The notes that follow were developed over the years to help my graduate and advanced undergraduate students at the University of Pittsburgh to see their way through the textual trees to the Sellarsian forest. They are meant to provide only a first take on the material, to indicate the most general outlines of the structure of the essay and of the thought behind it. To that end, many philosophically interesting issues and discussions have been brushed past. In particular, I have sedulously avoided discussing genuinely esoteric issues—such as the philosophical significance some have professed to find in the distinction between 'red' paragraphs and 'green' paragraphs. The formulations and characterizations that are provided are not intended to be definitive or authoritative. They aim to provide a place to start in reading this rich and difficult text.

The idea for such a document, and the notes to the concluding sections, had their origins in a handout Rorty circulated for similar purposes when I was a graduate student at Princeton in the '70s. I am grateful to my colleague John McDowell, and to our former student Danielle Macbeth, for many suggestions and improvements. It should be noted, though, that where their comments evidenced substantive disagreements about what Sellars is (and ought to be) saying—concerning in particular the intricacies of 'looks' talk in relation to reports of the presence of secondary qualities, and the various theses and commitments involved in scientific realism—I have stuck to my own readings. The errors that remain, both those of omission and of commission, should be charged to my account alone.

Note: Section numbers of "Empiricism and the Philosophy of Mind" are indicated in square brackets: [36]. On the rare occasions where sections of this guide must be referenced, I use double brackets: [[36]].

Wilfrid Sellars' "Empiricism and the Philosophy of Mind"

Study Guide

Bob Brandom

Part I [1]-[7] : An Ambiguity in Sense-Datum Theories

Section 1: Sellars announces that his project is to attack "the whole framework of givenness". By this he does not mean to be undercutting the distinction between judgments we arrive at noninferentially, paradigmatically through perception, and those that are arrived at as the conclusions of inferences. Indeed one of the positive tasks of the essay is precisely to tell us how to understand noninferential reports without insensibly sliding into the constellation of philosophical commitments Sellars calls "the Myth of the Given". Sense-datum theories, his immediate target, are important only as prominent and influential instances of the appeal to givenness. We will have to learn to recognize such appeals in many less obvious guises.

In these opening sections, the Myth of the Given shows up in the guise of the idea that some kind of non-epistemic facts about knowers could *entail* epistemic facts about them.[1] Epistemic facts about knowers are in the first instance facts about what someone *knows* (though we will come to see that facts about what one merely *believes* are equally 'epistemic' facts in Sellars' sense). One of Descartes' signal innovations was to define the mind in epistemic terms: for a state to be a *mental* state is for *being* in that state to entail *knowing* that one is in that state (transparency, ruling out ignorance) and for *believing* that one is in that state to entail *being* in that state (incorrigibility, ruling out error). The mind is the realm of what is known *immediately*, not just in the sense of noninferentially, but in the stronger sense that its goings-on are *given* to us in a way that banishes the possibility both of ignorance and of error. (Descartes' thought was that if anything is known to us *mediately*, that is, by means of representations of it, then something—some kind of representations—must be known to us *immediately*, on pain of an infinite regress.) Sellars will try to show us that the Cartesian way of talking about the mind is the result of confusion about the distinction between epistemic and non-epistemic items, and the roles they can play in various sorts of explanation.

In its most familiar form, the Myth of the Given blurs the distinction between sentience and sapience. This is the distinction between being aware in the sense of being merely *awake* (which we share with nondiscursive animals—those that do not grasp concepts), on the one hand, and being aware in a sense that involves *knowledge* either by

[1] The discussion of foundationalism in [32] shows that one can still be committed to the Myth of the Given even if one's foundations are conceived as epistemic facts—if the capacity to know those facts is thought of as independent of inferential capacities and hence the acquisition of ordinary empirical concepts.

being a kind of knowledge, or as potentially serving to *justify* judgments that so qualify. The "idea that a sensation of a red triangle is the very paradigm of empirical knowledge" [7], is a paradigm of the sort of conflation in question. The Myth of the Given is the idea that there can be a kind of *awareness* that has two properties. First, it is or entails having a certain sort of *knowledge*—perhaps not of other things, but at least that one is in that state, or a state of that kind—knowledge that the one whose state it is possesses simply in virtue of being in that state. Second, it entails that the capacity to have that sort of awareness, to be in that sort of state, does not presuppose the acquisition of any *concepts*—that one can be aware in that sense independently of and antecedently to grasping or mastering the use of any concepts (paradigmatically through language learning).[2] The conclusion of Sellars' critical argument is that these two features are incompatible: only what is propositionally contentful, and so conceptually articulated, can serve as (or for that matter, stand in need of) a justification, and so ground or constitute knowledge. Davidson expresses a version of this thought with the slogan "nothing can count as a reason for holding a belief except another belief". Sellars' thought is better captured by changing this to "nothing can count as a reason for endorsing a believable except another believable," where believables are the contents of possible beliefs, that is, what is propositionally contentful.[3]

Sellars understands propositional contentfulness, what is epistemic in the sense of being a candidate for knowledge, in terms of role in what he calls "the game of giving

[2] As McDowell puts the point: "The idea of the Given is the idea that the space of reasons, the space of justifications or warrants, extends more widely than the conceptual sphere." [*Mind and World* p. 7] That is, that what is Given can serve as a justification, without its being given requiring the exercise of conceptual capacities.

[3] The emendation allows that propositionally contentful items that are not believings might serve as epistemic justifiers—for instance, that *facts* could play this role.

and asking for reasons". "In characterizing an episode or 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," [36]. To treat something as even a candidate for knowledge is at once to talk about its potential role in *inference*, as premise and conclusion. Because a crucial distinguishing feature of epistemic facts for Sellars is that their expression requires the use of *normative* vocabulary, to treat something as a candidate for knowledge is also to raise the issue of its *normative* status. The Myth of the Given eventually appears as "of a piece with the naturalistic fallacy in ethics"—the attempt to derive *ought* from *is.*[4] This is because talk of knowledge is inevitably talk of what (conceptually articulated propositional contents) someone is *committed* to, and whether they are in various senses *entitled* to those commitments.

Section 2: Here Sellars distinguishes between the act or episode of sens*ing*, on the one hand, and the content of that act, what is sens*ed*, which is called a sense *content*, on the other. When one hallucinates a pink elephant, doing so is sensing, and the sense-content is what makes it an of-a-pink-elephant hallucination, rather than for instance an of-a-green-Norway-rat hallucination. In ordinary perception, the contents sensed must be carefully distinguished from the external objects sensed (which are entirely absent in the case of hallucinations).

[4] "The idea that epistemic facts can be analysed without remainder—even 'in principle'—into nonepistemic facts…is…a radical mistake—a mistake of a piece with the so-called 'naturalistic fallacy' in ethics" [5]. This theme arises very early in Sellars' writing. See for instance "A Semantical Solution to the Mind-Body Problem", reprinted in *PPPW*.

Section 3: Now consider the suitability of sensings of sense contents as foundations of knowledge and justification on the Cartesian model.

The general idea of a foundation for knowledge can be sketched as follows. Our beliefs constitute knowledge only insofar as they are not only true, but *justified*—lucky guesses don't qualify. One claim or belief can justify another to which it is inferentially related. If one is justified in a commitment to the claim that p, and q may be inferred from p, then one may for that reason be justified in a commitment to the claim that q. To say this is to offer a mechanism whereby justification can be inherited. But, the thought is, not *all* commitments that are justified can have inherited that status inferentially from others. There must be some other mechanism for acquiring positive justification status, to give the inheritance mechanism something to pass along. If p_1 inherits its status from p_2, and p_2 inherits it from p_3, and so on, then either:

a) at some point a claim is repeated (some p_n is identical with a p_m for m<n), in which case the 'justification' is circular, or

b) there never is a repetition, in which case an infinite regress arises, in which each p_n has the anomalous status of an unjustified 'justifier', which is not itself justified until an infinite number of other claims have been justified.[5]

The conclusion is that there must be some way of *being* justified without having to *be* justified. We ought to distinguish two senses of 'justification', one indicating a status

[5] This argument is obviously oversimplified in many ways. Of course justifications need not be single statements—but a corresponding dilemma occurs if sets of premises are allowed. The argument also ignores the fact that there is a regress on inferences in many ways analogous to this regress on premises, and that the two sorts of regress can interact in complex and significant ways.

(being justified), and the other making reference to a process (justifying) that can result in possession of the status.[6] Then the conclusion is that there must be some other way of acquiring positive justificatory status besides justifying it in the sense of offering a justification. Besides inferential inheritance, there must also be some noninferential acquisition mechanism for this epistemic status.

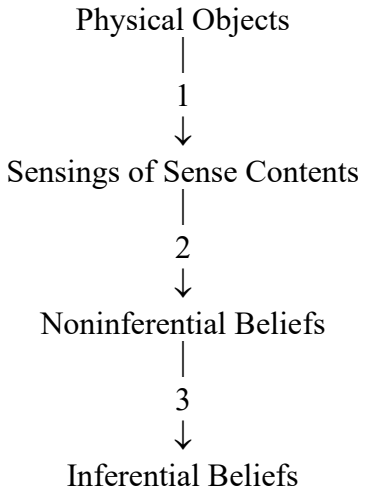
So far, so good. Descartes concluded from this line of thought that there is a kind of claim or belief, call them *basic* beliefs, that form the foundation of all other beliefs in the sense that they are the font from which the justificatory status of all the rest flows inferentially. This does not follow, but Sellars will not contest it.[7] Descartes believed further that unless those beliefs were certain (the ultimate positive justificatory status), none of those inferentially based upon them could even be probable (as C.I. Lewis put it in *Mind and the World Order*). Descartes gave philosophy a decisive epistemic turn which was, at least until Kant, confused with a subjective turn. The latter is a consequence only of Descartes' peculiar and optional way of working out the former. For he defined the mind by its epistemic status, as what is best known to itself by falling within the reach of the subject's incorrigibility and local omniscience. This epistemic definition is what motivates the assimilation of events whose contents are structured like *sentences*, such as thinking that Vienna is a city in Austria, and events whose contents are

[6] This is an instance of what Sellars calls "the notorious 'ing'/'ed' ambiguity" [24]. (See also [[35]]).

[7] See [32]. As it stands, the argument turns on an invalid quantifier inversion. What immediately follows from the foundationalist regress argument is at most that <u>for each</u> chain of justifications <u>there is</u> a belief that *is* justified (has positive justificatory status) without having to *be* justified (by appeal to another belief). It does not follow that <u>there is</u> a kind of belief such that <u>for each</u> chain of justification its terminus is a belief of that kind. A belief that stands in need of justification in one context might serve as an unjustified justifier in another. Compare: <u>For each</u> minute <u>there is</u> a woman somewhere having a baby at that time. This is true. It does *not* follow that <u>there is</u> a woman such that <u>for each</u> minute, she is having a baby at that time. If it *did* follow, then we could solve the problem of overpopulation by finding that woman and making her stop!

structured like *pictures*, such as imagining or seeming to see a red triangle inside a green circle.

To return to the idea of using sensings of sense contents as a foundation of knowledge, then, a process is pictured something like this:



In the standard perceptual case, it is *because* there is a red object with an octagonal facing surface in front of me that I find myself with a sensing of a red-and-octagonal sense content. It is *because* I have such a sense content that I acquire the noninferential belief that there is a red and octagonal object in front of me. And it is *because* I have this belief, together, perhaps, with other beliefs, that I am justified in the further inferential belief that there is stop sign in front of me.

The point to focus on is the nature of the 'because's. The first (arrow 1) can be understood as a causal notion, perhaps the sort studied by students of the neurophysiology of perception. As such, it relates particulars describable in a

nonnormative vocabulary. This is a matter-of-factual, nonepistemic relation. The final 'because' (arrow 3), on the other hand, indicates the sort of relation Sellars calls 'epistemic'. It is an inferential notion, relating sententially structured beliefs (or believables) which are repeatable abstracta—a matter of reasons rather than causes. This justificatory relation is not a natural one, but a normative one; it is not the empirical scientist, but the logician or epistemologist who has the final say about it..

The question is: what sort of relation is the middle one (arrow 2)? Does it belong in a box with the first, causal relation, or in a box with the third, inferential relation? How are the sensings of sense contents to be conceived as related to (potentially foundational) noninferential beliefs? Here is where the distinction between the epistemic and the nonepistemic, between particulars specified in the language of causes and believables specified in the language of reasons, comes into play.

Suppose that one understands the sensing of a sense content to be the existence of a nonepistemic relation between one particular, the sense content, and another, the person doing the sensing. (This is the position Sellars himself eventually endorses.) If so, then it is hard to see how the sensing of a sense content could entail or justify a claim, for instance a noninferential belief. For only things with sentential structure can be premises of inference, not nonepistemically specified particulars. For this reason sensings, understood in terms of nonepistemic relations between sense contents and perceivers, are not well suited to serve as the ultimate ground to which inferentially inherited justification traces back. Since the occurrence of such a sensing does not entail commitment to any claim, it would be possible to have one without coming to believe

anything, and certainly without coming to know anything (for this latter requires positive justificatory status). So it seems the foundationalist who wants to appeal to sensings as foundational must take the sensing of a sense content to be an epistemic fact about the sensing agent. But if so, what becomes of the particular?

Sections 4 & 5: The sense datum theorist can treat sensings as epistemic noninferential beliefs, from which inferences may be made and justification status inherited, so that sensings can perform their foundational function. To retain a role for the mental particulars that are sensed (sense data, sense contents conceived as a kind of sensed object), that theorist must be willing to say the following: "The primitive notion is believing that sense content x has property *F*. To sense the sense content x is to believe that it has some (no matter what) characteristic *F*. The sense content, which is a particular, is the intentional object of the epistemic sensing." The important thing to notice about this analysis is that epistemic notions are presupposed, not accounted for in terms of a supposedly antecedently understood nonepistemic notion of sensing a sense-content (thought of as a relation between a subject and a sense content, both of them particulars). In fact Sellars believes that no such reduction of the epistemic to the nonepistemic is possible, even in principle—though his arguments will not depend on this claim.

Section 6: At this point a further consideration is introduced: the ability to stand in the passive causal relations to the physical world envisaged by the fans of givenness is not something that must be *acquired* through experience or training. Organisms of the

right sort get it just by being awake. But the capacity to have beliefs of the form '*x* is *F*' involves classifying unrepeatables or particulars under repeatables or universals. It is natural to think that the capacity to classify is acquired, since one must learn by experience and training what the boundaries of the classes are. This line of thought results in the inconsistent triad of claims the sense datum theorist is committed to, and would like to be entitled to:

A. '*S* senses red sense content *x,'* entails '*S* noninferentially believes (knows) that *x* is red.'

B. The ability to sense sense contents is unacquired.

C. The capacity to have classificatory beliefs of the form '*x* is *F*' is acquired.

If A is given up, the sensing of sense contents becomes a nonepistemic event, which can at best be a logically necessary condition of knowledge or noninferential beliefs, not a logically sufficient condition of it. To take this way out would be to discard the line pursued in [4] & [5]. If B is given up, the sense datum theorist must either claim that we need practice to feel pain, hunger, itches, and so on when we are infants, or claim that feeling these things is not sensing. But then what is sensing? If C is given up, a story must be told about what universal concepts are innate (unacquired, inborn, wired-in) and which are not. This would require much more than even latter-day innatists such as Chomsky have claimed, since substantive concepts like *red* and *tall*, not merely grammatical forms, would have to be innate. A is the Myth of the Given, in one of its forms, and Sellars will give it up. He'll then owe (and provide) us a new account of both

thoughts and sensations, and of the origins (both in the order of causation and the order of justification) of knowledge.

Section 7: Sellars' diagnosis, which is not yet a treatment for the conceptual illness of givenness, is that it results from confusing two trains of thought, the first derived from an attempt to give a scientific account of perception and the acquisition of empirical information, and the second from an attempt to give a foundational epistemological account on the Cartesian model canvassed above in the discussion of [3]:

(1) The idea that there are certain inner episodes -- e.g. sensations of red or of C# which can occur to human beings (and brutes) without any prior process of learning or concept formation; and without which it would in some sense be impossible to see, for example, that the facing surface of a physical object is red and triangular, or hear that a certain physical sound is C#.

(2) The idea that there are certain inner episodes which are the noninferential knowings that certain items are, for example, red or C#; and that these episodes are the necessary conditions of empirical knowledge as providing evidence for all other empirical propositions.

The first class consists of particulars, picked out by their causal role. The second consists of claimings structured like sentences, picked out by their inferential or justificatory role. Sellars will offer an account (starting in [45]) of the genus, *inner episode,* to which these two species belong. He will call the first kind 'sense impressions', and the second kind 'thoughts', and will describe the roles they play. Finally, he will explain how they are related in human knowledge. (I have talked about belief so far, where Sellars talks about knowledge, in order to emphasize that the question of the *justification* of or warrant for noninferential beliefs has yet to be discussed.)

The result of running together these two lines of thought is "the idea that a sensation of a red triangle is the very paradigm of empirical knowledge." That idea is subject to precisely those related 'perplexities' Sellars has pointed out:

- Should we think of the sensation in question as a kind of *particular* (structured like a triangle), or as a kind of *belief* (structured like a sentence)?
- Is the capacity to have empirical knowledge like this acquired by experience, or prior to experience?
- Is it prior to the rest of our knowledge in the order of *causation*, or in the order of *justification* and evidence?

Part II [8]-[9] : Another Language?

Section 8: This and the next two sections (both marked in the original as Section 9!) are in one way an aside. The main thread is picked up again in [10]. The excursus is used to introduce some important ideas that will be discussed further along. The topic here is one possible form a sense datum theory might take to avoid the nonepistemically-specifiable-particular vs. only-epistemically-specifiable-sententially-structured-premise dilemma Sellars is constructing for it. One might give up entirely on the nonepistemic side of things, and embrace the foundational noninferential belief side. Thus Ayer sees sensing-of-sense-data talk as equivalent to and derivative from talk about how things look or seem to a subject. The suggestion comes in three parts:

a) There is a class of noninferential beliefs that form a justificatory basis for the rest of our empirical beliefs. [Note that this would be sufficient to respond to the regress argument sketched above in [[3]], though as suggested there, it is not a necessary condition for a response.]

b) Three nested descriptions of a phenomenon. First, a platitude: I may be mistaken that there is a red triangle in front of me. It is not possible for me to be mistaken about there *seeming* to be one. Next, a reifying move: an application of the Cartesian principle that although appearance must be distinguished from reality since subjects can be in error about the latter, on pain of an infinite regress it cannot be that one might be mistaken about the former also. Finally, a foundational claim: The class mentioned in (a) consists of beliefs that would be expressed by sentences used to make perceptual reports, prefixed by a special operator "It looks to me now that…," "It seems to me now that..." or "It now appears to me just as though…."

c) Sentences of the form "S is having (or is aware of) a sense datum that is *F*" (say, red and triangular) are by definition equivalent to sentence of the form "It seems to S that he senses something *F*". On this understanding, there are no particulars that are sense data -- the apparently referential singular terms that give the contrary impression must be understood contextually, like the 'it' in 'it is raining'.

Section 9: Here Sellars offers an observation about this approach, and then formulates a dilemma for it. The observation regards merely *generic* lookings. Something can *look* polygonal without there being any determinate number of sides that it looks to have. But nothing can *be* polygonal without there being a determinate number of sides that it has. (This contrasts will be explored in [17].) So the inferences one is

permitted to make in sense datum talk as introduced by the equivalence asserted by (c) are not the same as those licensed by the sense datum theorist's talk of sense data as particulars (for which the above 'inference to further determination' goes through). Thus the code is misleading.

Section 9 bis: The dilemma presents a more serious objection. If sense datum talk is just a code, it is redundant (insofar as it is not misleading). So what good is it? It can't *explain* anything about seemings or appearance. To do that it would have to be a *theory* of appearings, explaining them by relation to a certain kind of particular, namely sense data. (Sellars begins to explain how he thinks about theoretical explanation in [21] and [22]. We then hear a lot more about this topic in the second half of the essay, beginning at [39]-[44].) But this would reintroduce the strand of thought (1) above (in [7] and [[7]]), which the code theory is formulated precisely to avoid. The lesson is that that strand of thought is not altogether mistaken. The account Sellars will offer provides a theory of appearings, and will embrace and reconcile (1) and (2), properly understood. So (c) is not a way to avoid the problem. It allows us to look, however, at the assumptions (a) and (b) to which it was conjoined. Sellars' conclusion is that this line of thought is committed already at step (b) to the possibility of inferring from claims exclusively about how things *seem* to claims about how things actually *are*. But if, as (a) and (b) assert, all empirical evidence ultimately derives from how things seem, it is clear that such an inference cannot be warranted empirically, by inductive correlation of appearances and realities. The alternative seems to be to find a definitional reduction according to which "ordinary discourse about physical objects and perceivers could (in

principle) be constructed from sentences of the form 'There looks to be a physical object with a red and triangular facing surface over there'." Since commitment to (a) and (b) is much more widespread than commitment to (c), it is important to see what is wrong with the view they express—why the reduction they presuppose is impossible. To that end Sellars turns to the logic of 'looks' or 'seems' talk.

Part III [10]-[20] : The Logic of 'Looks'

Section 10: To get out of the trilemma of [6], it is necessary to "examine these two ideas [(1) and (2) in [7] and [[7]]] and determine how that which survives criticism in each is properly to be combined with the other." To begin with, consider the genus inner episode to which each subject has privileged access, which is common to sensations and thoughts.

a) Logical positivists have denied that there could be such episodes, because their existence is not intersubjectively verifiable or falsifiable. This is the source of the traditional problem of other minds, and of the inverted spectrum. To avoid entertaining such unverifiable hypotheses, one can reject idea (1).

b) Wittgenstein and some of his followers have attacked (2), the idea that inner episodes can be premises for inferentially based knowledge, because as private they escape the net of public discourse and language learning (the beetle in the box, and the private language argument).

Sellars disagrees with both of these. We'll return to the first later under the heading of 'behaviorism' ([54]-[55]). What one should reject in order to avoid the problem of other minds and the possibility of inverted spectra is the Myth of the Given, not the notion of inner episodes. Sellars will argue (in [45]-[47]) that (b) is both too strong and too weak. It is too strong in that inner episodes need not escape the net of public discourse. The second half of *EPM* has the task of showing us how to think of inner episodes—as theoretical entities that became observable. It is too weak in that this repudiation of inner episodes is not (as we see in the following sections) sufficient to avoid Sellars' foe, the Myth of the Given.

Section 11: The problem is that noninferential beliefs of the form "*X* looks *F* to *S*" can be held to be given in the bad sense even if, for inverted spectrum and beetle in a box reasons one refuses to talk about intrinsic properties of these lookings. That is, even if one does not assume (as one ought not, see [21]) that if anything *looks F* to *S*, something *is F* (the "sense datum inference"), one can still fall prey to the Myth. In order to attack this more insidious form of the Myth, then, Sellars considers the notion of 'looks' talk independently of any relation it might be taken to have to inner episodes as particulars.

Section 12: The question is: Does *looks-red* come before *is-red* conceptually (and so in the order of explanation)? That is, could the latter be defined in terms of the former in such a way that one could learn how to use the defining concept (*looking-F*) first, and only afterwards, by means of the definition, learn how to use the defined concept (*is-F*)? Descartes and his tradition claimed that *looks-F* talk, with which it is possible to form a class of statements about which subjects are incorrigible, is a foundation of knowledge,

and so must be prior in this sense to *is-F* talk, with which it is possible to express only corrigible, inferred beliefs. This view is the essence of Descartes' foundationalism.

Descartes was struck by the fact that the appearance/reality distinction seems not to apply to appearances. While I may be mistaken about whether something *is* red (or whether the tower, in the distance, *is* square), I cannot in the same way be mistaken about whether it *looks* red to me now.[8] While I may legitimately be challenged by a doubter: "Perhaps the item is not *really* red; perhaps it only *seems* red," there is no room for the further doubt, "Perhaps the item does not even *seem* red; perhaps it only *seems* to seem red." If it seems to seem red, then it really does seem red. The *looks*, *seems*, or *appears* operators collapse if we try to iterate them. A contrast between appearance and reality is marked by the distinction between *looks-F* and *F* for ordinary (reality-indicating) predicates '*F*'. But no corresponding contrast is marked by the distinction between *looks-to-look-F* and *looks-F*. Appearances are reified by Descartes as things that really are just however they appear. He inferred that we do not know them mediately, by means of representings that introduce the possibility of *mis*-representing (a distinction between how they really are and how they merely appear, i.e. are represented as being). Rather, we know them *immediately*—simply by having them. Thus appearings—thought of as a realm of entities *reported* on by noninferentially elicited claims about how things *look* (for the visual case), or more generally *seem*, or *appear*—show up as having the ideal qualifications for epistemologically secure foundations of knowledge: we cannot make

[8] I might be mistaken about whether *red* is what it looks, that is, whether the property expressed by the word 'red' is the one it looks to have. But that, the thought goes, is another matter. I cannot be mistaken that it looks that way, like *that*, where this latter phrase is understood as having a noncomparative use. It *looks-red*, a distinctive phenomenal property, which we may inconveniently only happen to be able to pick out by its association with a word for a real-world property.

mistakes about them. Just *having* an appearance ("being appeared-to *F*-ly", in one of the variations Sellars discusses) counts as *knowing* something: not that something is *F*, to be sure, but at least that something *looks-, seems-,* or *appears-F*. The possibility accordingly arises of reconstructing our knowledge by starting out only with knowledge of this sort—knowledge of how things look, seem, or appear—and building up in some way to our knowledge (if any) of how things really are (outside the realm of appearance).

This project requires that concepts of the form *looks-F* be intelligible in principle in advance of grasping the corresponding concepts *F* (or *is-F*). Sellars is a linguistic pragmatist about the conceptual order; that is, for him grasp of a concept just is mastery of the use of a word.[9] So he systematically pursues the methodology of translating questions of conceptual priority into questions about the relative autonomy of various language games. He will argue that in this case, Descartes got things backwards. 'Looks' talk does not form an autonomous stratum of the language—it is not a language-game one could play though one played no other. One must already be able to use 'is-*F* ' talk in order to master 'looks-*F'* talk, which turns out to be parasitic on it. In this precise practical sense, *is-F* is *conceptually* (Sellars often says 'logically') *prior* to *looks-F*.

Section 13: The definition being considered for exploitation in a Cartesian order of explanation (and so, ultimately, of justification) is:

x is red $=_{df.}$ *x* would *look* red under standard conditions.

[9] For it to be a *concept* one grasps thereby, the word must have an inferential role; must be usable in formulating premises and conclusions of inferences assessable as correct or incorrect. Thus acquiring the differential responsive dispositions required to use the word 'Ouch' does not qualify as grasping a concept. See [[16]] below.

Sellars will show how to acknowledge that this claim is definitionally true without countenancing the conceptual priority of 'looks' talk, and hence without giving aid and comfort to givenness and the sort of foundationalism it supports.

Section 14: Sentences can have reporting (noninferential) uses as well as (merely) fact stating (inferential) uses.[10] For reliable reporters, one may infer from the fact that one is disposed to say that *x* is *F* and that the conditions are as far as one knows standard and that when under those conditions one is so disposed it is usually the case that *x* is *F*, to the conclusion that *x* is in fact *F*. (The reporter's having to believe all of this, and so to understand it, is crucial to Sellars' later argument). Understanding the possibility of systematic error in the responsive dispositions of reporters introduces a new dimension in the relation between practices of *reporting* and those of *inferring*. Here Sellars introduces the illustrative parable of young John in the tie-shop.

Section 15: Where collateral beliefs indicate that systematic error is likely, the subject learns not to make the report '*x* is *F*', to which his previously inculcated responsive dispositions incline him, but to make a new kind of claim: '*x looks* (or seems) *F*'. Of course it is tempting to take this as a new kind of report, indeed a report of a special kind of particular, a sense datum. This report then is naturally thought of as reporting a minimal, noninferentially ascertainable, foundationally basic fact, about which each subject is incorrigible. There are two points here, which might be distinguished. First, it

[10] Sellars' terminology is strained here. There is no reason to deny that noninferential reports are in the fact-stating line of work, and so, when true, state facts. The preferred usage is to see the distinction between claims that are noninferentially elicited and those that arise as the conclusions of inferences as a distinction *within* fact-stating discourse.

is a mistake in any case to treat what is reported as foundationally basic facts, for the concepts needed to formulate them turn out to depend on other concepts, which are not formed using any analogue of a 'looks'-operator. (See [32].) It is a further claim that it is a mistake to treat these as reports at all—since they *evince* a disposition to call something *F,* but may not happily be thought of as *saying that* one has such a disposition. Sellars wavers on the second point, but he is firm on the first.

Section 16: Sellars' alternative analysis depends on distinguishing two different dimensions of the use of a noninferential report. First, each report is the manifestation of some *reliable differential responsive disposition*. That is, it is the result of one's being trained to behave in a certain way when in certain environmental situations (like a pigeon trained to peck at the red square when the red light comes on). What is the difference between a parrot trained to utter "That's red!" when and only when confronted by the visible presence of something red, and a genuine noninferential reporter of the same circumstance? Having the differential responsive dispositions is not enough to have the concept, else a chunk of iron that rusts in wet environments and not in dry ones would have to be counted as having the concepts of wet and dry environments. What more, besides the parrot's *sentience* is required for the *sapience* that consists in responding differentially by applying a *concept*? Sellars' answer, invoking the second dimension of reporting, is that the response must be taking up a position in the space of reasons—making a move in the game of giving and asking for reasons. The genuine noninferential reporter of red things has, and the parrot has not, mastered the *inferential role* played by reports of that type—where inferential role is a matter of what conclusions one is entitled

to draw from such a statement when it is overheard, what would count as a reason for it, and what is incompatible with it and so a reason against it. This is a matter of the inferentially articulated content of the assertional commitment undertaken by the reporter in virtue of the performance that is the reporting: what the reporter is responsible for. Sellars' term for this second dimension is *endorsement*, a matter of what one is linguistically *committed* to (the inferential consequences of one's claims) or responsible for (how it could be justified) in virtue of one's assertional performance. This notion of *responsibility*, or of what conclusions one has given others the *right* to draw, or has *obliged* oneself to draw, and what other commitments would count as *entitling* one to the commitment one has undertaken is the normative element in linguistic conduct, whose irreducibility to descriptive aspects (such as responsive dispositions) lies at the base of the epistemic/nonepistemic distinction, and is the source of Sellars' remark about the naturalistic fallacy at the end of [5].

On Sellars' understanding, the ability to use '*x* looks green' correctly appeals to the same responsive dispositions acquired in learning to use '*x* is green' correctly. But these two sorts of remarks elicited in accordance with those dispositions support quite different inferences. In particular, the parable of the tie shop shows that in saying that something merely *looks* green one can be understood to be doing two things: expressing one's noninferential differential responsive disposition to call it green (to commit oneself to the claim that it is green, with all of its inferential consequences and justificatory obligations), and at the same time explicitly *withholding* one's endorsement of that claim. For collateral beliefs concerning the possibility of systematic error under the prevailing circumstances of observation have undermined the reporter's confidence in his

reliability—that is, in the correctness of the inference from "X is disposed noninferentially to report the presence of something green (seen by electric lighting)," to "There is (probably) something green there."

This analysis of what one is doing in using 'looks' explains the incorrigibility of 'looks' talk. One can be wrong about whether something *is* green because the claim one endorses, the commitment one undertakes, may turn out to be incorrect. For instance, its inferential consequences may be incompatible with other facts one is or comes to be in a position to know independently. But in saying that something *looks* green, one is not endorsing a claim, but *withholding* endorsement from one. Such a reporter is merely evincing a disposition to do something that for other reasons (e.g. suspicion that the circumstances of observation lead to systematic error) he is unwilling to do—namely, endorse a claim. Such a reporter cannot be wrong, because he has held back from making a commitment. This is why the *looks*, *seems*, and *appears* operators do not iterate. Their function is to express the withholding of endorsement from the sentence that appears within the scope of the operator. There is no sensible contrast between 'looks-to-look *F*' and 'looks-*F*', of the sort there is between 'looks-*F*' and '(is-)*F*' because the first 'looks' has already withheld endorsement from the only content in the vicinity to which one might be committed (to something's being *F*). There is no further withholding work for the second 'looks' to do. There is nothing left to take back. Since asserting 'X looks *F*' is not undertaking a propositionally contentful commitment—but only expressing an overrideable disposition to do so—there is no issue as to whether or not that commitment (which one?) is correct.

Sellars accordingly explains the incorrigibility of appearance-claims, which had so impressed Descartes. He does so in terms of the practices of using words, which are what grasp of the relevant appearance concepts must amount to, according to his methodological linguistic pragmatism. But once we have seen the source and nature of this incorrigibility—in down-to-earth, practical, resolutely nonmetaphysical terms—we see also why it is precisely unsuited to use as an epistemological foundation for the rest of our (risky, corrigible) empirical knowledge. For, first, the incorrigibility of claims about how things merely *look* simply reflects their emptiness: the fact that they are not really claims at all. And second, the same story shows us that 'looks' talk is not an autonomous language game—one that could be played though one played no other. It is entirely parasitic on the practice of making risky empirical reports of how things actually are. Thus Descartes seized on a genuine phenomenon—the incorrigibility of claims about appearances, reflecting the non-iterability of operators like *looks, seems,* and *appears*—but misunderstood its nature, and so mistakenly thought it available to play an epistemologically foundational role for which it is in no way suited.

This analysis of "the logic of 'looks' talk", and the consequent diagnosis of the errors of a foundationalism based on the incorrigibility of our epistemic access to appearances is the constructive core of Sellars' critique of Cartesianism. It does not purport to be a knock-down argument; for it can only be as persuasive as its account of how 'looks' talk works, and alternatives are always possible.[11] What it is meant to do is

[11] "There are many interesting and subtle questions about the dialectics of 'looks'-talk, into which I do not have space to enter [17]." Sellars focuses on one sort of use that 'looks' has: first-person, noninferential uses. But as he points out in this section, 'Looks' also has third person, (merely) fact-stating uses, as when

to remove the temptation to go the Cartesian foundationalist route, by explaining, without explaining away, the exact nature of that temptation.

Section 17: (This line of thought is completed in [22]). We begin to look for some confirmation of the two-pronged account of 'looks' talk as expressing a differential responsive disposition to make a specified noninferential report, while withholding endorsement of that claim. The confirmation takes the form of explanations of otherwise puzzling features of appearance-talk. Consider the three sentences:

i) The apple over there is red.

ii) The apple over there looks red.

iii) It looks as though there were a red apple over there.

Utterances of these sentences can express the same responsive disposition to report the presence of a red apple, but they endorse (take responsibility for the inferential consequences of) different parts of that claim. (i) endorses both the existence of the apple, and its quality of redness. (ii) endorses only the existence of the apple. The 'looks'

S' says :"X looks *F* to S." The account generalizes to these cases straightforwardly. In making this claim, S' is doing two things: attributing to S a disposition noninferentially to report that X is *F*, and herself withholding endorsement from the claim that X is *F*. The words used to make this report do not settle whether S' would attribute to S the endorsement, or the withholding of endorsement, from the claim that X is *F* (that is, whether S' takes it that the disposition being she attributes to S issues in an endorsement, or is overridden and withheld by S). More expressively powerful and discriminating regimented uses that do mark such distinctions are easily contrived. (This same idea will account for the nontrivial iterated or embedded uses of 'looks' or 'seems' made possible by shifts of perspective: "It seems to S' that X looks *F* to S," and so on.)

Further uses of 'looks' that are more challenging to the Sellarsian account have been pointed out by Joe Camp. These are cases where we use 'looks *F*' without a corresponding practice of using the unmodified '(is) *F*'. Thus after the optometrist puts pupil-dilating drops in my eyes, I may say "Things look blurry." 'Blurry' does not express a way things could *be*. It is essentially an expression pertaining to images or representations. Taking this sort of use of 'looks' as central and paradigmatic could encourage the reintroduction of the idea that noninferential uses of 'looks' are genuine reports, reports of intrinsic features of appearances as such. Sellarsians will presumably see these 'intrinsic' uses of 'looks' rather as sophisticated late-coming possibilities, derivative from the central uses, and to be explained in terms of them. Sellars himself never discusses this issue.

In a note to the 1963 edition, Sellars suggests that one might distinguish between 'looks F' and 'looks to be F', in a way that corresponds to Chisholm's distinction between non-comparative and comparative 'appears' statements.

locution explicitly cancels the qualitative commitment or endorsement. (iii) explicitly cancels both the existential and the qualitative endorsements. Thus, if someone claims that there is in fact no apple over there, he is asserting something incompatible with (i) and (ii), but not with (iii). If he denies that there is anything red over there, he asserts something incompatible with (i), but not with (ii) or (iii). Sellars' account of the practice of using 'looks', in terms of the withholding of endorsement when one suspects systematic error in one's responsive dispositions, can account for the difference in scope of endorsement that (i)-(iii) exhibit. But how could that difference be accounted for on a sense datum approach?

In this section Sellars points out another virtue of his account, as opposed to sense datum theories, namely the possibility of reporting a merely *generic* (more accurately, merely determinable) look. (Recall that Sellars introduced this phenomenon already in [9].) Thus it is possible for an apple to look red, without its looking any specific shade of red (crimson, scarlet, etc.). It is possible for a plane figure to look many-sided without there being some particular number of sides (say 119) which it looks to have. But if 'looks' statements are to be understood as reports of the presence before the eye of the mind of a particular which *is F*, how can this possibility be understood? Particulars are completely determinate. A horse has a particular number of hairs, though as Sellars points out, it can *look* to have merely 'a lot' of them. It is a particular shade of brown (or several shades), even though it may look only darkly colored. So how are such generic, merely determinable, looks possible? Sellars' account is in terms of scope of endorsement. One says that the plane figure looks 'many-sided' instead of '119-sided' just

in case one is willing only to endorse (be held responsible for justifying) the more general claim. This is a matter of how far one is willing to trust one's responsive dispositions, a matter of the epistemic credence one feels they deserve or are able to sustain. Particulars, even if they are sense contents, cannot be colored without being some determinate color and shade. How then can the sense datum theorist—who wants to say that when something *looks F* to S, something in S *is F*—account for the fact that something can look colored without looking to be any particular color, or look red without looking to be any particular shade of red? So Sellars' account of 'looks' talk in terms of endorsement can account for two aspects of that kind of discourse that no theory that invokes a given can explain: the scope distinctions between qualitative and existential lookings, and the possibility of merely generic or determinable lookings.

Section 18: On this account, then, one must first acquire the practice of reporting red objects (getting both the appropriate responsive dispositions and an understanding of what one is endorsing by making such a claim), and only then can one learn to make reports expressing those same dispositions, but which are more guarded in their endorsement. As the argument of the previous section has shown, the mastery of different endorsements required can be quite sophisticated. One may, for instance, discriminate existential and qualitative lookings, and various grades of determinability. Thus to know when something *looks* red one must understand what it is to *be* red, and a good deal besides.[12] We can see at this point that the sentence we worried about in [13]

[12] In a footnote added to the 1963 reprinting of the essay, Sellars points out that his story is compatible with distinguishing a *rudimentary* concept corresponding to the use of 'green', which one can have *without* having mastered the use of 'looks green', and a *richer* concept (corresponding more closely to ours) which is achieved only once one has also mastered 'looks' talk. This observation opens the door to distinguishing

is true because it is a definition, not of *is-red*, but of *standard conditions*. For standard conditions are just those in which one's responsive dispositions can be trusted, and ought to be fully endorsed. Given such a definition, one can investigate empirically what those conditions are.

Sections 19 & 20: These sections return to the question of the acquisition of various capacities involved in mastering an observation concept, the question that set up the trilemma of [6]. We now know that these capacities involve both regular responsive dispositions and a capacity to manipulate endorsements inferentially, a nonepistemic and an epistemic skill respectively. The specifically *inferential* articulation required for endorsements to qualify as *conceptually* contentful introduces at least a limited *holism* into Sellars' picture: one could not have one concept unless one had many others to which it is inferentially related: "The essential point is that even to have the more rudimentary concept [of *green*, say] presupposes having a battery of other concepts." This entails rejecting the idea that "fundamental concepts pertaining to observable fact have that logical independence of one another which is characteristic of the empiricist tradition." These sections do not present Sellars' argument in a perspicuous, or even linear fashion, and the argument is repeated in more satisfactory form at [33]-[37], where we will discuss it.

Part IV [21]-[23] : Explaining Looks

(as McDowell does) concepts corresponding to *secondary qualities* as those mastery of which requires mastery of the associated 'looks' vocabulary—as, arguably, green does, while massive, and perhaps square do not.

Sections 21 & 22: [17] discussed the issue of the scope of endorsement, which is treated in the third paragraph of [21]. Sellars reformulates that account in terms of the events that are lookings and seeings, where earlier we discussed the reports causally occasioned by those events. The first two paragraphs of [22], and the third paragraph of [21] then present a preliminary account of a distinction between two ways in which some phenomenon can be explained:

i) By deducing it from some empirical generalization formulated entirely in terms of observables (things that can be noninferentially reported). Explaining a change in pressure of a gas sample by appealing to the law PV=kT, together with suitable background conditions, is an example.

ii) By postulating unobservable entities, and subsuming the phenomenon under laws involving those theoretical entities. Explaining the change in pressure of a gas by appealing to the kinetic theory of gases and its postulated molecules and their interactions is an example.

Once again, this discussion is really out of place here, serving merely as a dark foreshadowing of a line of argument that will be pursued in more detail later (beginning at [39]-[44]). Sellars does here raise the important question that remains even after we have understood existential, qualitative and unqualified 'looks' statements in terms of scope of endorsement (as in [17]), namely, what is it that is common to the three cases? The answer, we will see, is: "sense impressions of red." But that is an answer we will not be able to understand until the very end of the essay ([62]).

Section 23: This section discusses the question of what things can literally be red. Sellars' claim is that only physical objects can, and that it is a mistake to think that even facing surfaces can, except in a derivative sense. The discussion as it stands is unsatisfactory, since criteria of primacy of sense and literalness of attribution are not forthcoming. Once again, a later discussion is being prefigured, in this case concerning the 'of-red'ness of sensations in [60]-[61].

Parts V and VI (Sections [24]-[29]) offer a discussion of the British Empiricists' treatment of impressions. They fill in some of the historical background of Sellars' discussion, but are not central or essential to the development of his argument.

Part V [24]-[25] : Impressions and Ideas: A Logical Point

Sections 24 & 25: A discussion of the intentionality of 'sensation of...'. Sellars' view is that Descartes mistakenly assimilated sensations and thoughts because of the way in which each is 'of' or 'about' or directed at something. There need be no red triangle for me to have a sensation 'of' one, and there need be no golden mountain for me to have a thought 'of' one. But this is a superficial similarity, for the kind of aboutness is in fact quite different in the two cases. An equally important motivation for the assimilation, which Sellars does not mention here, is the incorrigibility and transparency, the epistemic privilege accorded to reports of sensations and of thoughts. Here Sellars point out "the notorious 'ing'/'ed' ambiguity" as it applies to the concept of *experience*.

Part VI [26]-[29] : Impressions and Ideas: A Historical Point

Section 26: The inverted spectrum problem cannot be stated without recourse to some version of the Myth of the Given.

Section 27: A central epistemological problem of the empiricists Locke, Berkeley, and Hume is taken to be the question how, given that we can be aware of completely *determinate* sense repeatables, we can come to be aware also of *determinable* sense repeatables. Determinate/determinable is like species/genus, except that there is no separately specifiable differentiating factor. Colors are the prime example: scarlet is a more determinate shade of the determinable color red, as red is a determination of the determinable colored.

Section 28: The British Empiricists "...all take for granted that the human mind has an innate ability to be aware of certain determinate sorts—indeed, that we are aware of them simply by virtue of having sensations and images." That is, they did not ask a corresponding question about how, given that we can be aware of particular unrepeatable *token* sense contents, we can come to be aware also of their repeatable *types*, even maximally determinate ones.

Section 29: Against this Sellars will argue for what he calls ‘psychological nominalism’ (not the best imaginable name) according to which all awareness of repeatables (whether determinate or determinable) is a linguistic affair, and hence may

not be presupposed in one's account of the acquisition and functioning of language. Sellars is proposing a linguistic, social theory of awareness. He has in mind more by this term than simply being awake (not being asleep): he is after awareness in the sense of *sapience*, not of *sentience*. It is classificatory awareness, awareness of something as something. But not all acts of classification are acts of awareness. As pointed out above in [[16]], anything with stable dispositions to respond differentially to stimuli can be thought of as classifying the stimuli according to the repeatable responses those stimuli elicit. A parrot trained to respond differentially to red things in its environment does not display the sort of awareness that Sellars is explaining. Such awareness, specifically *conceptual* awareness requires something beyond being awake and classifying by differential response.

Part VII [30]-[31] : The Logic of 'Means'

Section 30: Anyone whose account of the prelinguistic awareness that makes language acquisition possible assigns it an inferential structure (Sellars says 'logical', but that is just 1950s talk for 'conceptual', which for Sellars can be parsed as 'inferential') is committed to the Myth of the Given. Such conceptual awareness involves not only classification, but making the classifications significant in inferences. It is at this point that statements of fact are made, particulars referred to and classified under universals. On the sort of account Sellars opposes to the Myth, conceptual content is inferentially articulated. But inference is a process arising only within the "game of giving and asking for reasons", which essentially involves beliefs. This is a normative realm, of commitment and

entitlement to claims, of endorsement and justification. It is what Sellars has been calling the 'epistemic'. The Myth is to think that anything could intrinsically, naturally, or necessarily possess a particular significance for this realm, independently of the acquisition or deployment of concepts by the one for whom it has that significance. Acts of awareness as traditionally conceived, as entailing the existence of something sufficiently belief-like to serve as the ultimate inferential ground of empirical knowledge (never mind as themselves constituting knowledge), would have to have just such a property.

Section 31: So learning a meaning ought not to be understood as associating something one is already aware of with a verbal symbol. But isn't this the natural way to understand statements like "'Rot' (in German) means red"? How else can Sellars understand this sentence except as expressing an association between one's awareness of the determinable repeatable quality and the word 'rot'? His answer is that meaning claims like this really assert that the mentioned expression ('rot') plays the same conceptual functional role as the used expression ('red'). "These considerations make it clear that nothing whatever can be inferred…about the exact way in which the word 'red' is related to red things, from the truth of the semantical statement ' "red" means the quality *red*.'"

Part VIII [32]-[38] : Does Empirical Knowledge Have a Foundation?

Section 32: Another incorrect, foundationalist account is described here. Sellars only disagrees with one bit of this story, though it turns out to be an important bit. Foundationalism is the claim that there is a structure of particular beliefs such that:

1) Each one is noninferentially arrived at.

2) The beliefs in (1) presuppose no other belief, either particular or general.

3) These noninferentially acquired beliefs constitute the ultimate court of appeal for all factual claims.

Sellars accepts (1) and (3), but denies (2). His project at this point is to show how a bit of knowledge (belief) can, and indeed how all of it does, presuppose other knowledge (belief), even though it is not inferred from that other knowledge or belief. This possibility was not seriously examined by the classical epistemological tradition. It is a certain hierarchical picture of *understanding* (at this level a necessary condition of believing) that Sellars rejects. He does not object to a hierarchical picture of *justification*, once that has been suitably disentangled from bad foundationalism concerning the nature and acquisition of belief.

For Sellars, there is no such thing as a noninferential belief, if by that one means a belief one could have without grasping its inferential connection to at least some other beliefs. For to understand a sentence, to grasp a propositional content (a necessary condition of having a belief) is to place it in the space of reasons, to assign it an inferential role in the game of giving and asking for reasons, as entailing some other contents and being incompatible with others. A noninferential report or belief can properly be called ‘noninferential’ only in the sense that the reporter’s commitment to an

essentially inferentially articulated content is elicited noninferentially on this occasion—that is, that it is elicited as a response to some nonlinguistic, nonepistemic environing circumstance, rather than as a response to another belief or assertion. Noninferential beliefs do not form an autonomous discursive stratum: there could be no language game consisting entirely of noninferential reports. (Notice that this is a stronger claim than that made above in connection with 'looks'-talk. For this claim concerns *any* kind of noninferential report, whether what they report is inner or outer, appearances or empirical realities.) For any sentence to have noninferential uses, some sentences must have inferential ones. For the conceptual content expressed by a sentence (what is believ*ed*) essentially involves its potential as a premise and as a conclusion of inferences. Unless one can employ noninferentially acquired beliefs as the premises of inferences leading to further beliefs, their acquisition does not qualify as acquiring *beliefs* (something propositionally contentful) at all. On this inferentialist picture of conceptual content, one cannot have one concept without having many inferentially interrelated ones. This does not mean that there could not be a language consisting only of expressions for observables, however. For the concepts of observables are concepts that have noninferential, reporting, uses. The requirement is only that the concepts that can be used to make noninferential reports must also be available to be applied inferentially, as the conclusions of inferences whose premises are the noninferential applications of other concepts.

Sellars begins by asking about the nature of the *authority* (a patently normative notion) of noninferential beliefs, that is, their capacity to justify other claims. A

distinction is needed first between sentence *types* and sentence *tokens*: the type is repeatable and can be instantiated on different occasions, while the token is unrepeatable. It is the utterance or inscription of the sentence on a particular occasion. So if the distinction is applied to letters instead of sentences, the sequence 'aeaaeea' contains two letter types and seven letter tokens, four of one type and three of the other. Now it can be seen that it is sentence tokens whose justification is at issue. For while there are some sentences that are justified, if they are justified at all, whenever they are tokened, such as '2+2=4,' and 'Red is a color,' there are others that can be justified (and true) on one occasion and not justified or true on another. Then only the tokens and not the types can be said to be justified. These are sentences like 'That car is red,' or 'I'm hungry now,' which contain words whose reference is determined by the actual circumstances in which the sentence is tokened. These are called "token reflexive" expressions. Many, though not all, of the noninferential beliefs putatively described by (1)-(3) above are token reflexive. Authority or credibility (positive justification status) is either extrinsic, coming from something else, in this case by inferential inheritance, or intrinsic. Intrinsic credibility may be associated with types, as in meaning-analytic statements such as 'All bachelors are unmarried males,' or with tokens, as in 'This is red,' (or, given Sellars' account of 'looks', 'This looks red').

Section 33: Sellars now considers a line of thought according to which intrinsically credible types and intrinsically credible tokens, analytic claims[13] and observation reports,

[13] In "Two Dogmas of Empiricism" Quine objects to the notion of meaning-analytic claims (claims true in virtue solely of the meanings of their words) on the broadly pragmatist grounds that there is no practically discernible status corresponding to this supposed category. Claims taken to be analytic, such as "All bachelors are unmarried males," are not immune from revision, known a priori, or otherwise distinguished from statements of very general fact, such as "There have been black dogs." Sellars accepts analyticity, which he associates with the practical status of counterfactual robustness. This line of thought ties our

are similar in that they are both types such that their being *correctly* tokened, that is tokened according to the *rules* for the use of all the component expressions, is a sufficient, not just a necessary condition of their being true and justified (hence not just believed but known). Sellars can swallow all of this except the bit about rules. The idea he will reject is that analytic statements are true by virtue of discursive definition (definition of a linguistic expression in terms of other linguistic expressions), while observation reports are true by virtue of ostensive definitions. Ostensive definitions are the only sort we can give of terms like 'red'. They consist of defining the expression by exhibiting samples of the things it applies to (pointing to red objects). The usual foundationalist infinite regress argument can be applied to show that not all expressions of the language can be discursively defined on pain of circularity or infinite regress (in either case no definition is achieved). So there must be ostensive definitions in the language. These definitions, just like the discursive ones, codify the rules of appropriate usage of the expressions they define. Just as following those rules is sufficient for the truth of analytic statements, so following the 'rules' of ostensive definition is to be sufficient for the truth of observation reports. (Such a rule might look like the definition in [13].) At this point Sellars disagrees. One can imagine following the rules for the use of 'this', 'is', and 'green' only if one has some idea of prelinguistic awareness of green -- the Myth. For Sellars it is incoherent to talk of ostensive definitions setting up rules for

---

concepts to what we take to be laws of nature. (See Sellars' "Concepts as Involving Laws, and Inconceivable Without Them," in *PPPW*, and "Counterfactuals, Dispositions, and Causal Modalities" [ref.].) So conceived, analytic claims are neither immune from revision, nor known a priori.

using 'green', for there is no language available in which such rules could be stated. Ostensive definitions establish practices; they are regular, but not rule governed.[14]

Section 34: The notion that the authority of noninferential reports rests on episodes of nonverbal hence *nonconceptual* awareness, which verbal performances express, is a version of the Myth. From Sellars' point of view, such episodes are the tortoise underneath the elephant.

Section 35: Here Sellars presents his alternative view. It begins with the observation that " ... a token of 'This is green' in the presence of a green item...expresses observational knowledge [only if] it is a manifestation of a tendency to produce tokens of 'This is green'-- given a certain 'set' (context of collateral commitments and circumstances) if and only if a green object is being looked at in standard conditions...". That is, it must be the expression of a reliable differential responsive disposition. But photocells and parrots could satisfy this condition, which shows that so far only the responsive dispositions part of the observation report has been specified. It remains to add conditions to capture the epistemic side, the dimension of endorsement, of undertaking inferentially articulated commitments, of producing a performance with a distinctive kind of authority.

To have the authority of knowledge, the report must not only *be* reliable, it must be *taken to be* reliable. In fact Sellars claims that it must be *known* by the reporter to be reliable (and in this he perhaps goes too far): "...the perceiver must know that tokens of

[14] This distinction, and the need for some sense in which a practice (paradigmatically, a linguistic practice) can be governed by norms even though its practitioners cannot be said to be following rules, is Sellars' topic in his important essay "Some Reflections on Language Games".

'This is green' are symptoms of the presence of green objects in conditions which are standard for visual perception." ‘Justification’ has the 'ing/ed' ambiguity (cf. [24]): justifying, a practical activity, or being justified, a normative status. Sellars claims that one cannot have the status except when it is possible to redeem that claim to authority and epistemic privilege by engaging in the activity of justifying it. This claim of the priority of practice over status is a specific variety of pragmatism, to which Sellars adheres. The difference between a noninferential reporter and a photocell, or a parrot trained to utter ‘It's getting warmer,’ as the temperature rises, does not lie in the reliability or range of their responsive dispositions. It lies in the capacity of the reporter to redeem the commitment undertaken, the authority claimed by the reporting, by justifying the claim (if challenged) by giving reasons for it. The by now familiar basic point is that in order to count as making a claim (expressing a belief) at all, the reporter must be "in the space of giving and asking for reasons", in addition to having the right responsive dispositions. The further claim being forwarded here is that for a noninferential report to express knowledge (or the belief it expresses to constitute knowledge), the reporter must be able to justify it, by exhibiting reasons for it. This is to say that the reporter must be able to exhibit it as the conclusion of an inference, even though that is not how the commitment originally came about.

The inference in question is what might be called a “reliability inference.” One justifies a noninferentially elicited report that something is red by noting that one was disposed noninferentially to apply the concept *red* to it, and pointing out that one is a reliable reporter of red things in these circumstances. To say that one is reliable is just to say that the inference from one’s being disposed to call something red to its actually

being red is a good one. Thus the reliability of one's differential responsive dispositions, together with the report's being an exercise of those dispositions together justifies—offers good reasons for—the report. In insisting that in order properly to be credited with knowledge a reporter must be able to offer an inferential justification of the belief in question, Sellars is endorsing an epistemological internalism that puts him at odds with more recent reliabilist externalists in epistemology. Their claim is that the real function of the traditional justification condition on knowledge is to rule out *accidentally* true beliefs. If so, then the rationale for engaging in assessments of whether various beliefs qualify as knowledge is perfectly well-served by insisting only that candidate beliefs result from reliable belief-forming mechanisms—that is, mechanisms that are likely to lead to truths, whether or not the reporter knows that they are. Forming beliefs one can justify then appears as one reliable mechanism among others.

Of course, from Sellars' point of view it would be a mistake to conclude from this line of thought that one could trade inferential justification for reliable belief-formation in a wholesale fashion. For that it is *beliefs* one is forming, that what one is doing is applying *concepts*, is a matter of their specifically *inferential* articulation—their role in the game of giving and asking for reasons, justifying and demanding justifications. Against that background of inferential practice, however, it is not obvious why Sellars should resist the reliabilist's suggestion. Why isn't it enough that the *attributor* of knowledge know that the reporter is reliable, that the *attributor* of knowledge endorse the inference from the reporter's responsive disposition noninferentially to apply the concept *red* to the thing's (probably) being red? Why should the reporter herself have to be able

to offer the inferential justification for her noninferential report? (This is the thought behind the qualification offered parenthetically early in the second paragraph above.)

Section 36: The moral is that on the true view "one could not have observational knowledge of any fact unless one knew many other things as well". This is not to say that observation reports are somehow the results of inferences after all, but only that, though noninferential, they must be justifiable to be justified. The false view thinks it is supposed to give a causal description of how knowledge is possessed, but: "...in characterizing an episode or 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, or justifying and being able to justify what one says." Thus everything irrelevant to justification, either to knowing what would be a justification or to being entitled to produce one, is a noncognitive causal antecedent, perhaps a necessary condition of empirical knowledge, but not one that is constitutive of it. Nor is the general point is not specific to the normative, epistemic status of knowledge—though Sellars does not point this out. He could as well have said that in characterizing an episode or state as one of *believing*, or *applying concepts*, or *grasping propositional contents* we are not giving an empirical description of that episode or state but placing it in the logical space of reasons, or justifying and being able to justify what one says. For only what is inferentially articulated is conceptually contentful (and hence qualifies as a believable or claimable) at all. As we saw in the previous section, however, Sellars does wants to insist further that one cannot *know* noninferentially that something is green unless one also *knows* that one is a reliable reporter of green things under the prevailing circumstances.

Section 37: This view—the one Sellars endorses—seems to involve an infinite regress. For how could we have acquired knowledge that tokens of 'This is green,' are reliable symptoms... unless we had had knowledge of such facts as 'This is green,' and "This is a token of 'this is green'" beforehand?[15] Sellars' answer is that we can acquire knowledge of facts of these three types simultaneously, but that we can know facts that bear on events that occurred before we acquired any of this knowledge. Thus: "...it requires only that it is correct to say that Jones now knows, thus remembers, that these particular facts did obtain. It does not require that it be correct to say that at the time these facts did obtain he then knew them to obtain. And the regress disappears." Thus children at the age of six can know that at four they saw—in the sense of reliably responded to—a fire, although at the age of four all they could do was say 'fire' parrot-fashion, without knowing there was a fire.[16] The important difference is not one of responsive disposition but of capacity to endorse. The six year old has moved into the space of giving and asking for reasons; he can commit himself to a claim and be treated as authoritative; he is responsible for the claim he undertakes. For this he must at least be able to tell what he is thereby committing himself to and what evidence would entitle him to it, that is, he must understand his claim. But even that is not sufficient. For this new normative status is socially conferred. No nonepistemic description of the candidate reporter suffices for the conferral of this status, unless and only insofar as the community conferring that status, treating the individual as responsible, reliable, and so on, takes it to be sufficient.

---

[15] Notice that this is a problem Sellars need not have faced, had he endorsed the modified externalism offered to him in [[35]] above.

[16] Commenting on this point for the 1963 edition, Sellars said that his thought was that one could have direct (in the sense of noninferential) knowledge of a past fact which one could not conceptualize at the time that it occurred.

Compare achieving one's majority and being able for the first time to undertake contractual obligations. This status consists in the community's recognition of it. Some minors are more reliable at carrying out the commitments they undertake than many over the age of 21, but this fact does not make their signature mean that they have entered into a contract. This is how “the light dawns slowly over the whole”: at some point one masters the moves, inferential and noninferential, sufficiently that one’s noises come to be taken by one’s community as having the significance of making claims, undertaking commitments, giving reasons.

Section 38: The only sense in which there is no foundation for empirical knowledge is the sense in which the observation reports, which in a certain sense are its foundation, themselves rest (not inferentially, but in the order of *understanding* and sometimes of justification) on other sorts of knowledge. Observation reports, whether of inner episodes or outer happenings, do not constitute an autonomous stratum of the language—a game one could master though one had as yet not mastered the inferential use of any expressions. That is, Sellars rejects *only* claim (2) of the three foundationalist theses considered in [32]. But there is no need for a foundation in this sense: "Empirical knowledge is rational, not because it has a foundation but because it is a self-correcting enterprise which can put any claim in jeopardy, though not all at once."

Part IX [39]-[44] : Science and Ordinary Usage

Sections 39-44: Here Sellars sketches his Scientific Realism. He includes this discussion because if science is viewed in the opposite, positivist fashion, the notion of

inner episodes as theoretical entities, which he is about to introduce, is incoherent. Thus [42] claims that "science is the measure of all things...". This is a view about the authority of claims couched in scientific vocabulary relative to the authority of claims couched in other vocabularies. [43] briefly indicates the positivist view. According to the positivist scheme, there is an observation language, in which data are formulated and the results of experiments expressed. All we directly know about are the objects of observation (observation reports are the Konstatierungen of [33]). According to this account, a theoretical language is introduced in order to systematize our observations, and facilitate prediction and control. But the objects the theory postulates are virtual, mere calculational devices or instruments for the expression and systematization of observations. Theories are instruments, and their assertions should not be taken as entailing the existence of the objects they postulate. Sellars points out that only someone who thought that the observations themselves were given, not the product of the learning of concepts with which to report, would be tempted by this picture. Once it is discarded, another way of thinking about the distinction between theoretical and observable objects and concepts comes into view.

According to Sellars' view, the distinction between purely theoretical objects and observable objects is *methodological*, rather than *ontological*. That is, theoretical and observable objects are not different kinds of thing. They differ only in how we come to know about them. Theoretical objects are ones of which we can only have *inferential* knowledge, while observable objects can also be known noninferentially; theoretical concepts are ones we can only be entitled to apply as the conclusions of inferences, while

concepts of observables also have noninferential uses. But the line between things to which we have only inferential cognitive access and things to which we also have noninferential cognitive access can shift with time, for instance as new instruments are developed. Thus when first postulated to explain perturbations in the orbit of Neptune, Pluto was a purely theoretical object; the only claims we could make about it were the conclusions of inferences. But the development of more powerful telescopes eventually made it accessible to observation, and so a subject of noninferential reports. Pluto did not undergo an ontological change; all that changed was its relation to us. (Notice that this realism about theoretical entities does not entail scientific realism in the sense that privileges science over other sorts of cognitive activity, although Sellars usually discusses the two sorts of claims together.)

It might be objected to this view that when the issue of the ontological status of theoretical entities is raised, they are not considered merely as objects in principle like any others save that they happen at the moment to be beyond our powers of observation. They are thought of as *unobservable* in a much stronger sense: permanently and in principle inaccessible to observation. But Sellars denies that anything is unobservable in this sense. To be observable is just to be noninferentially reportable. Noninferential reportability requires only that there are circumstances in which reporters can apply the concepts in question (the dimension of inferentially articulated endorsement) by exercising reliable differential dispositions to respond to the objects in question (the causal dimension), and know that they are doing so. In this sense, physicists with the right training can *noninferentially* report the presence of mu mesons in bubble chambers.

In this sense of 'observation', nothing real is in principle beyond the reach of observation. (Indeed, in Sellars' sense, one who mastered reliable differential responsive dispositions noninferentially to apply normative vocabulary would be directly observing normative facts. It is in this sense that we might be said to be able to *hear*, not just the noises someone else makes, but their *words*, and indeed, *what they are saying*—their *meanings*.)

Once one sees that observation is not based on some primitive sort of preconceptual awareness (the tortoise beneath the elephant), the fact that some observation reports are riskier than others and that when challenged we sometimes retreat to safer ones from which the originals can be inferred will not tempt one to think that the original reports were in fact the products of inference from those basic or minimal observations. The physicist, if challenged to back up his report of a mu-meson may indeed justify his claim by citing the distinctively hooked vapor trail in the bubble chamber. This is something else observable, from which the presence of the mu meson can, in the right circumstances, be inferred. But to say that is not to say that the original report was the product of an inference after all. It was the exercise of a reliable differential responsive disposition keyed to a whole chain of reliably covarying events, which includes mu mesons, hooked vapor trails, and retinal images. What makes it a report of mu mesons, and not of hooked vapor trails or retinal images is the inferential role of the concept the physicist noninferentially applies. (It is a consequence of something's being a mu meson, for instance, that it is *much* smaller than a finger, which does *not* follow from something's being a hooked vapor trail.) If *mu meson* is the concept the physicist applies noninferentially, then if he is sufficiently reliable, when

correct, that is what he *sees*. His retreat, when a question is raised, to a report of a hooked vapor trail, whose presence provides good inferential reason for the original, noninferentially elicited claim, is a retreat to a report that is safer in the sense that he is a *more* reliable reporter of hooked vapor trails than of mu mesons, and that it takes less training to be able reliably to report vapor trails of a certain shape, so that is a skill shared more widely. But the fact that an inferential justification can be offered, and that the demand for one may be in order, no more undermines the status of the original report as noninferential (as genuinely an observation) than does the corresponding fact that I may under various circumstances be obliged to back up my report of something as red by invoking my reliability as a reporter of red things in these circumstances—from which, together with my disposition to call it red, the claim originally endorsed noninferentially may be inferred.

Part X [45] : Private Impressions: The Problem

Section 45: Sellars starts by setting the problem that will occupy him for the rest of the essay: "The problem of how the similarity among the experiences of *seeing that an object over there is red, its looking to one that an object over there is red* (when in point of fact it is *not* red) and *its looking to one as though there were a red object over there* (when in fact there is *nothing* over there at all). Part of this similarity, we saw, consists in the fact that they all involve the…proposition…that the object over there is red. But over and above this there is, of course, the aspect which many philosophers

have attempted to clarify by the notion of *impressions* or *immediate experience*." Sellars' response to this problem will not be fully in place until [62].

Next Sellars summarizes [32]-[38] (the meat of his epistemological discussion): we now recognize that instead of coming to have a concept of something because we have noticed that sort of thing, to have the ability to notice requires already having the concept, and cannot account for it. For to notice something—to be aware of it in the sense relevant to assessments of sapience, rather than of mere sentience—is to respond to it by applying a concept, making a noninferential judgment about it. So until one has the concept 'green' one cannot notice or be aware of green things, though one can respond differentially to them—obviously, in other ways than by applying the concept *green*. The title of this essay is "*Empiricism* and the Philosophy of Mind," but Sellars never comes right out and tells us what his attitude towards empiricism is. One might think he endorses it, misled by remarks such as he offers in [6] in discussing the inconsistent triad of commitments characteristic of classical sense-datum theories. For there he dismisses the option of rejecting the third element of the trilemma by doing no more than observing that to abandon it would "do violence to the predominantly nominalistic proclivities of the empiricist tradition" (proclivities that he discusses in more detail in [24]-[28]). But to interpret this remark as an endorsement *by Sellars* of the nominalistic proclivities of empiricism that he invokes here would be to mistake the role the remark plays in his argument. It is often hard to tell when Sellars is speaking in his own voice, and this is one of the occasions on which he is not. It is the classical sense-datum theorists who are committed to this tenet of empiricism, not Sellars—although, as will emerge just below,

he does as a matter of fact share with empiricists the belief that "The capacity to have classificatory beliefs of the form '*x* is *F*' is acquired."

Indeed, we can see at this point that one of the major tasks of the whole essay is to dismantle empiricism. For traditional empiricism depends on episodes of nonverbal, nonconceptual awareness, which serve both as the raw material for a process of abstraction by which concepts can be formed and grasped, and as our warrant for the ground-level (noninferential) applications of those concepts. (Compare [34].) This whole picture depends essentially on the Myth of the Given. Sellars' own view is one he is elsewhere happy to call 'rationalist'[17]: conscious experience presupposes that the experiencer already has concepts, and so cannot account for their acquisition. In this claim, Sellars aligns himself with the Leibniz of the *New Essays*, writing against his Lockean target. Sellars' task in the rest of the essay is to show how the philosophy of mind can understand inner episodes once one has rejected *both* Cartesianism *and* empiricism, having recognized that both depend upon the Myth of the Given.

The classical pre-Kantian rationalists, having won their way through to the realization that awareness in the sense that distinguishes us from pre-rational animals presupposes the possession of concepts, took it that that claim committed them to seeing concepts as *innate*—perhaps not all concepts, but at least the most basic or general ones. Sellars shows that that is not so. For he shows how to put together

a) reliable differential responsive dispositions, causally keyed to things, and

b) inferential uses of concepts, which actually apply to those things,

[17] For instance, in his important essay "Inference and Meaning" [ref.].

each of which can be acquired separately, to get the capacity for conscious conceptual awareness of things. He shows us how to build out of those ingredients *non*inferentially elicited reports in which the concepts are applied to the things that causally elicit the reports. In this way he can explain how concepts such as *red* and *green* can be acquired, by a route that does not presuppose preconceptual awareness of red and green things (though it does require the preconceptual capacity to discriminate them, and so to learn reliably to respond differentially to them). That allows him to agree with the empiricists (without indulging in their 'nominalist proclivities'—see [24]-[29]), that "the capacity to have classificatory beliefs of the form '*x* is *F*' is acquired," as he puts the point in [6]. In the rest of the essay, he is going to tell a corresponding story about the concepts *thought* and *sense impression*, ending with our capacity to be directly (in the sense of noninferentially) aware of them.

His question at this point is: If this rationalistic 'psychological nominalism'[18] is right (and Sellars insists that it is), how could we come to have the idea of an inner episode? Descartes thought it a satisfactory answer to this question that we get the idea just by having inner episodes. But this must now be rejected as a sufficient condition of our noticing (being aware of, believing that we have) them, for that is just the Myth. The empiricists thought we could get the concept of thoughts and impressions by abstraction from the thoughts and impressions we were already in any case aware of. That too is a version of the Myth. "In short, we are brought face to face with the general problem of understanding how there can be inner episodes—episodes that is, which somehow

[18] Here 'nominalism' has a sense quite distinct from that invoked in the previous paragraph, picking up on its use in [6]. See [29].

combine privacy in that each of us has privileged access to his own, with intersubjectivity, in that each of us can, in principle, know about the other's." In other words, how could we ever have come to know that reports of the form 'I'm seeing something that looks red,' or 'I am thinking that Vienna is in Austria,' were reliable signs of certain inner facts, given that we can make no empirical correlation by induction as we can with 'This is red'? The Jones myth is the answer to this question—indeed, the only answer available once we have given up both the self-authenticating nonverbal episode notion of Descartes and the empiricists, and the anti-inner-episode strain in Ryle and Wittgenstein (as promised in [10]).

Sellars will "use a myth to kill a myth" [63]. He will tell a story about how a community that turns out always already to have had thoughts and sense impressions might work its way up to having the concepts *thought* and *sense impression*, and then come to be able to apply them noninferentially and so for the first time to notice and be aware of those thoughts and sense impressions. This is explicitly put forward as a myth. Sellars is not claiming that things actually happened this way, that we really had Rylean ancestors, or owe our concepts to a primitive genius (never mind one called 'Jones'). Sellars' pragmatism dictates that issues of conceptual priority be translated into questions of the relative autonomy of different strata of language—that is, into questions concerning what language games can be played independently of and antecedently to which others. Telling an as-if historical, developmental story is a way of exhibiting those relations of conceptual dependency and presupposition.

Part XI [46]-[47] : Thoughts: The Classical View

Section 46: The previous section explained (contra Ryle) that there really are impressions to be accounted for. This section just says the same thing about thoughts. There isn't much in the way of argument here: Sellars points out that it is hard to explain these things away, and we may agree that a theory that can keep them is, other things being equal, superior to one that must deny them.

Section 47: This bit is directly addressed to Ryle, and dismisses his claims that:

a) 'privileged access' must mean invariable access—which Sellars rejects because often someone else can tell what I must have been thinking, even when I am not aware of having thought it, and
b) introspectible thoughts are just *sotto voce* verbal imagery: words running through one's head, 'perceived' as if the words were either heard or seen. (This point is discussed further in [56].)

We have to free ourselves from these preconceptions if we are to understand Sellars' positive story about thoughts and sense impressions.

Part XII [48]-[50] : Our Rylean Ancestors

Section 48: Sellars introduces the notion of 'our Rylean ancestors', who have and can talk about dispositional traits that are relatively long term, the sort of thing for which Ryle's account works well: beliefs, desires, hopes, fears, plans, moods, character traits, etc.. Ryle got these more or less right (we'd still have to put holist qualifications on his atomistic approach), but he injudiciously thought that his success at giving dispositional-behavioral accounts of this sort of mental phenomenon meant that anything that could not be explained this way must be metaphysical and illegitimate. Sellars, via Jones, will show that this is not so. Sellars insists (on the basis of the distinction between dispositions and episodes) that having subjunctive conditionals of the Rylean sort does not yet give the Ryleans the ability to talk about thoughts and experiences. Sellars is going to show what additional conceptual resources they need to develop the concept of *thoughts*, and then on that basis, the concept of *sense impressions*.

Section 49: The problem is, what would have to be added to the Rylean language so that those who speak it "might come to recognize each other and themselves as animals that think, observe, and have feelings and sensations as we use the terms." (The last clause is meant to eliminate the merely dispositionally analyzable bits of mentalistic discourse—items, paradigmatically such propositional attitudes as beliefs and desires, that are psychological, but do not qualify as mental *episodes*.) The first requirement is *semantic* discourse (see [30]). Semantic discourse falls on the side of the epistemic. It is not "definitional shorthand for statements about the causes and effects of verbal performances", although it may have such statements as contingent consequences. Semantic discourse is a kind of *normative* discourse, discussing how expressions *ought* to

be used, or are *properly* or *correctly* used. This is one of Sellars' most fundamental ideas, appearing in nearly all of his earliest essays. ([51] and [52] will tell us about the second requirement.)

Section 50: "My immediate problem is to see if I can reconcile the classical idea of thoughts as inner episodes which are neither overt behavior nor verbal imagery and which are properly referred to in terms of the vocabulary of intentionality, with the idea that the categories of intentionality are, at bottom, semantical categories pertaining to overt verbal performances." This latter idea is that thought must be understood by analogy to talk, in the sense that the concepts we put in play to talk about the meanings or contents of our thoughts are understood in terms of their role in their original or 'home' language game of talking about what we *say*, rather than about what we *think*. (Compare Dummett's commitment to understanding judging as the interiorization of an act of asserting, rather than understanding asserting as the exteriorization of an act of judging.[19])

Part XIII [51]-[52] : Theories and Models

Section 51: Returns to the discussion of theoretical language (discussed under the heading of "Scientific Realism" in Part IX). Theoretical discourse is just a sophistication of a dimension of ordinary empirical language. One way it can arise is by model and commentary. Sellars is telling us this because "the distinction between theoretical and observational discourse is involved in the logic of concepts pertaining to inner episodes."

[19] Michael Dummett, *Frege's Philosophy of Logic* [New York: Harper and Row, 1973] p. 362.

Section 52: So "the second stage in the enrichment of their Rylean language is the addition of theoretical discourse." This matters because Sellars claims that “the distinction between theoretical and observational discourse is involved in the logic of concepts pertaining to inner episodes.”

Part XIV [53]-[55] : Methodological Versus Philosophical Behaviorism

Section 53: Jones is a fore-runner of methodological behaviorism (which is clarified below, and which Sellars endorses).

Section 54: Behaviorists need not present their accounts as analyses of the concepts we already employ, nor need they introduce their theoretical notions by means of explicit definitions. The former would be analytic or logical behaviorism, the latter a kind of instrumentalism. Both are mistakes. Instead, the behavioristic requirement that all concepts should be introduced in terms of a basic vocabulary pertaining to overt behavior is compatible with the idea that some behavioristic concepts are to be introduced as *theoretical* concepts, relative to a behavioral observational vocabulary. This view becomes available once one sees (as we did in [39]-[44]) that the distinction between theoretical and observable objects is methodological, not ontological, i.e. that it has to do with our access to those objects, either purely inferential or also noninferential, and says nothing about the kind of object involved. To say that they are theoretical concepts in this sense is to say that (at this stage in the development of the language

game) they can only be applied as the conclusions of inferences. Thus they are not equivalent to any descriptions of behavior (which could be applied observationally). This idea is one of Sellars' cardinal innovations.

Section 55: Behaviorism in this methodological sense is compatible with physicalism, since the theoretical concepts it employs might turn out to refer to neurophysiologically describable items (just as 'Pluto', introduced as a name for whatever is perturbing the orbit of Neptune, might have turned out to apply to an astronomical ball of cheese). But it is also compatible with denying such physicalism. Behaviorism and physicalism are two different and independent sorts of commitment.

Part XV [56]-[59] : The Logic of Private Episodes: Thoughts

Section 56: Jones's model for thinking is inner speech. His commentary ensures that this is not conceived of as verbal imagery. What is objectionable about the verbal imagery proposal (introduced in [47]) is that it consists in the use of a quasi-perceptual model: hearing the wagging of an inner tongue.

Section 57: The model carries the applicability of semantical categories over from overt utterances to thoughts; so thoughts can be 'about' things.

Section 58:

1) This Jonesean theory is compatible with dualism as well as with materialism.

2) Inner episodes are to be unobservable the way molecules or the cause of the crack in the dam are, not the way ghosts are. That is, we happen (at this stage in the story) not to be able to report them noninferentially, though there is nothing that rules out such observation in principle. Thus they might turn out to be identical to physiological events. Nevertheless, at this point only the third person use is available, even for characterizing our own episodes.

3) One can't think until one has learned to speak—one can't assert anything 'mentally' (think to oneself that…) until one has caught on to the social practice of public assertion. Thus talk is prior to thought in the order of explanation. Once one has learned simultaneously to talk and think, however, thought often precedes talk in the order of causation.

4) So the notion of language having a meaning, being 'about' things, is not to be explained in terms of thoughts having meanings (for instance, in the Cartesian or Lockean fashion). The project must explain the meaning of thoughts in terms of the meaning of talk, which must be explained some other way (e.g. in terms of social practices).[20]

5) Jones does not think of these episodes as immediate experiences, that is things to which thinkers have privileged access, since he doesn't have this concept yet. His episodes are 'inner' only in the mundane sense of 'under the skin'.

Section 59: But it turns out that when Jones teaches his theory to other people they "can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe [their own] overt behavior." That is, one can develop a

[20] [ref. the Sellars-Chisholm correspondence]

conditioned reflex in someone (perhaps depending on some ultimately discoverable neurophysiological event related to his thought) to report noninferentially what heretofore could only be inferred. "What began as a language with a purely theoretical use has gained a reporting role." It might not have turned out this way. But insofar as Jones's theory is a good one (a question in principle independent of the eventual identifiability of these episodes with ones characterizable in neurophysiological terms), his fellows were *already* reliably differentially responding to these episodes. So one would expect they would be able to learn to expand their differential responses to include reports. This story explains why: recognizing "that these concepts have a reporting use in which one is not drawing inferences from behavioral evidence, [the account] nevertheless insists that the fact that overt behavior is evidence for these episodes is built into the very logic of these concepts, just as the fact that the observable behavior of gases is evidence for molecular episodes is built into the very logic of molecule talk."

Part XVI [60]-[63] : The Logic of Private Episodes: Impressions

Section 60: Jones now does for sense impressions what he previously did for thoughts. This category presupposes the category of thoughts. We start from a sub-class of thoughts called 'perceptions'. Seeing that something is the case is an inner episode in the Jonesian theory, which has as its model reporting on looking that something is the case. But these perceptions are not yet sense impressions. We still have a kind of *claim*, something in the epistemic order, not a kind of *particular*, something in the causal order. To get sense impressions we need the notion of a 'state of the perceiver' common to those

occasions when the perceiver is right and those occasions when he's wrong about there being something red and triangular. This will be the 'intrinsic characterization' of impressions that Sellars talks about in the third paragraph of [45], and [22]. Here is an outline of the theory of perception on the causal side that was appealed to in [7].

Section 61: Where thoughts were modeled on sentences, impressions are modeled on pictures, or more generally, replicas, which are particulars. The essential feature of the model is that visual impressions stand to one another in a system of ways of resembling and differing that is structurally similar to the ways in which the colors and shapes of visible objects resemble and differ. That is, there are states of the perceiver which, though neither red nor triangular, have features (call them 'of-red' and 'of-triangular') that are isomorphic to the kinds of features visible physical objects have. This is a sort of functionalism about sense impressions. The occurrence of these replicas is to be understood as a nonepistemic relation of particulars (which neurophysiology or dualistic mind science might further specify for us). "Thus the model for an impression of a red triangle is a red and triangular replica, not a *seeing* of a red and triangular replica," which would be an epistemic affair. Their overall explanatory role can be summarized thus : It is sense impressions "which (from the standpoint of the theory) are being responded to by the organism when it looks to the *person* as thought there were a red and triangular physical object over there."

Section 62: This section just does for sense impressions what [59] did for thoughts. It points out that people can be trained to develop conditioned reflexes for reporting these

theoretical entities called 'impressions'. (Perhaps some neurophysiological mechanism will be discovered eventually that explains the acquisition of such responsive dispositions.) At this point, since Jones's students can make noninferential reports of their sense impressions as well as their thoughts, they are directly (in the sense of noninferentially—the only sense available once the Myth of the Given has been rejected) aware of both sorts of inner episode. In the case of sense impressions, this is awareness of the impressions "of the sort which is common to those experiences in which we either see that something is red and triangular, or something merely looks red and triangular, or there merely looks to be a red and triangular object over there [45]." Such noninferential reports of sense impressions, reports of the form "I am now sensing a sense impression of a red triangle," are quite different from those made using 'looks', which were considered in the first half of the essay. A noninferential report using 'looks' takes a 'that' clause as its content-specification, and indicates the inferential potential that is being forwarded as a candidate for endorsement. A noninferential report of a sense impression takes a description of a particular of a sort modeled on replicas as its content-specification, and indicates the causal antecedent common to reports of how things are and of how things look. (Recall the diagnosis of [7].) Running the two together would re-enact the Myth of the Given. These two essentially derivative and parasitic strata of language, both centering on noninferential uses, express different aspects of perceptual experience. The conceptual awareness of sense impressions that Sellars has now made available in an unmysterious and unthreatening way is the "something more" that (according to the opening sentences of [16]) our perception involves, besides endorsement of propositional contents "wrung [noninferentially] from the perceiver by the object perceived." They

are what was promised in the first paragraph of [45]. The sense impressions of which we are aware (once both the concept of sense impressions and the corresponding noninferential reporting practices are fully in place) *explain* the fact that "when I say '*X* looks green to me now'…my experience is, so to speak, intrinsically, *as an experience* indistinguishable from a veridical one of seeing that *x* is green [16]." For both sorts of speech act arise as the result of exercising reliable differential dispositions to respond to the presence of sense impressions—as they did already before Jones gave us the concepts without which we could not be aware of them. All that needed to be added to those responsive dispositions is the new concept *sense impression*, with the kind of inferential articulation appropriate to its model of *replicas* of, e.g., visible surfaces.

Sellars has now completed his task. We now have recipes telling us how to diagnose and treat the Myth of the Given in all its multifarious manifestations, whether what is given shows up in the guise of particulars whose occurrence entails knowing or believing something (e.g. sense datum theories), or in the form of noninferentially acquired propositionally contentful beliefs (e.g. what is expressed by 'looks' talk). Epistemologically foundationalist appeals to the given of the Cartesian sort have been shown to fail because *non*inferential uses of concepts (no matter whether their subject matter is construed as 'inner' or 'outer') turn out to presuppose *inferential* uses of concepts. Empiricist appeals to the preconceptual given to explain concept acquisition (whether by abstraction or otherwise) fail because "We now recognize that instead of coming to have a concept of something because we have noticed that sort of thing, to have the ability to notice a sort of thing is already to have the concept of that sort of

thing, and cannot account for it [45]." Nonetheless, Sellars has shown us how we can make sense of the idea that we have direct awareness of mental episodes (the applications of inferentially articulated *concepts* of thoughts and sense impressions elicited noninferentially *by* thoughts and sense impressions), including the limited but very real privileged access each of us has to such inner episodes, without committing ourselves to the Myth of the Given.

Section 63: Grand Finale