

October 16, 2019

Week 8 Notes
 On “Phenomenalism” as Complementing and Completing the
 Argument of *EPM*

1. Recap:

a) Lessons from Sellars’s analysis of “looks-Φ.”

Relation of characterization of object-language *use* (what one is *doing* in applying object-language locutions expressing object-language concepts) in *pragmatic* metalanguage to characterization of *contents* expressed by (the use of) object-language locutions in *semantic* metalanguage is unobvious and complex.

Here: What one is *doing* is

- i) manifesting a reliable differential responsive *disposition* to make a claim (endorse a claimable content), while
- ii) substituting for the usual committive response a *noncommittal* response, overtly *withholding* the usual endorsement one is acknowledging being tempted by, on the basis of one’s collateral commitments concerning the standardness of conditions for observation, that is, reservations about how reliable a reporter one is in *these particular* circumstances.

We do not have a good meta-meta-vocabulary for discussing these relations. We are in the position of trying to think about examples of this phenomenon, and work our way towards a general vocabulary and account of it.

One big constraint (but also clue) is the Frege-Geach point. In addition to the free-standing use of expressions to make claims (claimables in claimings), there are **embedded uses**, from which the pragmatic force is (usually: exceptions include use of objectionable vocabulary like ‘Boche’, where the problem is with the content) stripped off, the standard assertional speech act not performed. Note that for “looks”-talk: I pointed out that Sellars’s account *does* tell us how to think about 3rd person *attributions* of “looks”. A: “It merely *looks* to John as though the tie were green.” This is ambiguous between a report of John having said “The tie looks green,” and him having said “The tie is green,” when *I* know that it is blue. What *I* am doing in the first case is *attributing* the withholding of a commitment, and in the second case *I* am withholding that commitment.

Sellars thinks of these expressions as *not descriptive*. One is *doing* something other than *describing* in performing the speech acts in which they appear asserted. And that is right. But he concludes that they do *not also* perform any descriptive function, or have any descriptive content. It is not so clear that that is not right.

A more promising strategy, I want to claim, is

α) to handle the embedded uses by attributing descriptive content to the claimable being used *in the first instance* non-descriptively, and

β) somehow (!) derive a residual, parasitic, secondary descriptive content from the principal expressive use of the expression. (This is just a hand-wave in the direction of a strategy: the suggestion of an approach.)

b) Foundationalism:

Principal lesson here is that non-inferentially arrived-at observation reports do not form a semantically autonomous stratum of language—a language-game one could play though one played no other, a set of uses of concepts one could engage in though one engaged in no other (inferential) uses.

They accordingly cannot play a *semantically* foundational role of any sort that requires such autonomy.

This is a *pragmatic* dependence claim, that underwrites a *semantic* dependence claims.

And this overarching claim about the dependence of non-inferential uses on inferential uses holds for observation reports *no matter what their subject-matter*. It does not in any way affect this argument if what is being reported is “inner episodes,” or experiences specified wholly in phenomenological or phenomenalist vocabulary (e.g. of how things merely seem—but this is just one example).

The large claim, about the *pragmatic*, and therefore *semantic* dependence of noninferential reporting practices on inferential uses of the same concepts, is the master argument against the Myth of the Given. Denying it, for any subject-matter the noninferential reports might have, whether outer or “inner”, is the Myth of the Given. This is the most general formulation of it.

Within the strictures of this insistence on the pragmatic and *therefore* semantic interdependence of noninferential and inferential uses, this appreciation of the *essential* role played by inferential involvements in the specifically *conceptual* contentfulness of any states (Rylean) or episodes (Jonesean)—the kind of contentfulness that is *cognitively* significant, because relevant to the role of such contentful items as *evidence*, as providing *reasons* for believing or doing—there is ample room for insisting on and articulating the distinctive and essential role played by noninferential reports in empirical knowledge.

Thus a carefully circumscribed empiricist epistemology can survive these observations, including views that articulate what is right about the original foundationalist epistemological enterprise.

c) Methodological behaviorism,

This is behaviorism without instrumentalism about theoretical entities. It is a real and important innovation of Sellars. It is epitomized in the analogy and the slogan: the mental is to the behavioral as theoretical entities are to observable ones. **Key here is the distinction between Rylean psychological states and Jonesean mental episodes.**

The reason behaviorists had not pursued methodological versions of behaviorism is the air of paradox associated with thinking of mental episodes (conceived of in Cartesian terms) as *theoretical*. For they were the paradigm of the observable. One wants to say, they are *obligatorily observational*: not only could one observe them, one *had to* (if they occurred), and indeed, if one did *not* observe them, they did not occur. And one could not be mistaken in one’s observation of them. In a characteristic exercise in the “new way of words” Sellars deploys his account of “looks-Φ”-talk to deflate this

incorrigibility and transparency. And he sees it as well downstream of the proper, methodological behaviorist way of understanding the conceptual *explanatory* role of thoughts and sense impressions.

Rorty's social pragmatist account of Cartesian **incorrigibility** as an optional structure of authority complements and completes this line of thought.

And Evans's suggestion as to how to understand the relation of **secondary-quality concepts** to the concepts of mere observables does the same, along a different but related dimension.

[I should rehearse this story (the account of givenness, foundationalism, and methodological behaviorism) as part of the run-up to my discussion of Rorty's eliminative materialism as a crucial stage in the development of his pragmatism, in my Spinoza lectures and the ensuing OUP publication of them.]

2. “**Phenomenalism**”: Here there are two points to emphasize.

- a) Sellars offers *another* anti-foundationalist argument here. It extends and deepens the argument of *EPM*.

I think of it as an argument for the claim that **even if, per impossibile, a stratum of phenomenal concepts and awareness that consists in applying such concepts could be semantically autonomous**—that is, that such concepts *could* intelligibly be the *only* concepts some subject could deploy [might mention Condillac's statue thought-experiment here]—**such an autonomous stratum of concepts and awareness still could not serve to so much as make intelligible our grasp of objective concepts: the concepts of objects and properties that can exist unobserved.** This includes having unobserved parts (interiors, backsides) and existing even when unobserved.

[I might just observe, without Freudian comment or judgment, that Critical Realists like Sellars's father Roy Wood were more or less obsessed with backsides.]

The strategy for doing such a construction—aiming, I emphasize, at making *objective* concepts *intelligible* in terms of *phenomenal* concepts (call this “the phenomenalist *semantic* project”)—is to use (infinite sets of—Lewis's “nonterminating judgements”) *subjunctive conditionals*, concerning what one *would* experience *if* one *were* in certain circumstances.

So order of exposition should be:

- 0) Suppose, for the sake of argument, that phenomenal concepts and the judgements that consist in applying them *could* intelligibly be supposed to be an autonomous discursive stratum—that it is intelligible to suppose that someone should be able to deploy *all* and (the crucial point) *only* concepts of this kind.
- 1) **The phenomenalist semantic project (PSP).** This must be understood as in the first instance a semantic project, of explaining how objective concepts are so much as *intelligible*. Only afterwards, secondarily, and downstream from that semantic project is there (can there be) an *epistemological* project concerning the *justification* of objective *claims* in terms of phenomenal evidence.
- 2) **The subjunctive strategy (SS)** for pursuing the phenomenalist semantic project.

This is really the “only straw floating,” the only idea *anyone* has had for how to bring off the PSP. (People find this thought in Berkeley and Hume, for instance.)

3) **Sellars’s fork (dilemma), the form of his argument against the SS for the PSP.** The two alternatives he considers concern the concepts employed in the *antecedents of the subjunctive conditionals in question*. (We suppose that the *consequents*, which would provide confirmatory evidence for objective claims, the circumstances of application of objective concepts, are couched in phenomenalist terms.) **Either they are or they are not expressible entirely in phenomenal terms.**

i. If not, then

- α) They could be **true**, but
- β) **they would not support the PSP.**

For, so construed, the subjunctive strategy would itself presuppose the intelligibility of the objective concepts deployed in articulating the antecedents of the subjunctive conditionals: “If I were to *turn my head*,” (“walk around the table,” “cut into the apple with a knife,”....).

ii. If so, then though

- β) **they would be suitable for use in the PSP**, unfortunately,
- α) they would **not be true**.

This last is the load-bearing member in Sellars’s argument.

He claims that there are no true, *subjunctively robust generalizations*, expressible wholly in phenomenal terms, except those that are what Hume would call “relations of ideas”: e.g. that if I *were* to be appeared to redly, I *would* be appeared to colorfully. But *not* “If I seemed to cut into the apple, I *would* seem to see white pulp rather than red skin.” For, notoriously, not all seemings *are* expressions of objective regularities: error, illusion, hallucination, dreaming....

I want to say that **this argument substantially supplements the argument of EPM concerning the nonautonomy of phenomenal concepts** (indeed, even of observational uses of objective concepts), and that these *two* arguments really drive a stake through the heart of the phenomenalist project. No-one should ever be tempted by it again.

But Frank Jackson is committed to it (in his Locke lectures), and it seems that David Chalmers is, too (though he is cagier about it than his teacher).

b) Sellars on *surfaces*.

The concept surface can be understood in two importantly different ways. Familiar but dangerous philosophical mistakes ensue if they are not clearly distinguished.

i. The ordinary idea that things have surfaces, and that in the ordinary case *in some sense* what is visible to us of things is their surfaces. If they are opaque, we cannot see the interiors of things such as apples. We see the red shiny surface, but not the white pulpy interior. The backs of things are not visible, not part of the instantaneously, visually presented surface available from the visual perspective occupied at a moment.

- ii. The philosophical idea that all we can strictly be said to see *is* surfaces. Everything else is known to us only inferentially, by drawing conclusions about non-surfaces from evidence that consists entirely of our knowledge of surfaces. This is the idea of the world being epistemically “painted in secondary qualities.” Sellars notes that in case (i), we can peel off the red surface, the skin, of the apple, and see what is beneath it. It has a thickness, and itself has a backside. In case (ii), the idea is, if we skin the apple, all we get is *another* surface. And surfaces in *this* sense do *not* have a thickness, or a back. They are, if one thinks about it, quite peculiar entities. How could they *not* have backs? *All* there is to them is the surface. How is that so much as *intelligible*?

And here Sellars has an answer. And *that* he has an answer, and *what* it is amount to a substantial *confirmation* of the Myth of Jones account of sense impressions. So in this way, too, “Phenomenalism” continues and extends the project of *EPM*.

For Sellars’s account of Jones’s invention (postulation) of sense impressions explains the peculiar “backsidelessness” of surfaces thought of in the phenomenalist way. Recall that Jones postulates sense impressions as having properties that are systematically analogous to, and similarly structured as (so *homologous* to) shaped and colored items. They are not red, but red*, not triangular, but triangular*.

The explanation of why we make perceptual mistakes is that the causal chain that mediates our RDRDs has as proximal members, red* triangulars*, which are *usually* brought about by red triangles upstream. But if they are not, we are still disposed to respond to *them* by reporting “red triangle.”

But the model comes with a commentary. Only *some* characteristics are carried over from the model to what is modelled on (understood in terms of) it. Recall that thoughts, modelled on overt linguistic claimings, are *not* to be thought of as “the wagging of an internal tongue.” *That* feature of the model has no analog in the *use* of the model. And just so for sense impressions. **The analogous, homologous features include only a surface*, which does not come with a back*, or a thickness*.** Those features are **simply not appealed to in the use made of them, according to the commentary, which tells us how to use the model. That is why they are surfaces without backs.**

And that is why the story the phenomenalist tells about surfaces in the second, philosophically loaded sense, is so much as *intelligible*.

But it is also why the phenomenalist’s *projection* of such surfaces into the objective world, treating them as “out there” as the only kind of object that is *really visible*—the phenomenalist’s *reification* of surfaces—is a mistake.

This is a diagnostic argument, by which Sellars explains the mistake (here, a confusion of two good ideas—compare §7 of *EPM*) behind a philosophical temptation. This can be laid alongside the account of looks-talk, and why its “incorrigibility” is *not* of use in epistemological foundationalist projects.

- 3. Recall that I want us to think about these different arguments that Sellars is making—all 5 of the arguments I’ve rehearsed here, (1a) on “looks”, (1b) on foundationalism, (1c) on methodological behaviorism, (2a) on the phenomenalist semantic project, (2b) on surfaces, both a

critical diagnosis of a mistake and a constructive account or explanation—on two different levels.

- I want us to learn appropriate philosophical lessons from them. They are richly resonant and suggestive, both for how we think about the philosophical topics they address and for the illumination they provide about the history of philosophy.
- But I also want us to be thinking about the *kind* of move he is making, the *way* he is thinking—not just to get clearer insights into the thinking of a great philosopher (I hope it is becoming clearer why I count him as one), but because making moves in some sense *like* these is what we should all be trying to do ourselves.

On CDCM:

Aside on every sufficiently serious reading being a rewriting.

CDCM is the *first draft* of an absolutely fabulous paper.

Outline:

4. From labeling to description: the necessity of location “in a space of implications.”
5. They must be subjunctively robust implications, of the sort expressed explicitly by subjunctive conditionals.
6. That is why “the descriptive and explanatory [expressive] resources of the language advance hand-in-hand.” It is essential, and not just accidental, to *descriptions* that they can play a particular role in *explanations*. That role is both as *explanans* and as *explanandum*, explaining and explained. Subjunctive conditionals, Sellars thinks, codify *rules* that license specifically *explanatory* inferential connections.
7. How this point matters for our understanding of possible worlds semantics.
8. His next move is to invoke the Kantian idea of categories, giving it a Carnapian, metalinguistic twist. The expressive function characteristic of specifically alethic modal vocabulary, whose paradigm for Sellars is subjunctive conditionals, is to allow us to make explicit, in the form of a claimable (conceptual content), our endorsement of a subjunctively robust implication. “Explicitness” here means claimability: something that can both serve as and stand in need of reasons, something that can be endorsed, challenged, and justified. This is to be a location in a space of reasons, or a position in a game of reasons. It is something that can be both a premise and a conclusion of an inference, the starting-point and terminus of a language-language move. (Note that, in the context of the way he has drawn a line around *describing*, in distinguishing it from mere labeling or classifying, one would think that this means it *does* have descriptive content. But he thinks that the fact that it is in a *metalinguage* undercuts this conclusion. He still sees (will always see) playing a descriptive role and playing a metalinguistic, rule-codifying role, as mutually exclusive.

9. The Kant-Sellars thesis about modality:

The use of alethic modal vocabulary is *elaborated from* and *explicative of* (“LX for”) the use of ordinary empirical descriptive (OED) vocabulary.

- i. The first (L) means that in knowing how to use *any* OED vocabulary, one already knows how to do everything one *needs* to know how to do in order to use alethic modal vocabulary. This is a sense in which the capacity to deploy alethic modal vocabulary is *a priori*. In this sense, the *a priority* claim is not that one could use it *first, before* using OED vocabulary. It is rather that there are no *particular* concepts or judgments expressed in OED vocabulary that one must have first. *No matter what* OED vocabulary one has learned to deploy, one has *thereby* learned how to do everything one needs to know how to do to use alethic modal vocabulary. This, I would claim, is the recoverable core of Kant’s notion of what is knowable *a priori*.
- ii. The second (X) means that the expressive function of modal vocabulary is of a certain metalinguistic kind: making explicit something that is implicit already in the use of *another* vocabulary: OED vocabulary. (See (i).)

So the *same* implicit abilities that are *elaborated into* the capacity to use alethic modal vocabulary (subjunctive conditionals) are what are *explicated* by that vocabulary.

Explicitness here is as above in (8): “Explicitness” means claimability: something that can both serve as and stand in need of reasons, something that can be endorsed, challenged, and justified. This is to be a location in a space of reasons, or a position in a game of reasons. It is something that can be both a premise and a conclusion of an inference, the starting-point and terminus of a language-language move.

Note that I will claim that there is a corresponding Kant-Sellars thesis about *normative* vocabulary.

10. Eventually: I am working my way up to (though most of this might have to wait for Week 9) saying how I think we can make progress on the challenge of completing and complementing an expressive story in a pragmatic metavocabulary about what we are *doing* in using alethic modal expressions (paradigmatically subjunctive conditionals) “free standing”, and what we are *saying* by doing that. This last is a *descriptive* component, which will be appealed to *inter alia* in explaining *embedded* uses. My story is in terms of normative governance and epistemic tracking.

In more detail:

11. Four moves:

- a) *Labeling* (mere classifying) vs. *describing*.
- b) *Describing* and *explaining* as two sides of one coin, intelligible only in terms of their relations to one another.
- c) *Explanation* essentially depends on *subjunctively robust implications*.
So, so does *description*.
- d) *Subjunctive conditionals* (the paradigm of alethic modal vocabulary)
 - i. in *one* sense codifies *lawlike* relations of *physical necessitation*, and

ii. in *another* sense expresses endorsement of a *rule* of inference.

I want to say that (ii) expresses what one is *doing* in asserting a subjunctive conditional.

That is something one would express in a *pragmatic* metalanguage.

Supposing that that is correct, the challenge is then to say what one is *saying*, the conceptual content one is assertorially endorsing (if any) in *doing* that.

This is something one would express in a *semantic* metalanguage.

I suggest that is what the sense of “codification” in (i) is after.

12. From Labels to Descriptions:

Labeling, nominalism, and the classificatory theory of consciousness:

a) What I'll call “*semantic nominalism*” was the universally held semantic theory until Kant, and is still probably dominant. It holds that:

i. Proper names are labels, stuck onto or otherwise associated with the objects named. This is the ‘Fido’/Fido theory.

ii. Predicates are like names, but they are *general* labels, labels that are stuck onto or otherwise associated with *many* objects (those they are true of), via the *properties* of those objects. They are general in that we stick them onto more than one thing. These labels specify *properties* of labelable objects, or their *kinds*. (Depending on whether they are sortals or not, that is, whether in addition to criteria of application—see below—they have criteria of identity and individuation—as Frege taught us in the *Grundlagen*. Cf. Sellars claim here: “The point is the more radical one that the relation of a thing-kind word to the criteria for belonging to that kind of thing is different in principle from the relation of words for characteristics of things to the criteria for the presence of these characteristics. “Lemon” and “bald” may both be vague, but they are so in radically different ways.” [§46] [But bracket all these considerations.]

Semantic nominalism is the view that the relation between a *name* and its *bearer*, what is a name of, construed on the model of *labeling*, is the paradigmatic semantic relation. Predicates name (label) properties, and sentences name possible states of affairs.

b) Complex labels are *descriptions*. So something can be *described* by pinning on it (associating with it in whatever the way distinctive of *semantic* association is) the labels ‘red’, ‘juicy’, ‘apple’.

c) Language, accordingly, consists of a bunch of *descriptive* terms, labels. And what one *does* with language is to *describe* things. The result is a picture of language as essentially a system of *classification*. The idea that this is what language is is *descriptivism*.

d) **So semantic nominalism is a principal route to descriptivism about language: the view that what language is for is to describe the world.**

e) Notice that the **language-as-labeling view is semantically atomistic**: Applying the label ‘red’ is *independent* of applying the label ‘apple’. Even though there turn out to be, as a matter of fact, connections between applying the labels ‘ripe’, ‘Macintosh’, and ‘red’, that sort of fact is not an *essential* feature of the semantic connection between the labels and what is labeled. For the relation between one label and what it labels does not depend on the relation between any other label and what it labels.

- f) ***Mere labels: Consider a tray of disparate objects, each of which is labeled with either a blue or a red dot. They have been labeled. Have they been described? Evidently not. For what have they been described as?***
- g) One way of seeing that such mere labels don't *mean* anything (or at least, that we don't *understand* them) is that we have no idea how to *go on* labeling things with red and blue dots. If a few more objects are added to the tray, we don't know which, if either, label is appropriate. The mere labeling of some objects does not establish a *standard*, *norm*, or *practice* we can appeal to in determining how it would be *correct* to continue labeling new objects. (This is *one* of the threads Wittgenstein is pursuing in his discussion of "going on in the same way" in the *PI*.)
- h) Be that as it may, for our purposes, we may conclude that the *only* labels that have any prospect of counting as *descriptions* are those associated with reliable differential responsive dispositions to apply them to new cases. These are *projectable* labels. They must be associated, explicitly or implicitly, with *standards*, or *norms*, or at least learnable-teachable *practices* that settle when it is and when it is not *correct* to *apply* the label to new cases. At the least, some notion of *mislabeling* must have been put in place, for labels to be even candidate descriptions. Descriptive terms, unlike mere labels in the thinnest sense, must at least come with *circumstances of appropriate application*.
- i) So, if we have such RDRDs embodying standards of appropriate application, will *such* labels be descriptions? It is not hard to see that they will not. 'Gleeb'. **Consider possession of an infallible 'gleebness' tester.** Point the device at something, and it lights up if and only if the thing is *gleeb*. This, by hypothesis, is projectible. It establishes a standard, with respect to which things can be *mislabeled* as 'gleeb'. But when one has found out that something is *gleeb*, what has one found out? (One can know what is a K without knowing what a K is.) One has not *described* it, but *merely labeled* it. One knows what things are *gleeb*, but has not thereby found out *anything* about them, since one does not, we want to say, yet know what *gleebness is*. Once again we can ask: What is it you are describing things *as* when you label them as 'gleeb'? The conclusion is that it is *not* enough to have a *description* of something that one have not only a label that has in the past been applied to some things and not to others, but also a *reliably differential responsive disposition* to discriminate things to which the label is and is not correctly applied.
- j) What more is needed? Consider a *parrot* who can respond to the visible presence of red things by uttering tokens of "Awrk! That's red!" And suppose that he does so in just the same circumstances in which *we* do. He reliably differentially responds to red things by correctly applying a vocalized label. [Kvetch about 'vocal' vs. 'verbal'.] Is he *describing* things as red? The noise he makes is *just* a noise to him, as 'gleeb' was to us. For the parrot, that label is not something that *contrasts* with other labels in that it excludes their proper applicability. And the applicability of that label does not have any further consequences, for instance, making further labels such as 'colored' appropriate.
- k) **Conclusion: For a description, we must have both circumstances of appropriate application and appropriate consequences of application.**
- l) **"It is only because the expressions in terms of which we describe objects, even such basic expressions as words for the perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label."¹**
- m) We saw that Sellars says [§107] that what is needed is that the description be put in a "space of implications". **We need inferential articulation to have description.** To be more than a *mere* label, the label must be one that one can offer *reasons* for applying in one case and withholding in another (corresponding to the *circumstances* of appropriate application) and whose

¹ Pp. 306-307 (§107) in: Wilfrid Sellars: "Counterfactuals, Dispositions, and Causal Modalities" In *Minnesota Studies in the Philosophy of Science, Volume II: Concepts, Theories, and the Mind-Body Problem*, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell (Minneapolis: University of Minnesota Press, 1958), p.225-308.

applicability can itself offer reasons for the application of *other* characterizations (corresponding to the appropriate *consequences* of application). This is what a *parrot* lacks, who can reliably differentially respond to red things by saying “That’s red” in the same circumstances we do, but who does *not*, just on that basis, count as *describing* anything as red, or *reporting* or *observing* that something is red. To be a *description*, the *label* must be situated in a web of *connections* to *other* labels/descriptions. And those connections are broadly *inferential*: a matter of what is *evidence* or *reason* for or against what, of what *obliges* one to apply further labels/descriptions, or *precludes* one from doing that. Absent that context, labels are not descriptions. That is what Sellars means by saying:

“It is only because the expressions in terms of which we describe objects, even such basic expressions as words for the perceptible characteristics of molar objects locate these objects in a space of implications, that they describe at all, rather than merely label.”

n) Since every expression must have *both* *circumstances of appropriate application and appropriate consequences* of application, each incorporates an *inference*: an inference *from* the obtaining of the *circumstances to* the obtaining of the *consequences*.

13. WS’s point about *describing/explaining* is an expression of the relation between *asserting/inferring*. The most basic point is that these are activities specifiable only in relation to one another.

This is the point that is the link between the diagnosis of the difference between labeling/describing and connection between describing/explaining.

It will then be a further claim that explaining requires not just a “space of implications” but a space of implications that are *subjunctively robust*.

14. One theme for first discussions of CDCM is his methodology of looking at what we are *doing* in *using* expressions of a certain kind, and then drawing conclusions about what we are or are not *saying* thereby. This is a kind of *pragmatism*. Pragmatism in *this* sense consists of giving explanatory priority to *pragmatics* over *semantics*.

Two varieties of this sort of explanation are:

- Looking at pragmatic *dependence* relations and drawing conclusions about *semantic* dependence relations. “looks-Φ” is a prime example.
- But the *interdependence* of *description* and *explanation* is another. For this is really (‘in’ / ‘ed’) about describing and explaining. Those are different doables, but one cannot globally be able to do one without globally being able to do the other.

The conclusion he formulates is then “the descriptive and explanatory resources of the language advance hand in hand.” He means *expressive* resources, and he is going to be asking about how we should understand the expressive resources provided by explanation (explaining): what we are *doing* when we *use* those explanatory resources. And his claim is going to be that we are being *semantically self-conscious* in a distinctive way: discussing what *rules* we should employ in reasoning/inferring.

The result, he says, should be understood as a kind of *reflective equilibrium* of (covertly metalinguistic) *rules* and particular language-language *moves*.

The rules both *govern* the moves and articulate and express them.

Cf. the *EPM* passage from sec. 38:

“For empirical knowledge, like its sophisticated extension, science, 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.”

- So this argument goes from the *activities* of *labeling* vs. *describing* (the one built on, and so dependent on, the other) to the *activities* of *describing* vs. *explaining*.

- d) Note that the conclusion WS arrives at from his discussion of labeling vs. describing is that in order to be more than mere labels in the sense that they have or express *descriptive conceptual content*, OED expressions must be “located in a space of implications.” This is on the side of *semantics*. But it is a semantic conclusion drawn from pragmatic premises.
- e) Again, he will conclude that the *material* implications involved must be *subjunctively robust*. The argument for *semantic* conclusion has to do with the *pragmatics of explanation* (explaining).

15. The argument for the *semantic* conclusion that the material implications that articulate the content of OED vocabulary whose use is to *describe*, not just to *label* (on the side of *pragmatics*) must be *subjunctively robust*, that is, must support (license) subjunctive and counterfactual reasoning [brief word on the difference] is that explaining and describing “go hand in hand,” are two sides of one coin, *activities* such that although any given act can be one rather than the other (explaining is a *move* and describing is a *position*), one cannot have the capacity to engage in the one kind of activity unless one has the capacity to engage in the other.

This is true whether we think of things from the side of the individual *abilities* or from the side of the social *practices*.

16. So what is the argument that the activity of explaining requires implication relations that are subjunctively robust? (See also the material on this point at (25) below.)

- a) First, I think, is the idea that, as Ryle puts it in his classic essay “If, So, and Because,” making an inferential move involves endorsing the license for it (the conditional that makes explicit one’s entitlement to the move) as expressing a *general* rule or principle, one that extends beyond this particular case. It involves “exhibiting it as an instance of a *pattern of good reasoning*.” Ryle’s example of a case that does not explicitly involve a quantifier, but is general nonetheless, is “If today is Wednesday, tomorrow will be Thursday.”

For Sellars, this is to say that it must be a *rule* that articulates the entitling license. The rule might *contingently* only apply to one instance, but its *rulishness* is inherently more general.

For Sellars, this is a Kantian point. Kantian concepts are *general*, they are *universals*, not just in the sense of being predicates rather than singular terms, but in the sense of embodying *rules* that can govern *multiple* judgments.

Here should acknowledge the three in fact orthogonal dimensions that Kant assimilates under the technical rubric of *concepts vs. intuitions*. Can think of him as just *confusing* these, or can think of him as having a substantive view that identifies them or at least lines them up. (I think the latter formulation is probably hermeneutically correct.) But they do not in fact line up, so this is not something we should follow him in.

Nonetheless, the underlying point about a kind of generality as a rule for making judgments, not just *this* judgment, *can* be underwritten. For it is a consequence of Gareth Evans’s **“Generality Constraint.”** Subsentential expressions *must* be capable of showing up in, contributing to the content of, be combinable with, *multiple* other components in determining the contents of more complex expressions.

But the current point at issue is more basic still. (The “Generality Constraint” on subsentential expressions is the manifestation of a more basic phenomenon at the level of *sentence* usage in asserting (describing) and inferring (explaining).)

For there is something wrong with implications asserted to be *merely* “hic et nunc,” as Abelard terms them. Insofar as such implications/inferences are intelligible at all, it is only as a degenerate limiting case. What is behind Kant’s “concepts are general in the sense of universal” thought is that the *license* *entitling* one to the conclusion, given (entitlement to) the premise must license something more than just *this particular* move. Otherwise, one is not really *justifying* the move (which one *would* be, if exhibiting it as good as an instance of a *pattern* of good inferences, a pattern of good inferences) by appeal to something *else*. One is *just* stamping one’s foot and insisting that *this* move is good.

Talk of (explicit) *rules* standing behind as licensing (governing, in the dual sense of providing standard for normative assessment of and being something the performances are subjunctively sensitive to) performances has to be cashable in terms of talk of (implicit) *norms* doing so. So what is wrong with a *norm* that *in principle* (not just as a matter of contingent fact) licenses or governs only *this one, particular, unrepreatable performance*?

What is the deep basis of this claimed fundamental *generality of norms*?

Note that even if we read “general” as *much* weaker than Kant’s “universal,” this question is still one whose answer will put us on the road that *can* lead to Kant’s claim, and so to his formal account of morality. (Note: I think we should take over Kant’s *categorial* conception of morality: as commitments that are implicit in engaging in any discursive activity at all. This is a different route that *could* lead to a generality or universality claim.)

It is equally a claim about the *generality of reasons*, since we are talking about the form of a *reason* for doing something. Such a reason must normatively license the doing, and it is such reasons that we are supposed to be subjunctively sensitive to (i.e. are rational insofar as we are so sensitive).

- b) The point here, I think, is that just as
 - i. One must, in order to count as able to use a descriptive term, be able to distinguish, in practice, however fallibly, between implications in which it occurs essentially that are and those that are not materially good. If one makes *no* such distinction, one is *not* using the expression to *describe*, but at most to label or classify.
 - ii. One must, in order to count as able to make a certain inference (endorse an implication) make *some* distinction, however rough and fallible, between additional auxiliary hypotheses (collateral premises) that *would*, and those that would *not*, infirm the implication.

These are two dimensions of what description and explanation (respectively) require.

- c) *Hic et nunc* implications are ones that one *both* treats as good *and* claims that *any* alteration of circumstances not explicitly included in the premises would infirm it.

We don't know what "follows" means if absolutely *any* and *every* variation or alteration in circumstances would infirm the implication.

These are the limiting case of *minimal* ranges of subjunctive robustness.

Candidate:

"If I look at this cow right now, under these particular circumstances, it looks really goofy."

I might claim to have *no* views about what one could change from the actual background circumstances without possibly infirming this conditional. Would the cow still look goofy in this light if it were Friday? Maybe my mood is crucial, and would be different.

Compare this to ordinary inferences (appeals to reasons).

If I light this dry, well-made match, *then* it will light.

I must know at least that there are conditions (the absence of oxygen, the presence of a strong magnetic field) that *would* infirm the implication.

And perhaps more obviously, I *must* know that there are *other* conditions that would be irrelevant: the day of the week, the exact height off the ground of a beetle on a tree on a different continent...

"If that lioness is too sick to hunt, she will starve to death."

I must know that this will still be true if it were Tuesday, if someone on a different continent stumbled on a curb, and would be true even if Julius Caesar had parted his hair differently, and would *not* be true if the lioness were in the mean time struck by lightning or rescued and fed by well-wishers...

The question is whether I can claim to even *understand* what is expressed by a conditional, never mind be entitled to the moves it licenses, if I have *no* idea at all about what collateral circumstances would and would not infirm it.

To be finitely statable, implications and inferences *must* leave the range of subjunctive robustness *implicit*. That is what is acknowledged by *ceteris paribus* clauses—which do not *cure* this difficulty, but merely *acknowledge* it.

So we can say that in endorsing a move, one must at least have *some* idea of what is acknowledged by the "other things being equal," bit—or at the least, know that there is such a distinction.

This is an essential feature of *material* implications. In general, we have no idea how to substitute for these *formally* good implications, which would *not* require even implicit *ceteris paribus* clauses.

d) This kind of "generality" of concept-use *is* on Kant's mind as an essential, categorial feature of concepts. How it relates to the "universality" he sees as part of the form of *reasons* (considerations that *can* function as reasons, which is essential to being conceptually articulated). See (a) above.

17.

a) Sellars's argument that having descriptive conceptual content (being OED vocabulary) *necessarily* involves standing in *subjunctively robust* implication relations to other such vocabulary is what I've called the "Kant-Sellars thesis about modality." For it follows

that *alethic modal* vocabulary (paradigmatically, subjunctive conditionals) makes explicit a feature of the framework of practices and abilities that makes empirical concept use possible. That is to say that these are *metalinguistic* concepts of the sort that Kant called “categorical,” or “pure concepts of the Understanding.”

- b) The insight that Sellars got from Carnap—though Carnap was never in a position to put it this way himself, and as a self-hating neo-Kantian (rejecting that label, even while doing things that justified its application to him)—and which opened his eyes to the possibilities afforded by the “new way of words,” (what struck *him* down on the road to Damascus) was precisely that Kant’s insight into the existence of framework-explicating, *non-descriptive* concepts (that is, concepts whose *principal* expressive use is not description but framework-explication) can fruitfully be understood to be *metalinguistic*.
- c) The issue I have been raising throughout is whether the metalanguages or metavocabularies involved should be understood in the first instance (that is, to begin with) as *pragmatic* metalanguages, or *semantic* ones. And I’ve claimed
 - i. first that the answer is, as Sellars *earlier* realized, but for some reason soon afterwards lost sight of or rejected [but why?] *pragmatic* metalanguages, and
 - ii. second, that there are intricate issues raised then about the relations between what one says in pragmatic and semantic metalanguages—issues that come up *only* if one worries about the sort of language-use that Sellars focuses on. And
 - iii. third, that the conceptual issues Sellars would wrestle with, mostly unsuccessfully, for the rest of his career stem precisely from these issues about the relations of the contents of claims in pragmatic and semantic metalanguages. (These are the issues I aimed to provide a framework and a regimented idiom for discussing, in my Locke lectures.)

18. This K-S thesis about modality has radical consequences. It is diametrically opposed to the conceptual framework within which possible worlds semantics domesticated alethic modality by offering an *extensional* metalanguage in which one could do *intensional* semantics—which is its great achievement.

For that strategy depends on starting from a notion of *non-modal* properties—what I shall call ***modally insulated* properties**. These are ones that, speaking in the terms of the framework, apply or don’t apply to things in one possible world independently of how things are in any other possible world. **Being spherical or made of copper** might be paradigms. The idea is that worlds can be identified and individuated entirely in terms of such intrinsic properties. Then, on the basis of that structure of worlds, we can introduce *modal* properties, whose possession by something in *one* world depends on what is true in *other*, related (accessible) worlds. Maybe being fragile (or water-soluble) in a sense that can apply to a diamond that is created and then melted in a volcano, without every having the opportunity to be shattered.

But the modal K-S thesis about OED vocabulary tells us that the notion of such properties makes no sense. To be determinately contentful as a description, an expression *must* be involved in *subjunctively robust* implications.

And I think that when we look at properties such as being spherical and being made of wood (or copper) that is just what we find. **All basic empirical predicates and relations have**

modal necessary conditions. That is, they do not apply to something unless other claims about how it *would* behave (or *would have* behaved) if some other conditions obtained are also true. Even being triangular has consequences for what *would* happen if one correctly counted its vertices, measured its angles, intersected it with a circle, and so on. Set membership *can be* “extensional” in the sense of modally insulated (depending on how the set is defined). But basic empirical properties (i.e. those not defined by using some such apparatus as set-theory) are not.

19. I will claim later, when we discuss Sellars’s *invidious Eddingtonian scientific naturalism*, that the modal K-S thesis also has radical consequences concerning the possibility of *cross-sortal identities*, of a kind that are fatal to essential elements of Sellars’s assimilation of the scientific image / manifest image distinction to Kant’s distinction between noumena and phenomena. The identities in question would connect sortals in the scientific framework to those in the manifest framework. The consequence I claim the modal K-S thesis has is that *no* such identities are ultimately so much as *intelligible*, never mind true. As such, they *could not* be true. If that is right, then Sellars’s reductive program is broken-backed and (so) ill-conceived. It is *this* tension that I claim keeps the strategy that organizes Sellars’s systematic ambitions from being realizable.

20. Challenges to the Metalinguistic Strategy:

Sellars’s attempt to make the metalinguistic understanding of modality work, in spite of the fact that in *saying* “All As are necessarily Bs,” for instance, “All samples of copper necessarily melt at 1084°C,” one is evidently *not saying that* “ $\forall x$ It is appropriate to infer B(x) from A(x),” i.e., as he put it earlier, propounding “a rule for the use of the expressions ‘A’ and ‘B’.” That is evident for two reasons:

- a) One could understand an expression of the law in another language, and have no views at all about the expressions ‘A’ and ‘B’.
Conversely, one could understand that Turks endorse the inference from applying ‘A’ to applying ‘B’ *without* understanding *that* they claim all A’s are B’s—because one did not understand their terms ‘A’ and ‘B’. [These should be Quinean corner-quotes.]
- b) The modal claim is evidently not *about* the use of linguistic expressions. It could be true even if there were not expressions ‘A’ and ‘B’—no term “copper” and no Celsius scale for measuring temperature.

These are the arguments from translation or language-relativity, and the argument from the wrong counterfactuals being underwritten, respectively.

21. One principal difficulty stems from the fact that although Sellars may be right that what one is *doing* in making a modal claim is endorsing a pattern of inference, it is clear that one is not thereby *saying that* an inference is good. When I say “Pure copper necessarily conducts electricity,” and thereby unrestrictedly endorse inferences from anything’s being pure copper to its conducting electricity, I have nevertheless *said* nothing about any inferences, explanations, justifications, or implications—indeed, have said something that could be true even if there had never been any inferences or inferrers to endorse them, hence no describers or discursive practitioners at all.² These two observations set the principal criteria of adequacy both for

² Sellars connects this obvious fact with the observation that:

Sellars's positive working-out of the pragmatist-inferentialist treatment of modal vocabulary, and for his argument that the purely descriptive base vocabulary invoked by the empiricist critic of the semantic credentials of modal vocabulary lacks the sort of discursive autonomy the empiricist criticism presupposes and requires.

Sellars's central rhetorical strategy in this essay is to address the issue of what is expressed by modal claims about necessary connections by offering:

... a sympathetic reconstruction of the controversy in the form of **a debate between a Mr. C (for Constant Conjunction) and a Mr. E (for Entailment) who develop and qualify their views in such a way as to bring them to the growing edge of the problem.**

Officially, he is even-handed in his treatment of the vices and virtues of the empiricist, who denies that the use of modal vocabulary can express any legitimate semantic content beyond that expressed by a descriptive, extensional universal generalization, and the rationalist, who understands that content in terms of entailments expressing rules of reasoning. In fact, however, as becomes clear when he launches into his own account, he is mainly concerned to develop a version of the rationalist account. As the second half of the essay develops, Sellars's marks his abandonment of the disinterested pose by an uncharacteristically explicit expository shift:

It is now high time that I dropped the persona of Mr. E, and set about replying to the challenge with which Mr. C ended his first critique of the entailment theory.³

Doing that requires careful investigation of the differences between and relations among four different sorts of item:

- **Practical endorsement of the propriety of an inference from things being A to their being B;**
- **The explicit statement that one may infer the applicability of 'B' from the applicability of 'A';**
- **The statement that A physically entails B;**
- **The statement that As are necessarily Bs.**

The first is the sort of thing Sellars takes to be pragmatically presupposed by the activity of describing, that is, deploying descriptive vocabulary. The second fails to capture such practical endorsements, because of the possibility of asserting such statements regarding the *expressions* 'A' and 'B' without understanding what they express.⁴

Idealism is notorious for the fallacy of concluding that because there must be minds in the world in order for *us* to have reason to make statements about the world, therefore there is no sense to the idea of a world which does not include minds. [CDCM Section 101]

³ CDCM Section 85. In fact, Sellars's 'defense' of Mr. C (see the passage from Section 82 quoted in note 3 above) consists of showing what concessions he needs to make to Mr. E. This proceeds first by Mr. C's qualification that "'A causes B' says that $(x)[Ax \rightarrow Bx]$ and *implies* that he latter is asserted on inductive grounds" [62], followed by the necessity of conceiving "of induction as establishing principles *in accordance with which* we reason, rather than as major premises *from which* we reason." [83] As will appear, the former concession, introducing the notion of what is *contextual implied* by contrast to what is explicitly said, is then dialectically made available to be pressed into service by Mr. E. This bit of dialectic is a pretty rhetorical flourish on Sellars's part, but I doubt that in the end it reflects any deep feature of the confrontation between the empiricist and rationalist approaches to modality.

⁴ As Sellars says:

But one can know that Turks, for example, ought to withdraw '...' when they commit themselves to '---' without knowing the language, whereas the statement that 'p entails q' contextually implies

The third sort of statement expresses Mr. E's initial stab at an analysis of the fourth. It is the answer to the question: what sort of entailment is it that modal statements are supposed to express?:

Mr. E has a ready answer. ...it might...be called '**natural**' or '**physical**' entailment, for while any entailment is a logical relation, we can distinguish within the broad class of entailments between those which are, and those which are not, a function of the specific empirical contents between which they obtain. The latter are investigated by general or formal logic (and pure mathematics). Empirical science, on the other hand, to the extent that it is a search for *laws*, is the search for entailments of the former kind. (Putative) success in this search finds its expression in statements of the form 'It is (inductively) probable that A physically entails B.'"⁵

The virtue of statements like "A physically entails B" is that they do plausibly codify the practical endorsement of an inference that is implicit in what one does in the form of something one can explicitly *say*, without bringing in irrelevant commitments concerning particular expressions, the activity of inferring, or discursive practitioners. The remaining difficulty is that they seem plainly not to have the same content, not to say the same thing, as explicitly modal statements of objective necessity.

Sellars's response to this problem is to acknowledge that modal statements do not say that some entailment holds, but to distinguish between what is *said* by using a bit of vocabulary and what is '*contextually implied*' by doing so. Sellars says very little about this latter notion, even though it bears the full weight of his proposed emendation of the rationalist account. It is recognizably the same distinction he had appealed to earlier, in "Inference and Meaning", as the distinction between what one *says* by making a statement and what one thereby *conveys*. **There his example is that in asserting "The weather is fine today," I say that the weather is fine today, but convey that I believe that it is fine.**⁶ That otherwise uninterpreted example suggests to me that what Sellars has in mind is the distinction between *semantic* and *pragmatic* inferences. That is the distinction between inferences underwritten by the *contents* of what is *said* or asserted, on the one hand, and inferences underwritten by what one is *doing* in saying them, on the other. The inference from "The weather is fine," to "It is not raining," is of the first sort; the inference from my asserting "The weather is fine," to "Brandom believes the weather is fine," is of the second sort. Inferences of these two kinds may generally be distinguished by the Frege-Geach embedding test: look to see whether those who make the inference in question also endorse the corresponding conditional. "If the weather is fine, then it is not raining," is generally true, while "If the weather is fine, then Brandom believes it is fine," is not generally true. (Compare the inference from my *saying* "That is an ugly tie you are wearing," to "Bob is annoyed with me.")

If that is in fact the distinction Sellars is after, then it seems to me that the view he is expounding and defending can be put less paradoxically if we do not take a detour through entailment statements, but concern ourselves directly with the relation between the endorsement of patterns of inference and modal statements. The underlying rationalist

that the speaker not only knows the language to which 'p' and 'q' belong, but, in particular, knows how to use 'p' and 'q' themselves. [CDCM Section 81]

⁵ CDCM Section 56.

⁶ Sellars, "Inference and Meaning", p. 280/332 in in J. Sicha (ed.) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars* [Ridgeview Publishing Company, Reseda CA, 1980]—hereafter *PPPW*.

insight is a pragmatist-inferentialist one: what one is *doing* in making a modal claim is endorsing a pattern of inference. Modal vocabulary makes possible new kinds of *sayings* that have the *pragmatic effect* of endorsing inferences. To say that is not yet to say what they say, it is only to say what one is *doing by saying them*. [Cf. Section 8 below.] But it does settle the *pragmatic significance* of such modal claims, in the sense of their appropriate circumstances and consequences of application.

It is the attempt to specify this peculiar and distinctive sort of pragmatically mediated relation between vocabularies that leads Sellars to say things like:

It is sometimes thought that modal statements do not describe states of affairs in the world, because they are *really* metalinguistic. This won't do at all if it is meant that instead of describing states of affairs in the world, they describe linguistic habits. It is more plausible if it is meant that statements involving modal terms have the force of *prescriptive* statements about the use of certain expressions in the object language. Yet there is more than one way of to 'have the force of' a statement, and failure to distinguish between them may snowball into a serious confusion as wider implications are drawn. [CDCM Section 81]

and

Shall we say that modal expressions are metalinguistic? Neither a simple 'yes' nor a simple 'no' will do. As a matter of fact, once the above considerations are given their proper weight, it is possible to acknowledge that the idea that they are metalinguistic in character oversimplifies a fundamental insight. For our present purposes, it is sufficient to say that the claim that modal expressions are 'in the metalanguage' is not too misleading if the peculiar force of the expressions which occur alongside them (represented by the 'p' and the 'q' of our example) is recognized, in particular, that they have 'straightforward' translation into other languages, and if it is also recognized that they belong not only 'in the metalanguage', but in discourse about *thoughts* and *concepts* as well. [CDCM Section 82]

and

We must here, as elsewhere, draw a distinction between what we are committed to concerning the world by virtue of the fact that we have reason to make a certain assertion, and the force, in a narrower sense, of the assertion itself. [CDCM Section 101]

If one practically endorses the pattern of inference that treats classifying or describing anything at all as an A as sufficient grounds ("all on its own", as Sellars says, in order to capture the way the pattern of inferences in question is counterfactually robust) for concluding that it is a B, then one is committed to the claim that all As are necessarily Bs. And commitment to that claim is commitment to practically ratify that pattern of inference. Assuming, as Sellars has claimed, that using ordinary, non-modal, descriptive vocabulary requires practically endorsing such patterns of inference ("situating descriptions in a space of implications"), that means that **anyone who has the practical ability to deploy "purely descriptive" vocabulary already knows how to do everything he needs to know how to do to deploy modal vocabulary as well**. He need not actually do so, since practically undertaking those inferential commitments does not require that one have available a language with vocabulary permitting one to *do* that by *saying* something. But *all* a practitioner lacks in such a circumstance is the *words* to hook up to

discriminative and responsive abilities he already possesses. In this precise sense, the ability to deploy modal vocabulary is *practically implicit* in the ability to deploy non-modal descriptive vocabulary.

22. Sellars strategy to avoid them is to acknowledge that they are not metalinguistic in Carnap's and Tarski's sense of syntactic or semantic metalanguages. They accordingly do not *say* anything *about* linguistic expressions. Nonetheless, it is in terms of the *use* of expressions in *other* vocabularies that we are to understand them. I do not know why he does not appeal here to the notion of pragmatic metalanguage that he introduces in his earlier work, for that seems clearly what is wanted.

What he *does* do is appeal to a notion of what is “conveyed” by saying something that is distinct from what is “said” by saying it.

I unpack this in terms of the distinction between *pragmatic* and semantic *implications*: what follows from one's saying something (a doing, an act) and what follows from the content of what is said.

What seems right about Sellars's modal expressivism is:

- a) His account of what we are *doing* in *using* modal expressions (asserting subjunctive conditionals), and
- b) The idea that to understand that use one must look at its relations to (how it governs) the use of *other* vocabulary: OED vocabulary.

The crucial fact in the vicinity is that there is no other vocabulary whose use stands to the use of OED vocabulary as the use of OED vocabulary stands to the use of modal vocabulary.

This fact, I want to say, is what gets *misunderstood* philosophically in the idea that possible worlds could each be *entirely specified* in vocabulary whose use had no intensional or modal necessary conditions. Note that this diagnosis would be very Sellarsian. Compare §7 of *EPM* on the confusion of two individually OK lines of thought, and the corresponding argument about two conceptions of *surfaces* in “Phenomenalism.”

23. Methodology and overview of Sellars's project:

On the *twin* difficulties of:

- a) **Descriptivism**, and
- b) **Abstract entities**

as *together* contributing to empiricists' “nothing-but-ism” in all its forms: emotivism [about norms], behaviorism [about the mind], phenomenism [in semantics and epistemology]:

103. It is...important to realize that **the presence in the' object language of the causal modalities** (and of the logical modalities and of the deontic modalities) **serves not only to express existing commitments, but also to provide the framework for the thinking by which we reason our way** (in a manner appropriate to the specific subject matter) **into the making of new commitments and the abandoning of old. And since this framework essentially involves quantification over predicate variables, puzzles over the 'existence of abstract entities' are almost as responsible for the prevalence in the empiricist tradition of 'nothing-but-ism' in**

its various forms (emotivism, philosophical behaviorism, phenomenism) as its tendency to assimilate all discourse to describing.

This is a *very* important methodological statement of how Sellars understood his overarching project circa 1957-1959 (the “wonder years”, *anni mirabile*). Though he was not (yet) writing about emotivism, it ties together his treatments of behaviorism in the philosophy of mind and phenomenism in semantics and epistemology, as *targets*. In the context of *this* essay it goes without saying that epistemological and ontological puzzles about *modal* discourse, which lead to empiricist nothing-but-ism in the form of Hume-Quine retreat to mere regularities, belongs in this target area.

His idea is that they *share* descriptivism and a puzzlement about abstract entities. His task is to show us how to be anti-descriptivists, and nominalists about abstract entities (thus assuaging empiricist epistemological worries and naturalist ontological worries), in *both* cases by understanding the discourses in question as *metalinguistic*.

24. This introduces the topic of the relation of what is expressed by modal vocabulary to *change of meanings*.

This is further the connection to *induction* as a distinctive mode of inference:

“The question as to the applicability of the terms ‘true’ and ‘false’ to the conclusions of inductive inferences is considerably more complex.” [§91]

“[T]he ‘conclusions’ of primary non-statistical inductions are decisions to espouse inference tickets.” [§92]

The solution of this puzzle lies in the fact that the logic of variables and quantification involves not only the *momentary* crystallized content of the language at a *cross section* of its history, but also its character as admitting—indeed demanding—modification, revision, in short, development, in accordance with rational procedures. In the case of variables the values of which are **descriptive constants**, these rational procedures can be summed up in the single word ‘Induction.’ [§105]

My general claim is that we should get clear about the previous topics before taking on induction as the empiricist’s account of our epistemic access to *laws*, and how that is consonant with understanding them as *rules of inference*.

25. Some notes on the technical language of CDCM;

- a) Sellars uses *Principia Mathematica* dot notation in lieu of (most) parentheses. And his use of it is not flawless. (I think he leaves a dot out of the last formula in §43, for instance.) Bernard Linsky’s Stanford Encyclopedia article on the Notation of *PM* offers a good primer.
- b) He carefully distinguishes between *subjunctive conditionals* (“if x were [phi]d it would [psi]”) and *counterfactual conditionals* (“if x had been [phi]d it would have [psi]d”).
- c) His analysis of dispositional talk essentially involves distinguishing four kinds of expressions: thing-kinds (sortals, not just predicates, and which typically do not take temporal qualifications), conditions (predicates which do take temporal qualifications), interventions ([phi]ing) and results ([psi]).

d) He distinguishes between *dispositions* and *capacities*: capacity claims say that there is a condition and an intervention that will have a result, while disposition claims presuppose that the condition obtains.

More material:

26. **Counterfactual robustness of the ‘implications’:**

a. The explanation-supporting space of implications that articulates the conceptual contents of ordinary empirical descriptive (OED) vocabulary **must be *counterfactually robust* inference/implications.**

That is, the commitment involved in using the descriptive expression in question is that if anything *were to* satisfy the circumstances of its application, it *would* satisfy the consequences. For otherwise, the term cannot be applied to *new* cases. For one would need to find out in advance if the inference held in *that* case. But the point is that the circumstances of appropriate application need *not* include checking whether the consequences also obtain.

b. In general, **one cannot count as understanding any descriptive expression (or the concept it expresses, what it describes something as being) unless one distinguishes at least *some* of the inferences it is involved in (some of the connections within the “space of implications” it is situated in) as *counterfactually robust*, at least in the sense that they would *remain* good inferences if some further premises were added that do not in fact obtain.** Thus one must know such things as that a lion would still be a mammal if the lighting were slightly different, it were a different day of the week, it was transported to a zoo, we clipped its fur....

c. Hume found that even his best understanding of actual observable empirical *facts* did not yield an understanding of *rules* relating or otherwise governing them. Those facts did not settle which of the things that *actually* happened *had* to happen (given others), that is, were (at least conditionally) *necessary*, and which of the things that did *not* happen nonetheless were *possible* (not ruled out by laws concerning what did happen). Though initially couched as an *epistemological* question about how one could *know* what rules or laws were in play, Hume’s worries run deeper, raising the *semantic* question of what it could so much as *mean* to say that the facts are governed or related by rules or laws. Hume (and, following him, Quine) took it that epistemologically and semantically fastidious philosophers faced a stark choice: either show how to explain modality in nonmodal terms or learn to live without it, to do what we need to do in science without making such arcane and occult supradescriptive commitments. But that challenge is predicated on the idea of an independently and antecedently intelligible stratum of empirical discourse that is purely descriptive and involves no modal commitments, as a semantically autonomous background and model with which the credentials of modal discourse can then be invidiously compared.

d. One of Kant’s most basic ideas, revived by Sellars, is that this idea is mistaken. **The ability to use ordinary empirical descriptive terms such as ‘green’, ‘rigid’, and ‘mass’ already presupposes grasp of the kind of properties and relations made explicit by modal vocabulary.** Sellars summed up the claim admirably in the title of one of his early papers:

“Concepts as Involving Laws, and Inconceivable Without Them.”⁷ This slogan is a good place to start in thinking about Kant’s point, but in fact Sellars’s own view is subtly but importantly different from Kant’s. For Sellars, the laws determining the truth of counterfactuals involving the application of a concept are part of the content of the concept. For Kant, modal concepts make explicit not something implicit in the *content* of determinate concepts, but something implicit in their *empirical use*, in *applying* them to make empirical *judgments*. That is why the pure concepts of the understanding—what he calls ‘categories’, such as possibility and necessity—both are to be understood in terms of the forms of judgment (the table of categories derives from the table of judgments) and express synthetic, rather than analytic necessities. From Kant’s point of view, a better slogan than Sellars’s would be “The *Use* of Concepts in Empirical *Judgments* as Involving Laws and Inconceivable Without Them.”

- e. Kant was struck by the fact that the essence of the Newtonian concept of mass is of something that by law *force* is both necessary and sufficient to *accelerate*. And he saw that all empirical concepts are like their refined descendants in the mathematized natural sciences in this respect: their application implicitly involves counterfactual-supporting dispositional commitments to what *would* happen *if*.... **Kant’s claim, put in more contemporary terms, is that an integral part of what one is committed to in applying any determinate concept in empirical circumstances is drawing a distinction between counterfactual differences in circumstances that *would* and those that *would not* affect the truth of the judgment one is making.** One has not grasped the concept cat unless one knows that it would still be possible for the cat to be on the mat if the lighting had been slightly different, but not if all life on earth had been extinguished by an asteroid-strike.
- f. **Grasp of a concept is mastery of the use of a word**, Sellars says. And for descriptive concepts, that use includes not only sorting inferences (however fallibly and incompletely) into materially good and materially bad ones, but also, among the ones one takes to be materially good, to distinguish (however fallibly and incompletely) between counterfactual circumstances under which they do, and counterfactual circumstances under which they do not, *remain* good. Part of taking an inference to be materially good is having a view about which possible additional collateral premises or auxiliary hypotheses would, and which would not, infirm it. Chestnut trees produce chestnuts—unless they are immature, or blighted. Dry, well-made matches strike—unless there is no oxygen. The hungry lioness would still chase the antelope if it were Tuesday or the beetle on the distant tree crawled slightly further up the branch, but not if lioness’s heart were to stop beating. The point is not that there is any particular set of such discriminations that one must be able to make in order to count as deploying the concepts involved. It is that if one can make *no* such practical assessments of the counterfactual robustness of material inferences involving those concepts, one could not count as having mastered them.
- g. Sellars says (in the Introduction to CDCM), that “the framework [note the word] of what objects of a certain kind K *would* do in circumstances C is *basic*.” (Q: In what sense ‘basic’? And why?) We have seen, in effect, that the implications in which genuine descriptive terms (as opposed to mere labels) are involved are *counterfactually robust*. That is, they must extend to *possible* cases. This is just another way of saying that there must be a *norm* or *standard* for the *correct* application of the term in cases that have not actually arisen. We are seeing the general shape of an argument that **modality (what is expressed by modal vocabulary, such as that used to express subjunctive conditionals—one kind of counterfactually robust conditional, as we shall see) is implicated in the framework that**

⁷ Reprinted at pp. 87-124 in J. Sicha (ed.) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars* [Ridgeview Publishing Company, Reseda CA, 1980]—hereafter *PPPW*.

makes *description* possible. Cf. Sellars's essay "Concepts as Involving Laws, and Inconceivable Without Them." And it is this same line of thought that will implicate *explanation* with *description*.

h. There are *two kinds* of counterfactually robust conditionals, which Sellars distinguishes [ref. in CDCM] as:

Genuine subjunctive conditionals, and

Subjunctive identicals.

[Explain this difference:]

Consider the difference between:

Copper (all samples of copper) melts at 1084° C.. and

All the coins in my pocket are copper.

The first, (iii), supports the subjunctive conditional:

If this coin (which is in fact a nickel) *were* copper, it *would* melt at 1084° C..

The second, (iv), does *not* support the corresponding subjunctive conditional:

If this coin (which is in fact a nickel) *were* in my pocket, it *would* be copper.

(iv) *does*, however support the claim

If I *were* to choose a coin at random from my pocket, it *would* be copper.

So we cannot distinguish between the "accidental generalization" or "contingent regularity" (iv) and the "lawlike statement" (iii) by saying something like "only laws support counterfactuals". (Though you do hear things like that a lot.)

Rather, we must distinguish what *kinds* of counterfactuals they support. Sellars does that by introducing the idea of "subjunctive *identicals*". The idea is that *all* the counterfactuals supported by statements such as (iv) must be derived via *identities* of the objects they concern with some of the members of the class *actually* picked out by the description in the generalization. The generalization (iv) supports counterfactuals concerning objects identical with one of the coins that in fact are in my pocket. It does *not* support counterfactuals concerning what would happen if *other* things *did* fit that description, i.e. *were* in my pocket. So, if any coin at all, in *any* possible world, *were* identical to one of the coins that is *actually* in my pocket, it *would* be copper. For all *those* coins, the ones *actually* in my pocket *are* copper. But the generalization in (iv) tells us *nothing at all* about any coins in other worlds just on the basis of the fact that in *that* world, *their* world, the description "coin in Bob's pocket" applies to them. Sellars gets at this difference in the significance of the two sorts of claims—genuine subjunctive *conditionals* and mere subjunctive *identicals*—by saying that the former, but not the latter, tell us something about what follows from the applicability of a description "*just on the basis of the applicability of that description*."

Sellars's distinction between *subjunctive identicals* and *subjunctive conditionals* is not nearly as straightforward as he suggests.

- a) When it is drawn, as I did, in terms of the example of "All of the coins in my pocket are copper," and "Copper melts at 1084° C.," it seems sharp enough. Though both do support counterfactuals, the former supports counterfactuals about objects that are *in fact* in my pocket, and the latter about *any* copper.
- b) But if we look at generalizations such as "The Baltic is less salty than the Atlantic," and "Homeotherms have higher metabolic rates than poikylotherms," we seem to have intermediate cases. Past, probable-future, and at least some non-actual water samples drawn from the two oceans will conform to the former. And the latter, while almost certainly not an exceptionless generalization across *possible* species (and perhaps not even for some actual dinosaurs) is true of more than just the currently extant species.
- c) But it seems to me I did—briefly—say the right thing about this distinction, while not emphasizing that saying that involves conceiving of it somewhat differently from the way Sellars presents it. The issue is whether there is some specification, typically a description, such that one proceeds by determining what things *actually* satisfy that description, and then looking at counterfactuals that are true of *them* in other worlds. That is a *de re* procedure. Or whether one looks at what things satisfy that specification in *other* worlds, and then at the counterfactuals true of *them*. That is a *de dicto* procedure. Almost any given generalization can be read either way—that is, is subject to both kinds of readings. We can think of the distinction, accordingly, not so much as concerning the generalization as concerning the connection between it and counterfactuals that it supports. Sellars is just pointing out that his claims—paradigmatically about what *induction* supports—are intended to address only the counterfactuals *de dicto* that a generalization supports.
- d) This is a distinction of the order in which two operations are performed (so, using the precise word for such an issue, a "scope distinction"). If the generalization is $\forall x[Fx(\Box) \rightarrow Gx]$:

- i. One can *first* figure out which things are F in *this* world, and *then* follow *them* to other possible worlds in order to assess the truth-values of counterfactuals involving *them*—the *de re* method of assessing counterfactuals on the basis of that generalization; or
- ii. One can *first* move to the world with respect to which one wants to assess the truth-value of some counterfactual, and *then* see what things are F in *that* world.
- e) Carnap (and Russell, I think)—Hempel discusses this view in *Aspects of Scientific Explanation*—argues that genuine laws, of the kind (the best sort of) science aims at and employs in its explanations, do not make any reference to *particular* objects, events, and so on. Thus, they do not refer to the Earth, to North America, to Julius Caesar, to me, or to now. The point is sometimes put by saying that they are “purely universal”. But the distinction aimed at *could* be the one above—or rather, the distinction between extracting counterfactuals from generalizations in the *de re* way and in the *de dicto* way could be what lies behind the thought Carnap (and Russell, and others—Reichenbach, Nagel?) were expressing. For what they forbid, in the statement of laws, is the use of *proper names*, *demonstratives*, and *indexicals*, all of which are (and are paradigms of) *rigid designators*. And that is to say that they *must* be read *de re* (and so have problematic A-intensions, in Jackson’s sense). If this is right, then there is a progression of interpretation:
 - i. Carnap (and the others) offer as a *necessary* condition of being a “law-like” rather than an “accidental” generalization, that proper names, demonstratives, and indexicals, as expressions picking out particular things, do not occur in the statement of law-like generalizations.
 - ii. Sellars renders this distinction in terms of that between “subjunctive identicals” and “subjunctive conditionals”, namely, in terms of whether counterfactual consequences can only be derived via *identities* linking the objects of the counterfactual with objects the antecedent predicate of the generalization actually applies to.
 - iii. I then read Sellars’s point in terms of the order of application (hence *scope*) of two operations: moving to a counterfactual world, and applying the predicate of the antecedent.
- f) The next question, then, is: when the distinction is read as I suggest above, in what sense is it *true* that induction supports only *de dicto* counterfactuals?
- g) The idea is that if we look not at the probability of *all* of the coins in my pocket being copper being increased by checking them, one-by-one (since finding that the 9 checked so far of the 10 are *does* increase the probability that *all* of them are, assuming independence, and regardless of the antecedent probability of each), but at the probability that the *next* coin in my pocket that we check will be copper, if the generalization is really accidental, then we can look to the case where each is independent, and there is some antecedent probability for each, and finding that the first 9 checked of the 10 are copper says *nothing* about the probability of the 10th. To think that it does is the gambler’s fallacy. (Compare coin-flipping.)
- h) But does this point really generalize as it is claimed to (on the reading I offered) to *de re* and *de dicto* counterfactuals supported by generalizations generally?