

The Unity of the Manifest and Scientific Image by Self-Representation *

Keith Lehrer[†]

lehrer@email.arizona.edu

ABSTRACT

Sellars (1963) distinguished in *Empiricism and Philosophy of Mind* between ordinary discourse, which expressed his “manifest image”, and scientific discourse, which articulated his “scientific image” of man-in-the-world in a way that is both central and problematic to the rest of his philosophy. Our contention is that the problematic feature of the distinction results from Sellars theory of inner episodes as theoretical entities. On the other hand, as Sellars attempted to account for our noninferential knowledge of such states, particularly in correspondence with Castañeda, discussed by Lehrer and Stern (2000), he is lead to account of representation of such states that incorporates the states into what Lehrer has called *exemplar representation* (2004, 2011a) and Ismael *reflexive self-description* (2007). What is common to the three accounts, with some differences, is that such states may be function reflexively in selfrepresentation. Our argument is that the elaboration of this account, suggested in Sellars, shows how the discourse of the manifest image can be transformed into the discourse of the scientific image as self-representations of scientific entities.

Sellars (1963) distinguished in *Empiricism and Philosophy of Mind* (EPM hereafter) between ordinary discourse, which expressed his “manifest image”, and scientific discourse, which articulated his “scientific image” of man-in-the-world in a way that is both central and problematic to the rest of his philosophy. Our contention is that the problematic feature of the distinction results from

* I have used the first plural form in this paper to express my indebtedness to David Stern for earlier discussion leading to our earlier joint paper and to Jenann Ismael for more recent discussion leading to this paper, though I alone remain responsible for the final content of it.

[†] University of Arizona.

Sellars theory of inner episodes as theoretical entities. On the other hand, as Sellars attempted to account for our non-inferential knowledge of such states, particularly in correspondence with Castañeda, discussed by Lehrer and Stern (2000), he is led to an account of representation of such states that incorporates the states into what Lehrer has called *exemplar representation* (2004, 2011a) and Ismael *reflexive self-description* (2007). What is common to the three accounts, with some differences, is that such states may function reflexively in self-representation. Our argument is that the elaboration of this account, suggested in Sellars, leads to an explanation, which is central to Sellars, of how this special form of self-representation can explain how the self-represented states can provide an anchor for theory in experience, in the manifest image and the scientific image as well, and an explanation of how the discourse of the manifest image can be transformed into the discourse of the scientific image as self-representations, which give us what Ismael (2007) has called a phenomenal profile, can be decoupled from the discourse of the manifest image and added the discourse of the scientific image. We shall focus on the discourse of color, including appearances and sensations of color, as the test case for our argument that self-representation effects the transformation of the discourse of the manifest image to that of the scientific image into a coherent, though dynamically changing, image of man-in-the-world.

Our argument may strike some as close to a position that Sellars (1963) attacked as the “Myth of the Given”. Self-represented states may be described, and are by Ismael, as *self-presenting*, which is a vocabulary Sellars would have eschewed. However, our notion of self-representation is contained in Sellars in the Castañeda correspondence, cited in Lehrer and Stern (2000) and suggested already in EPM in his discussion of the reporting role of discourse of thoughts and feelings used in first-person reports. So, how can Sellars consistently combine self-representation of inner episodes with his rejection of the Myth of the Given and account of inner episodes originating as theoretical entities? It is important, and useful for our exposition, to clarify the matter at the outset. First of all, the target of the Myth of the Given is a specific claim that the existence of inner episodes *entails* knowledge and conception of them. His argument against the view is that conception and knowledge require the learning of language supplying a conceptual framework of reasons and justification. Since it is possible, and he holds, in fact the case, that inner episodes, sensations, for example, may occur prior to the acquisition of language, the occurrence of such episodes does not logically entail conception

or knowledge that they occur. Our account of self-representation does not conflict with Sellars rejection of the Myth of the Given.

The account of self-representation we advocate and believe to be central to explaining how the manifest image can be transformed into the scientific image, does not have the consequence that the occurrence of the episodes that are self-represented entail conception and knowledge of them. The occurrence of the episodes, sensations of color, for example, does not entail that they are self-represented or that we have knowledge of such states. The self-representation of inner episodes by exemplarizing them or profiling them is a process that uses the states as vehicles of representation reflexively representing themselves. It is possible that a human being could have sensations but lack the capacity to use those sensations as vehicles of representation. Sellars consistently held that such episodes, even if originally introduced in the manifest image to explain behavior within a theory postulating them as theoretical entities, do not by their simple occurrence logically entail representation of them, not even self-representation.

In terms of what Sellars was denying in the “Myth of Agreement”, he is consistent in suggesting a form of self-representation as a method of using inner episodes as reflexive vehicles to represent themselves. We concede, however, a possible area of disagreement. Sellars was convinced that the capacity to represent the world, our own states included, was acquired by learning a conventional language. Since his time, the arguments in favor of innate representational systems, especially by Fodor (1983), makes Sellars insistence on the actual role of language acquisition on the capacity to represent the world highly controversial. However, in fairness to Sellars, the argument against the Myth of Given is an argument against the view that the occurrence of inner episodes logically entails the representation of them. For that argument to succeed all that is required is the logical possibility that language learning should be necessary for acquiring the ability to represent or self-represent inner episodes. We acknowledge that logical possibility, which suffices for us to agree with Sellars. The conditions necessary for realizing the capacity for self-representation may be left open for the purposes of the present discussion, though the insistence of Sellars on the necessity for meta-representation for genuine descriptive use of language may return us to the issue.

We shall need to present that argument again published in Lehrer and Stern (2000) for the thesis that Sellars held the view that, although the

description of inner episodes was originally introduced in the manifest image as inferred theoretical entities, such episodes become self-representational as description of inner episodes takes on a reporting role in which first person reports are reliable and knowledge of the inner episodes becomes non-inferential. Castañeda asked Sellars for an account of how this transition occurs. Lehrer and Stern (2000) argue that Sellars answer is that the inner episodes act like exemplars used to represent themselves, as well as other phenomenal states, in the process of exemplarization. Lehrer (2007) suggested that self-representation, exemplarization, served the purpose of anchoring discourse to experience in response to the Quine-Neurath claim that we must rebuild the discourse of science floating at sea. Ismael (2007), in her writing directly on Sellars argues that once phenomenal states take on this role of self-representation, showing us what properties the episode exemplify by a kind of indexical ostension, the phenomenal profiles of self-representation can be decoupled from one form of descriptive discourse in which it anchors a descriptive map in experience to attach it another descriptive map again securing the anchor.

So both Lehrer and Ismael are implying, Ismael in more direct connection to Sellars whom she discusses, that the transition from description in the manifest image to description in the scientific image connects with the world and our place in the world by self-representation. As we move from the discussion of Eddington's two tables to the discussion of the relationship between inner episodes such as thoughts and sensations to states of the brain as neurological states and their subatomic constituents, we can decouple, in Ismael's terminology, the self-representations of the former and transfer them to the latter. So the self-representation of the same state can be used to show us what a reddish appearance is like in the descriptive discourse of the manifest image as well as being used to show us what red things look like under standard conditions, in that use of discourse, while showing us at later time in scientific understanding what a specific kind of r-brain state is like that occurs. We shall depend, at a certain stage at correlating the reddish appearance with the r-brain state to effect the transition, but once effected, we shall be able to report when we are in the r-brain state from the self-representation of the exemplarized state in our phenomenal profile. It is as though, given the profile, we shall be in position to report, "There it is, the r-brain state", because we have decoupled the self-representation from one form of report to use it for another. Moreover, though the transition may at first be inferential, based on the observed

correlation, it may become noninferential much in the way that we must at first infer where we have to strike a key to type a letter and subsequently perform the action noninferentially. Sellars suggested that we could recondition ourselves to respond to experience with the description in the scientific image instead of the one in the manifest image. That would leave us with a problem, as he would concede, of how we know noninferentially what the r-brain state described in scientific discourse is like when we report the occurrence of it from our experience. The answer, explained by Sellars in the Castañeda correspondence, is that the self-representation of the state anchors the description in our scientific image of the world and ourselves in the world.

Let us now turn to a more detailed argument that Sellars appealed to self-representation to give an account of how reports of inner episodes can constitute noninferential knowledge. Sellars is famous for his thesis in EPM that the language of inner episodes is introduced as a theory to explain behavior. This account has the advantage that the justification of claims about inner episodes, thoughts, for example, have the same structure as the justification of claims about theoretical entities, electrons, for example. This initial idea contains the suggestion that the justification of claims in the manifest image and the conceptual framework of it have the same structure in principle as claims in the scientific image and the conceptual framework of it. Moreover, his appeal to a behavioral theory of conditioned responses giving rise to the functional role of discourse about thoughts and sensations ties the meaning of discourse of inner episodes to behavior in a way that blocks the problem of the justification of claims about the thoughts and sensations of others. Such claims may be defeated, of course, but the meaning of claims about thoughts and sensations of others, inner episodes of others, commits one to considering behavior as evidence for claims of such episodes because such claims obtain their meaning from their role in explaining behavior. Behavior is evidence for claims about inner episodes by virtue of the meaning of the words used to describe them. The problem of our knowledge of other minds becomes the problem of our knowledge of theoretical entities. The unity of our descriptions of inner episodes in the manifest image with our description of theoretical entities in the scientific image seems embedded in the way such discourse is introduced to achieve explanation.

Though this has some plausibility, Sellars confronted an obvious objection, provided an answer, and was pressed to elaborate it by Castañeda, as presented in Lehrer and Stern (2000). Sellars acknowledged the problem. The problem

simply put is that first person reports are reports of noninferential knowledge of inner episodes, thoughts and sensations, for example, that the person does not infer from his behavior. So, the theory that introduces talk about inner episodes, in what Sellars formulates as his myth of Jones, who introduces the discourse of inner episodes to explain behavior, must add something to the role of discourse to explain how someone can know noninferentially about the existence of his own inner episodes. Sellars solution, which he calls the denouement in the saga of Jones, introduces a contrast between theoretical and nontheoretical in the language of thought as follows:

For once our fictitious ancestor, Jones, has developed the theory that overt verbal behaviour is the expression of thoughts, and taught his compatriots to make use of the theory in interpreting each other's behavior, it is but a short step to the use of this language in self-description. Thus, when Tom, watching Dick, has behavioural evidence, which warrants the use of the sentence (in the language of the theory) "Dick is thinking 'p'" (or "Dick is thinking that p"), Dick, using the same behavioral evidence, can say, in the language of the theory, "I am thinking 'p'" (or "I am thinking that p"). And now it turns out – need it have? – that Dick can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe his own behavior. [...] Our ancestors begin to speak of the privileged access each of us has to his own thoughts. *What began as a language with a purely theoretical use has gained a reporting role.* (EPM, 1963, p. 189)

Two observations are in order. The first is that the "need it have?" expression is intended to block the objection that the existence of the inner episode *entails* that Dick or the rest of us have the ability to give self-descriptions of inner states. The obvious suggestion, elaborated in terms of conditioning, which Tom effects, is that it is a contingent matter whether Dick can master the use of the reporting role and has the cognitive ability to obtain noninferential knowledge of the inner episode, of his thought in this case.

The second observation leads us to the objection Castañeda raised in the correspondence (Castañeda and Sellars, 1961–1962), namely, that Sellars has not given us any explanation of how Dick can master the reporting role. Castañeda asks,

What exactly is what Jones reports in the new use of "I am thinking that p"? How is it that he can make correctly such a statement without observing his behavior? It is not easy to see how on your view these questions can be answered, if they are not answered, it is difficult to see exactly what your view is

accomplishing. (Castañeda and Sellars, 1961–1962, letter of March 6, 1961; cf. Lehrer and Stern, 2000, pp. 205–206)

He concludes by suggesting that to complete his account Sellars must fall back into the Myth of the Given and to admitting mental entities whose «*occurrence entails my knowledge of them*» (2000, p. 206). That is the challenge to Sellars. Either explain how the discourse of inner episodes, thoughts, acquires a reporting role as the result of conditioning resulting in noninferential knowledge or acknowledge that it is logically impossible to have the inner episodes, thoughts, without knowing that you have them.

Sellars reply will take us to the theory of self-representation, exemplarization and phenomenal profiles. Here is the core of Sellars reply.

The important thing to note is that the *core* of Dick’s learning to report what he is thinking is a matter of his acquiring a tendency (*ceteris paribus*) to respond to his thought that-p by saying “I am thinking that-p”. Everything hinges on the force of word “respond” in this connection. *It is being used as a technical term borrowed from learning theory.* The following diagram will help clarify matters:

$$\begin{array}{c} MQ_i \rightarrow MV_i \\ \uparrow \\ Q_i \end{array}$$

where Q_i is a thought that-p, MQ_i is a meta-thought “I am thinking that-p” and MV_i is a meta-statement “I am thinking that-p”. (Castañeda and Sellars, 1961–1962, letter of April 3, 1961; cf. Lehrer and Stern, 2000, p. 207)

Sellars insists that the arrows are in the first instance the result of conditioning, and he concludes,

The above type of account explains the “privileged access” a person has to his own inner episodes. For (although worlds are possible in which this is not the case) only the person who has a thought that-p can respond to it [...] with the thought that he has the thought that-p. (Lehrer and Stern, 2000, pp. 207–208).

That is the account the Sellars offers in his reply to Castañeda’s objection that Sellars must offer an account of our self-description that avoids the Myth of the Given.

Brief reflection on the diagram above will reveal that Sellars answer rests on an assumption of the self-representation of inner episodes. He thereby avoids the Myth of the Given and sustains the role of self-representation in the

transition from the manifest to the scientific image. The short argument for self-representation, noted by Lehrer and Stern (2000), is that MQ_i contains Q_i in a generalization that with the use of dot quotes converts Q_i into a sortal, that is, a general representation that uses Q_i as an exhibit of the kind or class of objects it represents. We note the reflexive character of this form of representation, the use of Q_i as an exemplar to represent the class represented. Therefore, self-representation is involved in Sellars explanation of self-description. The vertical arrow in the first line of the diagram indicates the direction of the externalization of the internal representation, reflexive self-description or exemplarization, to the statement, which may be expressed as an utterance in the conventional language.

There are details in this account of self-representation, which are a modification of the original account offered by Lehrer (1996) and contained in Lehrer and Stern. The modifications were evoked by the work of Ismael (2007) and Fuerst (2010). Papineau (2002, 2007) and Lehrer (2000) had suggested that the representation of consciousness states using the conscious state to represent itself, as well as other states and objects, was best explained as something like quoting a word to obtain a representation of the word and then disquoting it to use it in self-representation. This is somewhat misleading in the case of reflexive representation as Ismael and Fuerst argued leading Lehrer to alter his view. The issue is that quotation and disquotation involve two tokens, the one quoted and the one disquoted as is apparent in a simple example from Sellars,

“red” means red

in which quotation and disquotation involve two different tokens, one quoted, the other not, in the sentence. There is no individual that refers to itself in the sentence. In exemplarization, by contrast, an exemplar is being used as an exhibit of a class of things represented and refers reflexively to itself. To take a different example, suggested by Goodman (1968), if I use a piece of cloth as a sample to represent a kind of cloth, Harris Tweed, for example, the sample refers to instances of cloth, and it is itself an instance of that cloth. The one sample is used as an exemplar to represent instances of cloth, and refers to the token of itself in the use of the token. The token loops back onto itself in exemplarization and reflexive self-description. Moreover, there is some security in the token loop of reference that is lost when one token refers to another. We are not arguing that it is logically impossible for the token to fail to represent itself, however. The token becomes representational and

represents what it does because of the way that it is used representationally. The token reflexive loop of self-representation depends on the exercise of a human capacity which, like all such capacities, lacks the logical infallibility of representation assumed by the Myth of the Given. The security of reference obtained from reflexive representation or exemplarization is that success depends only on the process or activity of reflexive representing and exemplarizing.

The correspondence with Castañeda we have discussed is focused on the question of our representation of thoughts and our noninferential knowledge of them. Sellars gave the same account of sense impressions or sense contents. Jones initially introduces them as explanatory entities, indeed, as inner states. In the case of the impression of a red triangle, it is, Sellars writes:

That state of the perceiver – over and above the idea that there is a red and triangular physical object over there – which is common to those situations in which

- (a) he sees that the object there is red and triangular
- (b) the object over there looks to him to be red and triangular
- (c) there looks to him to be a red triangular physical object over there. (EPM, 1963, p. 190)

He continues to say about the theoretical entities introduced by Jones,

[T]he hero of my myth 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 process on the body, [...] the eye. (EPM, 1963, p. 191)

Finally, Sellars adds the reporting role as in the case of thoughts,

Jones teaches his theory of perception to his followers. As before in the case of *thoughts*, they begin by using the language of impressions to draw theoretical conclusions from appropriate premises. [...] Finally, he succeeds in training them to make a *reporting* use of this language.... (EPM, 1963, p. 194)

Once again the myth helps us to understand that concepts pertaining certain inner episodes – in this case *impressions* – can be primarily and essentially *intersubjective*, without being resolvable into overt behavior symptoms, and that the reporting role of these concepts, their role in introspection, the fact that each of us has a privileged access to his impressions, constitutes a dimension of these concepts which is *built on* and *presupposes* their role in intersubjective discourse (EPM, 1963, p. 195).

Thoughts and sense impression are, according to Sellars, introduced in his myth of Jones as theoretical inner episodes. But these entities are real, not fictional, and the impressions like the thoughts, we propose from the above analysis of the correspondence concerning thoughts, become self-representational as they are exemplarized and become reflexive self-descriptions. Sellars insists that it takes a theory to produce, a conceptual framework, to effect the self-representation. He denies that we begin with self-representation of inner episodes as data in search of a conceptual framework. We locate the inner episodes within a conceptual framework. Jones may introduce a conceptual framework of thoughts and impressions, but the question that confronts Dick is how to find his thoughts and impressions in this framework? He needs to find fixed points of reference to locate his thoughts and impressions in the framework. Those fixed points enable Dick to report his location in the conceptual framework of his world, in himself in his world and in his world in himself.

How does self-representation solve the problem of how the framework represents the world? The exemplarized entities can exhibit at one and the same time what they are like, what a red impression is like, for example, and what a red object is like. The impression, Sellars insists, is not an object of perception. It is a state that can be used in representation to represent itself, giving us noninferential knowledge of what it is like when the representation finds a place in a conceptual framework. At the same time, it can be used to represent what a red object is like in the external world, referring at the same time to itself and to something beyond itself. So the inner episode, exemplarized in reflexive representation, shows us both what the external object is like, exhibiting what it is like for us, and how we represent the external object, exhibiting how we represent it. In that way it exhibits to us what we are like as we represent our world.

It is important for understanding Sellars to appreciate that fixed points of self-representation of thoughts and impressions are not the chronological starting points of representation and knowledge. They are found within discourse as we seek fixed points of reference and representation for ourselves. However, having found them when seeking fixed points for discourse and our conceptual framework, they may take us beyond the discourse and framework. For, we may use those self-representations as fixed points, as exemplarized reflexive representations, knowing what the states represented are like and how we use them to represent the world, in the conceptual framework of the

manifest image and beyond that in the scientific image. Self-representations, once found within a conceptual framework to show us what our experience of our world is like may be decoupled, as Ismael (2007) has put the matter, from the initial framework and transferred to another as Lehrer (2011b) formulates the transition.

The coupling and transfer is familiar enough from everyday experiences of art as well as science. Look at a painting of a historical figure, Madame Pompadour by Boucher, for example, notice how she is portrayed in the painting. You have a conception of her in the framework of the painting from your impression and thoughts of the painting. Now the question arises of whether to transfer your impressions and thoughts from the framework of the painting to the historical figure. You know what she is like in the framework of painting even as you turn away from the painting itself. For you know what your thoughts and impressions are like as you carry away your exemplarized reflexive representations of them. Now you may decouple those self-represented states and transfer them to your conception of her in the court of Louis XV. Whether you transfer them would influence how you acted in the court if you were a member of it, and your conception of the historical events in present day.

This familiar reflection about decoupling and transfer of the content of representation shows us how to bridge the gap between the conceptual framework of the manifest image and the conceptual framework of the scientific image. Having located thoughts and impressions in the framework and descriptions of the manifest image, you exemplarize those states into reflexive self-representations and in phenomenal profiles. But having located them in self-representation, you may decouple them from the words “thoughts” and “impressions”. Those states represent themselves whatever words we attach to them. As we transfer those self-representations to the discourse of science, to the discourse of “neural activation”, for example, they connect that discourse with the same fixed points of self-representation in our experience of our world, of ourselves, and of ourselves in our world. They provide an arch of representation connection between the conceptual framework of the manifest image and the scientific image. The unity of two frameworks does not consist of the reduction of one to the other or even the inclusion of one in the other. It consists instead of finding an arch of self-representation that connects the one with the other. Self-representation, exemplarization and reflexive self-description, can be decoupled from

discourse to connect the manifest image with scientific image with a representational arch.

As Sellars closed EPM, he asked concerning the myth of Jones,

But is my myth really a myth? Or does the reader not recognize Jones as Man himself in the middle of his journey from the grunts and groans of the cave to the subtle and polydimensional discourse of the drawing room, the laboratory, and the study, the language of Henry and William James, of Einstein and of the philosophers who, in their efforts to break out of discourse to an *arché* beyond discourse, have provided the most curious dimension of all. (EPM, 1963, p. 196)

The most curious dimension of them all, articulated by Sellars in his correspondence with Castañeda, may be the dimension of self-representation of inner episodes given birth to within the society of intersubjective discourse which, once mature, is free to move in a private or social manner from one form of discourse to another.

Finally, Sellars argues that the efforts to break out of discourse, which may succeed in taking us to self-representation, will not by itself take us to *knowledge* of the self-represented states. Knowledge, even noninferential knowledge of inner episodes, is tied to the framework of others concepts and claims within the framework «placing it in the logical space of reasons, of justifying and being able to justify what one says» (EPM, 1963, p. 169). We conclude with our concurrence revealed in our own writings on knowledge and acknowledge gratefully the precedence of *Empiricism and Philosophy of Mind*.

REFERENCES

- Castañeda, H., & Sellars, W. (1961–1962). *Correspondence between Hector Castañeda and Wilfrid Sellars on Philosophy of Mind*. Ed. by A. Chrucky. Available on-line at: <http://www.ditext.com/sellars/corr.html> Accessed on : April 15th 2012
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fürst, M. (2010). *A Dualist Version of Phenomenal Concepts*. In Draft.
- Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Bobbs-Merrill.

- Ismael, J. (2007) *The Situated Self*. Oxford: Oxford University Press.
- Lehrer, K. (1996). Skepticism, Lucid Content, and the Metamental Loop. In A. Clark, J. Ezquerro & J. M. Larrazabal (Eds.), *Philosophy and Cognitive Science*, Dordrecht: Kluwer, 73–93.
- Lehrer, K. (2000). Meaning, Exemplarization and Metarepresentation. In D. Sperber (Ed.), *Metarepresentation*, a volume of Vancouver Studies in Cognitive Science. New York: Oxford University Press, 299–310.
- Lehrer, K. (2004). Representation in Painting and Consciousness. *Philosophical Studies*, 117(1–2), 1–14.
- Lehrer, K. (2007). Loop Theory: Knowledge, Art and Autonomy. John Dewey Lecture. *Proceedings and Addresses of the American Philosophical Association*, 81(2), 121–136.
- Lehrer, K. (2011a). What Intentionality is Like. *Acta Analytica*, 26(1), 3–14.
- Lehrer, K. (2011b) Art. Self and Knowledge. Oxford: Oxford University Press.
- Lehrer, K., & Stern, D.G. (2000). The “denouement” of “Empiricism and the Philosophy of Mind”. *History of Philosophy Quarterly*, 17(2), 201–216.
- Papineau, D. (2002). *Thinking about Consciousness*. Oxford: Oxford University Press.
- Papineau, D. (2007). Phenomenal and Perceptual Concepts. In T. Alter & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge. New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press, 111–144.
- Sellars, W. (1963). Empiricism and the Philosophy of Mind [EPM]. In W. Sellars, *Science, Perception and Reality*. New York: Humanities Press, 127–196.

