a process which begins with certain inner episodes. . . . [His] model for these episodes which initiate the events which culminate in overt verbal behavior is that of overt verbal behavior itself. (1963a: 186)

Jones’s theory earns its epistemic credentials by enabling him, and his fellow Ryleans who master it, successfully to explain and anticipate intelligent behavior conducted in silence, that is, unaccompanied by overt verbal episodes of the sort that we would recognize as expressing an agent’s conduct-rationalizing beliefs and desires.

The new idioms of Jones’s theory, for example, “is thinking ‘. . .,’” initially have a purely theoretical use, being ascribed on the basis of inferences from observable

behavior in observable circumstances. But crucially, Sellars argued, what begin as purely theoretical terms can *acquire* a first-person reporting role. For it turns out to be possible for Jones to train his compatriots, in essence by a process of Skinnerian operant conditioning, to respond reliably and directly (i.e. non-inferentially) to the occurrence of such an “inner episode” with a judgment to the effect that it is occurring. That is, they can respond to one thought with a second (meta-)thought to the effect that they are thinking it; this matches the *de facto* phenomenology of first-person “privileged access” sufficiently to account for the Cartesian illusion of mental transparency. The Jonesean story thus shows how the essential intersubjectivity of language can be reconciled with the “privacy” of inner episodes.

Scientific realism

Sellars’s novel appeal to forms of theoretical reasoning in his myth of Jones reflected his broader philosophical concern with the nature, structure, and role of theories in the natural sciences. On the received, positivist, view explanation was identified with derivation. Singular matters of empirical fact were to be explained by deriving descriptions of them from (“inductive”) empirical generalizations (along with appropriate statements of initial conditions), and these “empirical laws” in turn were to be explained by deriving them from theoretical postulates and correspondence rules. On the positivist view, in consequence, theories (e.g. microtheories) explain observational matters of fact only indirectly, by implying the (observation-language) generalizations that explain them directly.

Sellars, in contrast, argued that this “levels picture” of theories was fundamentally misleading. Theories do not explain laws by entailing them. Rather, “theories explain laws by explaining why the objects of the domain in question obey the laws that they do to the extent that they do” (1963c: 123).

[That is,] they explain why individual objects of various kinds and in various circumstances in the observation framework behave in those ways in which it has been inductively established that they do behave. Roughly, it is because a gas is . . . a cloud of molecules which are behaving in certain theoretically defined ways, that it obeys the empirical Boyle–Charles Law. (1963c: 121)

On Sellars’s view, then, “theoretical entities” are not merely convenient fictions, enabling us to abbreviate complicated and unwieldy stories about entities that we have good, (observational) reasons to believe actually exist. Theoretical entities are rather those entities we justifiedly believe to exist for good and sufficient *theoretical reasons*.

Sellars thus advocated a robustly realistic epistemology of scientific inquiry and, correlatively, an understanding of its ultimate outcomes as ontologically definitive: “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” (1963a: 173).

(This is his “*scientia mensura*.”) Scientific theories indeed explanatorily “save the appearances,” but they do so precisely by describing the reality of which the appearances are appearances. This robust realism, combined with a thoroughgoing

naturalism, in fact set Sellars's metaphilosophical agenda for postwar analytic philosophy *per se*.

Metaphilosophical views

Sellars saw contemporary philosophy as confronted by two "images," each of which purported to be a complete picture of man-in-the-world, which need to be unified into a single synoptic vision. The "manifest image" is the conception of the world and the place of persons in it that has descended from the great speculative systems of ancient Greece, through the dialectics of a "perennial philosophy," to the dimensions of contemporary Anglo-American thought that emphasize "ordinary usage" and "common sense." Its primary objects are persons, beings who, *inter alia*, reflectively conceive of themselves as being in the world both as sentient perceivers and cognitive knowers of it, and as agents capable of affecting it through deliberate and rational elective conduct.

In contrast, the "scientific image" is the complex new understanding of man-in-the-world that is still in the process of emerging from the fruits of theoretical reasoning, in particular, from the processes of postulational theory construction. Although this image is "methodologically dependent on the world of sophisticated common sense," Sellars argues that

it purports to be a complete image, i.e. to define a framework which could be the whole truth about that which belongs to the image. Thus although methodologically a development within the manifest image, the scientific image presents itself as a rival image. From its point of view the manifest image on which it rests is an "inadequate" but pragmatically useful likeness of a reality which first finds its adequate (in principle) likeness in the scientific image. (1963c: 57)

The leading challenge for contemporary philosophy, he concluded, is to show how the inevitable tensions between these two images can be resolved by a "stereoscopic understanding" in which they come to be "fused" into a single synoptic vision of man-in-the-world. Sellars's philosophy is usefully viewed as a fuller articulation of this confrontation of the two images and a detailed dialectical engagement with the philosophical agenda to which it gives rise: that places be found within the context of a thoroughgoing naturalism for the intentional contents of language and thought, for the normative dimensions of knowledge and action, and for the sensuous contents of perception and imagination. Consonant with such a naturalism, the sought synoptic story must find a place for mind without assigning an independent, autonomous, and irreducible, ontological status to intentional states or entities, and it must eschew any ontological view of abstracta as real that fails to deliver an adequate account of their role within the causal order, broadly construed.

Semantic meaning

The centerpiece of Sellars's response to both of these naturalistic challenges was a sophisticated theory of conceptual roles, concretely instantiated in the conducts of linguistic communities and socially transmitted by modes of cultural inheritance. At the

heart of this theory was an increasingly refined account of meaning as functional classification, more precisely, of the “meaning” idiom as, in the first instance, a context of translation in terms of which structurally distinct “natural-linguistic objects” (e.g. utterings or inscribings) are classified in terms of their roles or functions *vis-à-vis* the organized behavioral economies of families of speaking organisms. In short, Sellars interpreted “means” as a specialized form of the copula, tailored to metalinguistic contexts, according to which the right side of the superficially relational form “– means . . .” is also properly understood as mentioning or exhibiting a linguistic item.

Ordinary quotation, argued Sellars, suffers from a systematic ambiguity regarding the criteria according to which linguistic tokens are correctly classifiable as belonging to this or that linguistic type. He therefore introduced the straightforward device of two separate styles of quotation marks – star-quotes and dot-quotes – to differentiate between two different ways of sorting and individuating lexical items. Star-quotes form common nouns that are true of items belonging to the spatiotemporal causal order (“tokens”) which are appropriately *structurally* isomorphic to the tokens exhibited between them, while dot-quotes form common nouns true of tokens that, in some specified language, are appropriately *functionally* isomorphic to (i.e. play the role or do the job performed by) the tokens exhibited between them in *our* language. Sellars then proposed to transcribe such semantic claims as

- (1s) (In French) “rouge” means *red*,
and
(2s) (In German) “Es regnet” means *it is raining*,

by the more perspicuous formulations

- (1*) (In the French linguistic community) *rouge*s are ‘red’s
and
(2*) (In the German linguistic community) *Es regnet*s are ‘it is raining’s.

Roles and rules

To sort and classify descriptively individuated families of natural items in terms of their linguistic jobs, roles, or functions, is to sort and classify them *normatively*. Sellars’s overall philosophical agenda thus required a complementary thoroughly naturalistic account of the normative dimension of language, and he indeed offered one, basing it on the notion of what he called *pattern-governed behavior*. The basic concept of pattern-governed behavior is

the concept of behavior which exhibits a pattern, not because it is brought about by the intention that it exhibit this pattern, but because the propensity to emit behavior of the pattern has been selectively reinforced, and the propensity to emit behavior which does not conform to this pattern selectively extinguished. (1974: 423)

Pattern-governed behavior can arise from processes of natural selection on an evolutionary time-scale as a characteristic of a species, for example, the dance of the bees,

but it can also be developed in individuals, “trainees,” by deliberate and purposive selection on the part of other individuals, the “trainers.” In this connection, Sellars introduced a distinction between “rules of action” and “rules of criticism.”

Rules of action specify what someone ought to do, for example, “*Ceteris paribus*, one ought to say such and such if in circumstances C.” They can be efficacious in guiding *linguistic* activity only to the extent that their subjects already possess the relevant concepts, such as concepts of “saying such-and-such,” of “being in circumstances C,” and, indeed, of obeying a rule (i.e. doing something because it is enjoined by a rule). Rules of criticism, in contrast, specify what ought to be the case, for example, “Westminster clock chimes ought to strike on the quarter hour” (1975d: 95). The items whose performances may legitimately be appraised according to such rules, however, need not themselves have the concept of a rule nor, indeed, even be capable of having any concepts at all.

Thus trainers can be understood in the first instance as acting in accordance with rules of conduct whose authority derives from rules of criticism, that is, as aiming at bringing about the pattern-governed behaviors which their trainees’ conduct ought to manifest: “Patterned-behavior of such and such a kind ought to be exhibited by trainees, hence we, the trainers, ought to do this and that, as likely to bring it about that it is exhibited” (1974: 423).

If training is successful, then, in consequence of the conducts of trainers under the guidance of such rules of action, the behavior of a language-learner can come to conform to the relevant rules of criticism without his, in any other sense, grasping them himself.

[The] members of a linguistic community are first language learners and only potentially “people,” but subsequently language teachers possessed of the rich conceptual framework this implies. They start out being the subject matter of the ought-to-be’s and graduate to the status of agent subjects of the ought-to-do’s. (1975d: 100)

The modes of pattern-governed behavior relevant to semantic meaning and, correspondingly, the relevant families of rules of criticism, Sellars proposed, are threefold:

- 1 Language Entry Transitions: It ought to be the case that speakers respond to objects in perceptual situations and to certain states of themselves with appropriate linguistic-conceptual activity.
- 2 Intra-linguistic Moves: It ought to be the case that speakers’ linguistic-conceptual episodes tend to occur in patterns of valid inference (theoretical and practical, formal and material), and tend not to occur in patterns which violate logical principles.
- 3 Language Departure Transitions: It ought to be the case that speakers respond to such first-person linguistic-conceptual episodes as “I shall now raise my hand” with an upward motion of the hand, etc. (Cf. 1974: 423–4)

These transitions are respectively the core elements of perceptual takings, inferences, and volitions, and, although they are acts (*qua* both actualities and actualizations of behavioral dispositions), Sellars insisted that they are not themselves actions. They are acquired as, and remain, pattern-governed activities, but form the basis of linguistic

actions proper as “the trainee acquires not only the repertoire of pattern-governed linguistic behavior which is language about non-linguistic items, but also that extended repertoire which is language about linguistic as well as non-linguistic items” and thus becomes “able to classify items in linguistic kinds, and to engage in theoretical and practical reasoning about his linguistic behavior” (1974: 425).

Linguistic roles or functions are individuated in terms of the structure of positive and negative uniformities generated in the natural order by these pattern-governed activities of perception, inference, and volition. Analyzing sameness of linguistic role as sameness of place in the complex relational structure generated by conducts that have been causally shaped in these ways by systems of espoused linguistic norms equipped Sellars with a functional conception of semantics. This conception neither presupposed nor unavoidably led back into the domains of either ontological abstracta or irreducibly intentional mental entities.

The intentionality of thought

Instead, Sellars’s alternative account of the distinctive intentionality of thought was itself drawn in terms of the forms and functions of natural linguistic items, modeled by what he came to call “verbal behaviorism” (VB):

According to VB, thinking “that-p,” where this means “having the thought occur to one that-p,” has as its primary sense [an event of] saying “p”; and a secondary sense in which it stands for a short term proximate propensity [disposition] to say “p.” (1974: 419)

The roots of Sellars’s mature verbal behaviorism reach back to the myth of Jones in “Empiricism and the Philosophy of Mind.” The thought-episodes postulated by Jones on the model of overt verbal behavior are introduced by a purely functional analogy. The concept of an occurrent thought is not that of something encountered *propria persona* but rather that of a causally-mediating logico-semantic role-player, whose determinate ontological character is initially left entirely open.

Since on Sellars’s account the concept of a thought is fundamentally the concept of a functional kind, no ontological tension is generated by the identification of items belonging to that functional kind with states and episodes of an organism’s central nervous system. The manifest image’s conception of persons as thinkers can consequently fuse smoothly with the scientific image’s conception of persons as complex material organisms having a determinate physiological and neurological structure. Sellars’s conviction that the fundamental characteristic of semantic discourse is its ineliminable appeal to functional considerations, and his correlative pioneering analyses of the intentional categories of the mental in terms of epistemologically theoretical transpositions of the semantic categories of public language earn him the title of the first *functionalist* in contemporary analytic philosophy of mind. This functionalist view is one whose implications and influence have not yet begun to be exhausted.

Categorial ontology

The parallels between semantic discourse and the classical ontological idioms of Platonistic discourse, ostensibly designating abstract entities, have not gone unnoticed.

Consistent with his global commitment to naturalism, Sellars exploited these parallels to construct his own unique variant of linguistic nominalism, a view according to which

the abstract entities which are the subject matter of the contemporary debate between platonistic and anti-platonistic philosophers – qualities, relations, classes, propositions, and the like – are linguistic entities. (1967a: 229)

In first approximation, Sellars proposed to analyze the ostensibly relational ontological claims

(1o) (The French word) “rouge” stands for *redness*,

and

(2o) (The German sentence) “Es regnet” expresses the proposition *that it is raining*,

precisely as he had analyzed the corresponding semantic claims (1s) and (2s). These two will be analyzed as (1*) and (2*), claims that specify the functional roles of determinate families of structurally-individuated tokens. This strategy of understanding traditional ontological discourse as classificatory discourse within a functional meta-language transposed into the “material mode of speech” had been pioneered by Carnap (see CARNAP). But unlike Carnap, Sellars refused to identify the formally definable constructs of a “pure” syntax or semantics with the corresponding syntactical and semantical terms in everyday, pre-philosophical usage. Such a facile interpretation of the relationship between “pure” and “descriptive” syntactic and semantic discourses, he argued, fails to do proper justice to the essential normative dimension of the latter. Thus, while Sellars was prepared to interpret such paradigmatic ontological terms as “universal,” “individual,” “kind,” “quality,” “proposition,” and “fact” by appealing to syntactic and semantic counterparts (e.g. “predicate,” “singular term,” “common noun,” “monadic predicate,” “sentence,” and “true sentence”) he insisted that these syntactic and semantic terms

have a conceptual role which is no more reducible to [non-syntactical and] non-semantical roles than the role of prescriptive terms is reducible to non-prescriptive roles. . . . [The] empirical (in the broad sense) character of statements in descriptive (historical) [syntax and] semantics does not entail that [syntactical and] semantical concepts, properly so called, are descriptive. (1975a: 459)

Like Sellars’s account of the distinctive intentionality of thought, then, his account of discourse ostensibly about the entities and categories of classical ontology is also drawn in terms of the forms and functions of natural linguistic items. “Abstract entities,” too, consequently constitute no obstacle to the sought fusion of the manifest and scientific images.

Sensations

Surprisingly it is when Sellars turns from intentional thought and ontological abstracta to *sensations* that significant complications to his synoptic project first come into view.

Like Kant, Sellars rejected the Cartesian picture of a sensory-cognitive continuum. The “of-ness” of a sensation – e.g. its being of a red triangle or of a sharp shooting pain – he insisted, is not the intentional “of-ness” (“aboutness”) of a thought. “The ‘rawness’ of ‘raw feels’,” he wrote, “is their non-conceptual character” (1967c: 376). Consequently, although Sellars’s *epistemological* story about sensations also begins with a strategic appeal to the unique epistemic status of postulated theoretical entities, his account of the *ontology* of sensations diverges significantly from his semantic and functionalist account of intentional thoughts.

In a final episode of the Myth of Jones the hero . . . postulates a class of inner – theoretical – episodes which he calls, say, impressions, and which are the end results of the impingement of physical objects and processes on various parts of the body. (1963a: 191)

Jones postulates impressions as elements of an explanatory account of the occurrence in various circumstances of perceptual cognitions with determinate semantic contents. In this case, however, the *model* for Jones’s theory is not functionally-individuated families of sentence tokens, but rather “a domain of ‘inner replicas’ which, when brought about in standard conditions share the perceptible characteristics of their physical sources” (1963a: 191). Although the entities of this *model* are particulars, the entities introduced by Jones’s *theory* are not particulars but rather non-conceptual (non-intentional) states of a perceiving subject. Thus, while talk of the “of-ness” of sensations, like that of the “of-ness” of thoughts, is fundamentally classificatory, the classification of sensations is not functional, but rather based on analogies that are initially extrinsic and causal, and ultimately intrinsic and contentive.

In the first instance, then, the concept of a person’s having an of-a-red-triangle sensation – an adjectival idiom contrived to highlight the classificatory role of “of-ness” – or the concept of a person’s sensing [red triangle]_sly – an adverbial idiom contrived to reflect the fact that “sensation” is a “verbal noun” – is the concept of her being in the sort of state that is brought about in normal perceivers in standard conditions by the action of red triangular objects on the eyes. The point of the model of “inner replicas,” however, is to insist that such states can discharge their explanatory jobs in relation to cognitive perceptual takings (especially non-veridical perceptual judgments) only if they are conceived as having themselves determinate intrinsic characters and, in particular, as resembling and differing from other sensory states in a manner formally analogous to the way in which objects of the “replica” model (e.g. “wafers” of various colors and shapes) are conceived to resemble and differ from one another.

Sellars proceeded to develop this core account of sensations in two different directions, in consequence of which it has come to be regarded as one of the most difficult and controversial aspects of his philosophy. The first line of development turned on his conclusion that the fundamental concept pertaining to color within the manifest image is the concept of a kind of stuff. It is the concept of a quantum of red in space, an expanse or volume consisting of red, and is a *basic* concept in the sense that there is “no . . . determinate category prior to the concept of red as a physical stuff, as a matter for individuated physical things” (1981, I: 84). When dialectical pressures generate worries about the ontological status of the red which one ostensibly sees when it is *not* a constituent redness of a physical object, however, no alternative category can simply

be “read off” from an introspective scrutiny of color quanta. The idea that a person is always also aware of the actual categorial status of the items that he encounters in perception or introspection, Sellars suggests, is only a particularly pernicious form of the myth of the given.

All that is available is such transcendentals as “actual,” “something” and “somehow.” The red is something actual which is somehow a portion of red stuff, somehow the sort of item which is suited to be part of the content of a physical object, but which . . . is not, in point of fact, a portion of physical stuff. (1981, I: 90)

It then becomes the job of analogical thinking to construct new categorial forms of concepts pertaining to color.

The first complication of Sellars’s account of sensation resulted from his conviction that, in the case of sensations, Jones’s theory takes this *interpretive* form. It does not introduce new domains of entities, but rather new forms of concepts.

[Jones’s] theory of sense impressions . . . reinterprets the categorial status of the cubical volumes of pink of which we are perceptually aware. Conceived in the manifest image as, in standard cases, constituents of physical objects and in abnormal cases, as somehow “unreal” or “illusory,” they are recategorized as sensory states of the perceiver and assigned various explanatory roles in the theory of perception. (1981, III: 44)

The crux of the Jonesean theory, in other words, is the thesis that the very color quanta of which we are perceptually aware as being in space are instead actually states of persons-qua-perceivers. It follows that, already within the manifest image, the ontological status ultimately accorded to sensory “content qualia” is in fact incompatible with their actually being instantiated in physical space. “[The] *esse* of cubes of pink is *percipi* or, to use a less ambiguous term, *sentiri*. Of course . . . we are not perceptually aware of [them] as states of ourselves, though that is in point of fact what they are” (1981, III: 66).

Absolute processes

The second complication of Sellars’s account of sensations then arose from his further conclusion that the scientific image’s commitment to the idea that perceivers are complex systems of micro-physical particles constitutes a barrier to any straightforward synoptic assimilation of this manifest image conception of sensory contents as states of perceiving subjects. On the one hand, Sellars observed that Jones’s analogical treatment of sensory contents as states of perceivers formally preserves the “ultimate homogeneity” of those contents as originally conceived as space-filling stuffs. No defined states of a system or multiplicity of logical subjects, he argued, could continue to do so.

On the other hand, Sellars contended, we cannot simply adopt a “reductive materialist” view according to which “what really goes on when a person senses a-cube-of-pinkly consists in [a certain] system of micro-physical particles being in a complex physical-2 state” (1981, III: 79), where “physical-2” states are those definable

in terms of theoretical predicates necessary and sufficient to describe non-living matter. (To be “physical-1,” in contrast, is simply to belong in the space-time network.) For such reductive materialism amounts to the rejection of the idea that a (Jonesean-theoretical) state of, for example, sensing a-cube-of-pinkly is itself something actual in any categorial guise, and this fails properly to respect the philosophical demands of an adequate sensory phenomenology.

What Sellars notoriously concluded was that sensory contents could be synoptically integrated into the scientific image only if *both they and* the micro-physical particulars of that image as well were subjected to yet another ontological reconception in terms of a categorially-monistic framework whose basic entities were all “absolute processes.” Only when perceivers themselves had been reconceived as systems or “harmonies” of absolute processes, including the ultimate conceptual descendants of sensory contents, would the way be cleared for a unification of the two images. Thought of as absolute processes, sensings would be physical

not only in the weak sense of not being mental (i.e. conceptual), for they lack intentionality, but in the richer sense of playing a genuine causal role in the behavior of sentient organisms. They would . . . be physical-1 but not physical-2. Not being epiphenomenal, they would conform to a basic metaphysical intuition: to be is to make a difference. (1981, III: 126)

Intention and action

In contrast to the integrative challenges posed by thoughts and sensations, the challenge of integrating *actions* properly so called, that is, conducts informed by intention and volition, into the scientific image is not fundamentally ontological. Although they exhibit quite special features when considered functionally, regarded from the ontological perspective, intentions and volitions are simply species of occurrent thought-episodes. What makes such thoughts *practical* are their special relationships to conduct or behavior, analogous to the way in which their status as non-inferential responses to sensations confers on particular thoughts the functional role of perceptual judgments.

Sellars characteristically signals the special conduct-determining role of practical thoughts by a contrived use of the auxiliary verb “shall” as an operator on thought contents expressed as sentences. Categorical intentions are temporally determinate first-person future-tensed practical thoughts. They have the canonical form (INT): Shall (I will do X at t). Volitions (“acts of will”) are special cases of such intentions, whose time determination is the indexical present, thus, (VOL): Shall (I will *now* do X).

Such practical thinkings, on Sellars’s view, mediate between deliberative reasoning and overt behavior by being appropriately caught up in a network of acquired causal propensities that guarantee, roughly, that intentions of the form (INT) regularly give rise, at time t , to volitions of the form (VOL), which, absent paralysis and the like, in turn regularly give rise, then and there, to bodily movements that are (further circumstances being appropriate) the initial stages of a doing of X . Such practical thinkings are governed according to a single principle which unites practical and theoretical reasoning: If $\lceil p \rceil$ implies $\lceil q \rceil$, then $\lceil \text{Shall}(p) \rceil$ implies $\lceil \text{Shall}(q) \rceil$.

Persons

Here again, then, Sellars concluded, what the manifest image contains is not the concept of something with a determinate intrinsic character with which we are acquainted but rather the functional conception of a causally-mediating logico-semantic role-player. Practical thinkings can thus be ontologically accommodated within the scientific image along the lines already sketched for cognitive thoughts in general. But here, he continued, ontological accommodation cannot be the end of the synoptic story. If we take seriously the idea that the scientific image purports to be a complete image of man-in-the-world and a candidate ultimately to replace the manifest image, then the latter's categories pertaining to *persons* will need to reappear within the sought synoptic fusion as such. We need to reconcile "the idea that man is what science says he is" with "the categories pertaining to man as a person who finds himself confronted by standards (ethical, logical, etc.) which often conflict with his desires and impulses" (1963c: 38).

On Sellars's view, the basic concept of a person is irredeemably social. To think of an entity as a person is essentially to think of it as actually or potentially a member of a community, "an embracing group each member of which thinks of itself as a member of the group" (1963c: 39), and it is the most general shared intentions of its members that fundamentally define the structure of norms and values in terms of which their conducts come to be appraised as "correct" or "incorrect" or "right" or "wrong."

Thus the conceptual framework of persons is the framework in which we think of one another as sharing the community intentions which provide the ambience of principles and standards (above all, those which make meaningful discourse and rationality itself possible) within which we live our own individual lives. (1963c: 39–40)

As we have seen, Sellars interpreted the framework of thoughts as founded within the manifest image on a series of ontologically noncommittal functional analogies to which we can readily imagine an emerging scientific understanding progressively supplying structural (e.g. neurophysiological) form. In contrast, he argued that accommodating the manifest image's sensory contents within a synoptic fusion would require the conceptual transposition of some of its ontologically basic entities into new categorical forms. Unlike the frameworks of thoughts and sensations, however, Sellars contended that the conceptual framework of persons as such "is not something that needs to be reconciled with the scientific image, but rather something to be joined to it" (1963c: 40). To achieve a genuinely synoptic vision of man-in-the-world, he concluded, the scientific image needs to be enriched

not with more [or different] ways of saying what is the case, but with the language of community and individual intentions, so that by construing the actions we intend to do and the circumstances in which we intend to do them in scientific terms, we directly relate the world as conceived by scientific theory to our purposes, and make it our world and no longer an alien appendage to the world in which we do our living. (1963c: 40)

Such an ultimate unification of the manifest and scientific images, the world of persons with the world of science, was the controlling vision of Sellars's philosophy.

What makes him one of the towering figures of postwar analytic philosophy, however, is not just the grand scope of his enterprise, but the profound originality of his specific conclusions, the sophisticated dialectically argued and historically informed reasoning with which he supported them, and the exemplary thoroughness with which he painstakingly developed something very rare in the analytic tradition, a principled and consistent *systematic* philosophical view.

Bibliography of works by Sellars

- 1963a: "Empiricism and the Philosophy of Mind," in 1963d, pp. 127–96. (Formerly published in *The Foundations of Science and the Concepts of Psychoanalysis*, Minnesota Studies in the Philosophy of Science, vol. I, ed. H. Feigl and M. Scriven, Minneapolis: University of Minnesota Press, 1956.)
- 1963b: "Phenomenalism," in 1963d, pp. 60–105.
- 1963c: "Philosophy and the Scientific Image of Man," in 1963d, pp. 1–40. (Formerly published in *Frontiers of Science and Philosophy*, ed. Robert Colodny, Pittsburgh, PA: University of Pittsburgh Press, 1962.)
- 1963d: *Science, Perception and Reality*, London: Routledge and Kegan Paul, and New York: Humanities Press. (Reissued by Ridgeview Publishing.)
- 1963e: "The Language of Theories," in 1963d, pp. 106–26. (Formerly published in *Current Issues in the Philosophy of Science*, ed. H. Feigl and G. Maxwell, New York: Holt, Rhinehart and Winston, 1961.)
- 1967a: "Abstract Entities," in 1967b, pp. 229–69. Reprinted in *Review of Metaphysics* 16 (1983).
- 1967b: *Philosophical Perspectives*, Springfield, IL: Charles C. Thomas. (Reissued in 2 vols, *Philosophical Perspectives: History of Philosophy* and *Philosophical Perspectives: Metaphysics and Epistemology*, by Ridgeview Publishing.)
- 1967c: "The Identity Approach to the Mind–Body Problem," in 1967b, pp. 370–88. (Formerly published in *Review of Metaphysics* 18 (1965).)
- 1968: *Science and Metaphysics: Variations on Kantian Themes*, London: Routledge and Kegan Paul and New York: Humanities Press. (Reissued by Ridgeview Publishing.)
- 1974: "Meaning as Functional Classification," *Synthese* 27, pp. 417–37.
- 1975a: "Empiricism and Abstract Entities", in 1975b, pp. 245–86. (Formerly published in *The Philosophy of Rudolph Carnap*, ed. P. A. Schilpp, La Salle, IL: Open Court, 1963.)
- 1975b: *Essays in Philosophy and its History*, Dordrecht: Reidel.
- 1975c: "Induction as Vindication," in 1975b, pp. 367–416. (Formerly published in *Philosophy of Science* 31 (1964).)
- 1975d: "The Structure of Knowledge," in *Action, Knowledge, and Reality: Studies in Honor of Wilfrid Sellars*, ed. H. N. Castaneda, Indianapolis: Bobbs-Merrill, pp. 295–347.
- 1975e: "Language as Thought and Communication," in 1975b, pp. 93–117. (Formerly published in *Philosophy and Phenomenological Research* 29 (1969).)
- 1979: *Naturalism and Ontology*, Reseda, CA: Ridgeview Publishing.
- 1980: *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, ed. J. F. Sicha, Reseda, CA: Ridgeview Publishing. (Also contains a long introductory essay by Sicha and an extensive bibliography of Sellars's work through 1979.)
- 1981: "The Carus Lectures for 1977–78," *The Monist* 64/1. (Citations by lecture and numbered paragraph.)
- 1989: *The Metaphysics of Epistemology, Lectures by Wilfrid Sellars*, ed. P. Amaral, Reseda, CA: Ridgeview Publishing. (Also contains a complete bibliography of Sellars's published work through 1989.)