

## On the Way to a Pragmatist Theory of the Categories

### 1. Introduction

Several decades ago, Richard Rorty suggested that philosophical admirers of Wilfrid Sellars could be divided into two schools, defined by which of two famous passages from his masterwork “Empiricism and the Philosophy of Mind” are taken to express his most important insight. The two passages are:

In the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not" (§41).

and

[In] characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says” (§36).<sup>1</sup>

The first passage, often called the “*scientia mensura*,” expresses a kind of scientific naturalism. Its opening qualification is important: there are other discursive and cognitive activities besides describing and explaining. The second passage says that

---

<sup>1</sup> In Herbert Feigl and Michael Scriven, eds., *Minnesota Studies in the Philosophy of Science*, vol. I (Minneapolis: University of Minnesota Press, 1956); reprinted in Sellars’s *Science, Perception, and Reality* (London: Routledge and Kegan Paul, 1956; reissued Atascadero, Ridgeview, 1991); reprinted as a

characterizing something as a knowing is one of them. And indeed, Sellars means that in characterizing something even as a believing or a believable, as conceptually contentful at all, one is doing something other than describing it. One is placing the item in a normative space articulated by relations of what is a reason for what. Meaning, for him, is a normative phenomenon that does not fall within the descriptive realm over which natural science is authoritative.

Rorty called those impressed by the scientific naturalism epitomized in the *scientia mensura* “right wing Sellarsians,” and those impressed by the normative nonnaturalism about semantics expressed in the other passage “left wing Sellarsians.” Acknowledging the antecedents of this usage, he used to express the hope that right wing and left wing Sellarsians would be able to discuss their disagreements more amicably and irenically than did the right wing and left wing Hegelians, who, as he put it, “eventually sorted out their differences at a six-month-long seminar called ‘the Battle of Stalingrad.’” According to this botanization, I am, like my teacher Rorty and my colleague John McDowell, a left wing Sellarsian, by contrast to such eminent and admirable right wing Sellarsians as Ruth Millikan, Jay Rosenberg, and Paul Churchland.

While I think Rorty’s way of dividing things up is helpful (there really are “41-ers” and “36-ers”), I want here to explore a different perspective on some of the same issues. I, too, will focus on two big ideas that orient Sellars’s thought. I also want to say that one of them is a good idea and the other one, on the whole, a bad idea—a structure

---

monograph, with an Introduction by Richard Rorty and a Study Guide by Robert Brandom (Cambridge, Mass.: Harvard University Press, 1997).

that is in common between those who would self-identify as either right- or left-wing Sellarsians. And the one I want to reject is near and dear to the heart of the right wing. But I want, first, to situate the ideas I'll consider in the context of Sellars's neo-Kantianism: they are his ways of working out central ideas of Kant's. Specifically, they are what Sellars makes of two fundamental ideas that are at the center of Kant's transcendental idealism: the metaconcept of *categories*, or *pure concepts of the understanding*, and the distinction between *phenomena* and *noumena*. The latter is a version of the distinction between appearance and reality, not in a light epistemological sense, but in the ontologically weighty sense that is given voice by the *scientia mensura*. I cannot say that these fall under the headings, respectively, of What Is Living and What Is Dead in Sellars's thought, since the sort of scientific naturalism he uses to interpret Kant's phenomena/noumena distinction is undoubtedly very widespread and influential in contemporary Anglophone philosophy. My aim here is to explain what I take it Sellars makes of the most promising of these Kantian ideas.

When asked what he hoped the effect of his work might be, Sellars said he would be happy if it helped usher analytic philosophy from its Humean into its Kantian phase. (*A propos* of this remark, Rorty also said, not without justice, that in these terms my own work could be seen as an effort to help clear the way from analytic philosophy's incipient Kantian phase to an eventual Hegelian one.<sup>2</sup>) Sellars tells us that his reading of Kant lies at the center of his work. He used that theme to structure his John Locke lectures, to the point of devoting the first lecture to presenting a version of the Transcendental Aesthetic with which Kant opens the *Critique of Pure Reason*. Those lectures, published as

*Science and Metaphysics: Variations on Kantian Themes*, are Sellars's only book-length, systematic exposition of his views during his crucial middle period. The development of Kantian themes is not only self-consciously used to give that book its distinctive shape, but also implicitly determines the contours of Sellars's work as a whole. I think the best way to think about Sellars's work is as a continuation of the neo-Kantian tradition. In particular, I think he is the figure we should look to today in seeking an appropriation of Kant's theoretical philosophy that might be as fruitful as the appropriation of Kant's practical philosophy that Rawls initiated. On the theoretical side, Sellars was the greatest neo-Kantian philosopher of his generation.<sup>3</sup>

In fact, the most prominent neo-Kantians of the previous generation: C. I. Lewis and Rudolf Carnap were among the most immediate influences on Sellars's thought. Kant was the door through which Lewis found philosophy, and later, the common root to which he reverted in his attempt to reconcile what seemed right to him about the apparently antithetical views of his teachers, William James and Josiah Royce. (Had he instead been trying to synthesize Royce with Dewey, instead of James, he would have fetched up at Hegel.) In his 1929 *Mind and the World Order*, Lewis introduced as a central technical conception the notion of the sensory "Given", which Sellars would famously use (characteristically, without mentioning Lewis by name) as the paradigm of what he in *EPM* called the "Myth of the Given." (Indeed, shortly after his 1946 *An Analysis of Knowledge and Valuation*, which Sellars also clearly has in mind in *EPM*,

---

<sup>2</sup> In his Introduction to my Harvard University Press edition of "Empiricism and the Philosophy of Mind."

<sup>3</sup> His only rival for this accolade, I think, would be Peter Strawson, who certainly did a lot to make us realize that a reappropriation of some of Kant's theoretical philosophy might be a viable contemporary

Lewis wrote a piece addressing the question “Is The Givenness of the Given Given?” His answer was No: It is a necessary postulate of high philosophical theory, which dictates that without a sensory Given, empirical knowledge would be impossible.)

In the book I argue that Sellars modeled his own Kantian “metalinguistic” treatments of modality and the ontological status of universals explicitly on ideas of Carnap. Although, like Lewis, Carnap is not explicitly mentioned in *EPM*, his presence is registered for the philosophical cognoscenti Sellars took himself to be addressing there by the use of the Carnapian term “protocol sentence” (as well as Schlick’s “Konstatierung”) for noninferential observations. Unlike Lewis, Carnap actually stood in the line of inheritance of classical nineteenth-century German neo-Kantianism. His teacher, Bruno Bauch, was (like Heidegger), a student of Heinrich Rickert in Freiburg—who, with the older Wilhelm Windelband, led the Southwest or Baden neo-Kantian school. In spite of these antecedents, Bauch was in many ways closer to the Marburg neo-Kantians, Hermann Cohen and Paul Natorp, in reading Kant as first and foremost a philosopher of the natural sciences, mathematics, and logic. I suppose that if one had asked Carnap in what way his own work could be seen as a continuation of the neo-Kantian tradition of his teacher, he would first have identified with this Marburg neo-Kantian understanding of Kant, and then pointed to the *logical* element of his logical empiricism—itsself a development of the pathbreaking work of Frege, Bauch’s friend and colleague at Jena when Carnap studied with both there—as giving a precise and modern

---

project. But I do not think of Peter Strawson’s work as *systematically* neo-Kantian in the way I want to argue that Sellars’s is.

form to the conceptual element in empirical knowledge, which deserved to be seen as a worthy successor to Kant's own version of the conceptual.

If Lewis and Carnap do not immediately spring to mind as neo-Kantians, that is because each of them gave Kant an empiricist twist, which Sellars was concerned to undo. If you thought that Kant thought that the classical empiricists' Cartesian understanding of the sensory contribution to knowledge was pretty much all right, and just needed to be supplemented by an account of the independent contribution made by a conceptual element, you might well respond to the development of the new twentieth century logic with a version of Kant that looks like Lewis's *Mind and the World Order*, and *An Analysis of Knowledge and Valuation*, and Carnap's *Aufbau* (and for that matter, Nelson Goodman's *Structure of Appearance*). That assumption about Kant's understanding of the role played by sense experience in empirical knowledge is exactly what Sellars challenges in *EPM*.

One of the consequences of his doing that is to make visible the neo-Kantian strand in analytic philosophy that Lewis and Carnap each, in his own way, represented—and which Sellars and, in our own time, John McDowell further developed. Quine was a student of both Lewis and Carnap, and the Kantian element of the common empiricism he found congenial in their thought for him drops out entirely—even though the logic remains. His Lewis and his Carnap are much more congenial to a narrative of the history of analytic philosophy initiated by Bertrand Russell and G.E. Moore, according to which the movement is given its characteristic defining shape as a recoil from Hegel (seen

through the lenses of the British Idealism of the waning years of the nineteenth century). They understood enough about the Kantian basis of Hegel's thought to know that a *holus bolus* rejection of Hegel required a diagnosis of the idealist rot as having set in already with Kant. This narrative does pick out one current in the analytic river—indeed, the one that makes necessary the reappropriation of the metaconceptual resources of Kant's theoretical philosophy in the late twentieth and early twenty-first centuries. But it was never the whole story.<sup>4</sup> The neo-Kantian tradition comprising Lewis, Carnap, and Sellars can be thought of as an undercurrent, somewhat occluded from view by the empiricist surface.

## 2. Categories in Kant

Many Kantian themes run through Sellars's philosophy. My book is oriented around two master-ideas, each of which orients and ties together a number of otherwise apparently disparate aspects of his work. One is a strand of scientific naturalism, which I reject, on behalf of Sellars own better wisdom--so I claim. The one I'll focus on here is the good idea that besides concepts whose characteristic expressive job it is to describe and explain empirical goings-on, **there are concepts whose characteristic expressive job it is to make explicit necessary structural features of the discursive framework within which alone description and explanation are possible.** Failing to acknowledge and appreciate this crucial difference between the expressive roles different bits of vocabulary play is a perennial source of distinctively philosophical misunderstanding. In

---

<sup>4</sup> Paul Redding begins the process of recovering the necessary counter-narrative in the Introduction to his *Analytic Philosophy and the Return of Hegelian Thought* [Cambridge University Press, 2010].

particular, Sellars thinks, attempting to understand concepts doing the second, framework-explicating sort of work on the model of those whose proper use is in empirical description and explanation is a fount of metaphysical and semantic confusion.<sup>5</sup> Among the vocabularies that play the second sort of role, Sellars includes *modal* vocabulary (not only the alethic, but also the deontic species), *semantic* vocabulary, *intentional* vocabulary, and *ontological-categorial* vocabulary (such as ‘proposition’, ‘property’ or ‘universal’, and ‘object’ or ‘particular’). It is a mistake, he thinks, to understand the use of any of these sorts of vocabulary as fact-stating in the narrow sense that assimilates it to *describing* how the world is. It is a corresponding mistake to recoil from the metaphysical peculiarity and extravagance of the kinds of facts one must postulate in order to understand statements couched in these vocabularies as fact-stating in the narrow sense (e.g. normative facts, semantic facts, conditional facts, facts about abstract universals) by denying that such statements are legitimate, or even that they can be true. (Though to say that they are true is not, for Sellars, to describe them.) Both mistakes (the dogmatic metaphysical and the skeptical), though opposed to one another, stem from the common root of the *descriptivist fallacy*. That is the failure to see that some perfectly legitimate concepts do not play a narrowly descriptive role, but rather a different, explicative one with respect to the practices of description and explanation. Following Carnap, Sellars instead analyses the use of all these kinds of vocabulary as, each in its own distinctive way, “covertly metalinguistic.”

---

<sup>5</sup> Distinguishing two broadly different kinds of *use* bits of vocabulary can play does not entail that there are two corresponding kinds of *concepts*—even in the presence of the auxiliary Sellarsian hypothesis that grasp of a concept is mastery of the use of a word. Though I suppress the distinction between these two moves in these introductory formulations, it will become important later in the story.

In opposing a Procrustean descriptivism about the expressive roles locutions can play, Sellars makes common cause with the later Wittgenstein. For Wittgenstein, too, devotes a good deal of effort and attention to warning us of the dangers of being in thrall to (“bewitched by”) a descriptivist picture. We must not simply assume that the job of all declarative sentences is to state facts (“I am in pain,” “It is a fact that ...”), that the job of all singular terms is to pick out objects (“I think...,” “I have a pain in my foot,”), and so on. In addition to tools for attaching, detaching, and in general re-shaping material objects (hammer and nails, saws, draw-knives...) the carpenter’s tools also include plans, a foot-rule, level, pencil, and toolbelt. So, too, with discursive expressive ‘tools’.<sup>6</sup> Wittgenstein’s expressive pluralism (language as a motley) certainly involves endorsement of the anti-descriptivism Sellars epitomizes by saying

[O]nce the tautology ‘The world is described by descriptive concepts’ is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an *ungrudging* recognition that many expressions which empiricists have relegated to second-class citizenship in discourse are not *inferior*, just *different*.<sup>6</sup>

But Sellars differs from Wittgenstein in characterizing at least a broad class of nondescriptive vocabularies as playing generically the *same* expressive role. They are broadly *metalinguistic* locutions expressing necessary features of the framework of discursive practices that make description (and—so—explanation) possible. Of this broad binary distinction of expressive roles, with ordinary empirical descriptive

---

<sup>6</sup> CDCM §79.

vocabulary on one side and a whole range of apparently disparate vocabularies going into another class as “metalinguistic”, there is, I think, no trace in Wittgenstein.<sup>7</sup>

The division of expressive roles that I am claiming for Sellars binds together modal, semantic, intentional, and ontological-categorial vocabulary in opposition to empirical-descriptive vocabularies traces back to Kant’s idea of “pure concepts of the understanding,” or categories, which play quite a different expressive role from that of ordinary empirical descriptive concepts. The expressive role of pure concepts is, roughly, to make explicit what is implicit in the use of ground-level concepts: the conditions under which alone it is possible to apply them, which is to say, use them to make judgments. Though very differently conceived, Kant’s distinction is in turn rooted in the epistemological difference Hume notices and elaborates between ordinary empirical descriptive concepts and concepts expressing lawful causal-explanatory connections between them. Hume, of course, drew skeptical conclusions from the observation that claims formulated in terms of the latter sort of concept could not be justified by the same sort of means used to justify claims formulated in terms of empirical descriptive concepts.

Kant, however, looks at Newton’s formulation of the best empirical understanding of his day and sees that the newly introduced concepts of force and mass are not intelligible apart from the *laws* that relate them. If we give up the claim that  $F$  equals  $m \cdot a$  then we do not mean force and mass, but are using some at least slightly different

---

<sup>7</sup> The best candidate might be the discussion of “hinge propositions” in *On Certainty*. But the point there is, I think, different. In any case, Wittgenstein does not *generalize* the particular expressive role he is

concepts. (Galileo's geometrical version of the (late medieval) observable concept of acceleration *is* antecedently intelligible). This leads Kant to two of his deepest and most characteristic metaconceptual innovations: thinking of statements of laws formulated using alethic modal concepts as making explicit rules for reasoning with ordinary empirical descriptive concepts, and understanding the contents of such concepts as articulated by those rules of reasoning with them.

This line of thought starts by revealing the semantic presuppositions of Hume's epistemological arguments. For Hume assumes that the contents of ordinary empirical descriptive concepts are intelligible antecedently to and independently of taking them to stand to one another in rule-governed inferential relations of the sort made explicit by modal concepts. Rejecting that semantic atomism then emerges as a way of denying the intelligibility of the predicament Hume professes to find himself in: understanding ordinary empirical descriptive concepts perfectly well, but getting no grip thereby on the laws expressed by subjunctively robust rules relating them. Even though Kant took it that Hume's skeptical epistemological argument rested on a semantic *mistake*, from his point of view Hume's investigation had uncovered a crucial *semantic* difference between the expressive roles of different kinds of concepts. Once his attention had been directed to them, he set himself the task of explaining what was special about these *nondescriptive* concepts.

Two features of Kant's account of the expressive role distinctive of the special class of concepts to which Hume had directed his attention are of particular importance

---

considering to anything like the extent I am claiming Sellars does.

for the story I am telling here. They are *categorial* concepts, and they are *pure* concepts. To say that they are ‘categorial’ in this context means that they make explicit aspects of the *form* of the conceptual as such. For Kant concepts are functions of judgment, that is, they are to be understood in terms of their role in judging. Categorial concepts express structural features of empirical descriptive judgments. What they make explicit is implicit in the capacity to make any judgments at all. This is what I meant when I said above that rather than describing how the world is, **the expressive job of these concepts is to make explicit necessary features of the framework of discursive practices within which it is possible to describe how the world is.** The paradigm here is the alethic modal concepts that articulate the subjunctively robust consequential relations among descriptive concepts.<sup>8</sup> It is those relations that make possible *explanations* of why one description applies because another does. That force *necessarily* equals the product of mass and acceleration means that one can *explain* the specific acceleration of a given mass by describing the force that was applied to it. (Of course, Kant also thinks that in articulating the structure of the judgeable as such, these concepts *thereby* articulate the structure of what is empirically *real*: the structure of *nature*, of the *objective world*. But this core thesis of his understanding of empirical realism within transcendental idealism is an optional additional claim, not entailed by the identification of a distinctive class of concepts as categories of the understanding.)

To say that these concepts are ‘pure’ is to say that they are available to concept-users (judgers = those who can understand, since for Kant the understanding is the faculty

---

<sup>8</sup> Note that these concepts are *not* those Kant discusses under the heading of “Modality”, but rather concern the hypothetical form of judgment.

of judgment) *a priori*.<sup>9</sup> Since what they express is implicit in any and every use of concepts to make empirical judgments, there is no *particular* such concept one must have or judgment one must make in order to be able to deploy the pure concepts of the understanding. To say that judges can grasp these pure concepts *a priori* is *not* to say that they are *immediate* in the Cartesian sense of nonrepresentational. Precisely not. The sort of self-consciousness (awareness of structural features of the discursive as such) they make possible is mediated by those pure concepts. What was right about the Cartesian idea of the immediacy of self-consciousness is rather that these mediating concepts are available to every thinker *a priori*. Their grasp does not require grasp or deployment of any *particular* ground-level empirical concepts, but is *implicit* in the grasp or deployment of *any* such concepts. The way I will eventually recommend that we think about this distinctive *a prioricity* is that in being able to deploy ordinary empirical descriptive concepts one already knows how to do everything one needs to know how to do in order to be able to deploy the concepts that play the expressive role characteristic of concepts Kant picks out as “categorical” (as well as some that he does not).

### 3. Categories in Sellars

Sellars’s development of Kant’s idea of pure concepts of the understanding is articulated by two master ideas. First, his successor metaconception comprises concepts that are in some broad sense *metalinguistic*.<sup>10</sup> In pursuing this line he follows Carnap,

---

<sup>9</sup> I take it that Kant always uses “*a priori*” and “*a posteriori*” as adverbs, modifying some some verb of cognition, paradigmatically “know”.

<sup>10</sup> In Chapter Three I discuss the sense in which “metalinguistic” should be understood in such formulations..

who besides ground-level empirical descriptive vocabulary allowed metalinguistic vocabulary as also legitimate in formal languages regimented to be perspicuous. Such metalinguistic vocabulary allows the formulation of explicit rules governing the use of descriptive locutions. Ontologically classifying terms such as ‘object’, ‘property’, and ‘proposition’ are “quasi-syntactical” metavocabulary corresponding to overtly syntactical expressions in a proper metalanguage such as ‘singular term’, ‘predicate’, and ‘declarative sentence’. They are used to formulate “L-rules”, which specify the structure of the language in which empirical descriptions are to be expressed.<sup>11</sup> Alethic modal vocabulary is used to formulate “P-rules”, which specify rules for reasoning with particular empirically contentful descriptive vocabulary. Carnap’s neo-Kantianism does not extend to embracing the metaconcept of categories, which he identifies with the excesses of transcendental idealism. But in the expressions Carnap classifies as overtly or covertly metalinguistic, Sellars sees the raw materials for a more thoroughly Kantian successor conception to the idea of pure categories of the understanding.

The second strand guiding Sellars’s reconceptualization of Kantian categories is his *semantic inferentialist* approach to understanding the contents of descriptive concepts. Sellars picks up on Kant’s rejection of the semantic atomism characteristic of both the British empiricism of Locke and Hume that Kant was reacting to and of the logical empiricism of Carnap that Sellars was reacting to.<sup>12</sup> The way he works out the anti-

---

<sup>11</sup> Chapter Seven discusses Sellars’s view about this kind of locution.

<sup>12</sup> “Another feature of the empiricist tradition is its ‘logical atomism,’ according to which every basic piece of empirical knowledge is logically independent of every other. Notice that this independence concerns not only *what* is known, but the *knowing* of it. The second dimension of this ‘atomism’ is of particular importance for understanding Kant’s rejection of empiricism...”[“Towards a Theory of the Categories” §16]

atomist lesson he learns from Kant is in terms of the essential contribution made to the contents of ordinary empirical descriptive concepts by the inferential connections among them appealed to in *explanations* of why some descriptions apply to something in terms of other descriptions that apply to it.

Although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*. It is only because the expressions in terms of which we describe objects, even such basic expressions as words for perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label. The descriptive and explanatory resources of language advance hand in hand.<sup>13</sup>

This is a rich and suggestive passage. It is worth unpacking the claims it contains. It is framed by a distinction between a weaker notion, *labeling*, and a *stronger* one, *describing*. By ‘labeling’ Sellars means discriminating, in the sense of responding differentially. A linguistic expression is used as a label if its *whole* use is specified by the *circumstances* under which it is applied—the *antecedents* of its application. We might distinguish between three kinds of labels, depending on how we think of these circumstances or antecedents. First, one could look at what stimuli as a matter of fact elicit or in fact have elicited the response that is being understood as the application of a label. Second, one could look *dispositionally*, at what stimuli *would* elicit the application of the label. Third, one could look at the circumstances in which the label is *appropriately* applied. What the three senses have in common is that they look only *upstream*, to the situations that have, would, or should prompt the use of the label. The

first provides no constraint on future applications of the label—*que sera sera*—as familiar gerrymandering arguments about “going on in the same way” remind us. The second doesn’t fund a notion of mistaken application. However one is disposed to apply the label is proper, as arguments summarized under the heading of “disjunctivitis” make clear. Only the third, normatively richer sense in which the semantics of a label consists in its circumstances of *appropriate* application (however the proprieties involved are understood) makes intelligible a notion of *mislabeling*.

Sellars wants to distinguish labeling in *all* of these senses from *describing*. The idea is that since labeling of any of these sorts looks only to the *circumstances* in which the label is, would be, or should be applied, expressions used with the semantics characteristic of labels address at most one of the two fundamental aspects of the use characteristic of descriptions. The rules for the use of labels tell us something about what is (or would be or should be) in effect so described, but say nothing at all about what it is described *as*. That, Sellars thinks, depends on the *consequences* of applying one description rather than another. The semantics of genuine descriptions must look downstream, as well as upstream. It is this additional feature of their use that distinguishes descriptions from labels. (Here one might quibble verbally with Sellars’s using ‘label’ and ‘description’ to describe expressions whose semantics depends on only one or on both of these dimensions of use. But it seems clear that a real semantic distinction is being marked.)

---

<sup>13</sup> CDCM §108.

Making a further move, Sellars understands those consequences of application of descriptions as essentially involving *inferential* connections to other descriptive concepts. This is what he means by saying that what distinguishes descriptions from labels is their situation in a “space of implications.” Paralleling the discussion of circumstances of application, we can think of these implications (consequences of application) as specifying what other descriptions do, would, or should *follow from* the application of the initial, perhaps responsively elicited, description. As he is thinking of things, a description (correctly) applies to a range of things (for descriptive concepts used observationally, including those that are appropriately noninferentially differentially responded to by applying the concept), which are described *by* it. And it describes them *as* something from which a further set of descriptions (correctly) follows. Crucially, these further descriptions can themselves involve applications of descriptive concepts that also have *non-inferential* (observational) circumstances of application. Descriptive concepts that have *only* inferential circumstances of application he calls ‘theoretical’ concepts.

In the opening sentence of the passage Sellars includes *understanding* as one of the phenomena he takes to be intricately with *description* in the way *explaining* is. Understanding a descriptive concept requires being able to place it in the “space of implications,” partly in virtue of which it has the content that it does. This is in general a kind of knowing *how* rather than a kind of knowing *that*: being able to distinguish in practice the circumstances and consequences of application of the concept, when it is appropriately applied and what follows from so applying it. Grasping a concept in this

sense is not an all-or-none thing. The ornithologist knows her way around inferentially in the vicinity of terms such as ‘icterid’ and ‘passerine’ much better than I do. A consequence of this way of understanding understanding is that one cannot grasp one concept without grasping many. This is Sellars’s way of developing Kant’s anti-atomist semantic insight.

Taking a further step (undertaking a commitment not yet obviously entailed by the ones attributed so far), Sellars also thinks that the inferences articulating the consequences of concepts used descriptively must always include *subjunctively robust* inferences. That is, the inferences making up the “space of implications” in virtue of which descriptive concepts have not only potentially atomistic circumstances of application but also non-atomistic relational consequences of application must extend to what other descriptions *would be* applicable if a given set of descriptions *were* applicable. For what Sellars means by ‘explanation’ is understanding the applicability of some descriptions as *explained by* the applicability of others according to just this kind of inference. This is, of course, just the sort of inferential connection that Hume’s empiricist atomistic semantics for descriptive concepts, construing them as *labels*, could *not* underwrite. Sellars’s conception of descriptions, as distinguished from labels, is his way of following out what he sees as Kant’s anti-atomist semantic insight. *Modal* concepts make explicit these *necessary* inferential-consequential connections between descriptive concepts. They thereby perform the expressive role characteristic of Kantian categories: expressing essential features of the framework within which alone genuine description is possible.

All of this is meant to explicate what Sellars means by saying that “the descriptive and explanatory resources of language advance hand in hand.” In addition to Kant’s idea, Sellars here takes over Carnap’s idea of understanding concepts whose paradigm is modal concepts as (in some sense) *metalinguistic*. The principal class of genuinely intelligible, nondefective nondescriptive vocabulary Carnap allows in *The Logical Syntax of Language* is syntactic metavocabulary and what he there calls “quasi-syntactic” vocabulary, which is covertly metalinguistic. For Sellars, the *rules* which modal vocabulary expresses are rules for deploying linguistic locutions. Their “rulishness” is their subjunctive robustness. Following out this line of thought, Sellars takes it that “grasp of a concept is mastery of the use of a word.” He then understands the metalinguistic features in question in terms of rules of *inference*, whose paradigms are Carnap’s L-rules and P-rules. His generic term for the inferences that articulate the contents of ordinary empirical descriptive concepts is “material inferences.” The term is chosen to contrast with inferences that are ‘formal’ in the sense of depending on *logical* form. In another early essay he lays out the options he considers like this:

...we have been led to distinguish the following six conceptions of the status of material rules of inference:

- (1) Material rules are as essential to meaning (and hence to language and thought) as formal rules, contributing to the architectural detail of its structure within the flying buttresses of logical form.

(2) While not essential to meaning, material rules of inference have an original authority not derived from formal rules, and play an indispensable role in our thinking on matters of fact.

(3) Same as (2) save that the acknowledgment of material rules of inference is held to be a dispensable feature of thought, at best a matter of convenience.

(4) Material rules of inference have a purely derivative authority, though they are genuinely rules of inference.

(5) The sentences which raise these puzzles about material rules of inference are merely abridged formulations of logically valid inferences. (Clearly the distinction between an inference and the formulation of an inference would have to be explored).

(6) Trains of thought which are said to be governed by "material rules of inference" are actually not inferences at all, but rather activated associations which mimic inference, concealing their intellectual nudity with stolen "therefores".<sup>14</sup>

His own position is that an expression has conceptual content conferred on it by being caught up in, playing a certain role in, material inferences:

...it is the first (or "rationalistic") alternative to which we are committed. According to it, material transformation rules determine the descriptive meaning of the expressions of a language within the framework provided by its logical transformation rules... In traditional language, the "content"

---

<sup>14</sup> Sellars, "Inference and Meaning" *PPPW* pp. 265/317, reprinted in *In the Space of Reasons*.

of concepts as well as their logical "form" is determined by the rules of the Understanding.<sup>15</sup>

By “traditional language” here, he means Kantian language. The talk of “transformation rules” is, of course, Carnapian. In fact in this essay Sellars identifies his “material rules of inference” with Carnap’s “P-rules.” (‘Determine’ is--here, as generally--crucially ambiguous between ‘constrain’ and ‘settle’—the difference corresponding to that between what I have elsewhere called ‘weak’ and ‘strong’ semantic inferentialism.)

As already indicated, the material inferential rules that in one or another of these senses “determine the descriptive meaning of expressions” are for Sellars just the subjunctively robust, hence explanation-supporting ones. As he puts the point in the title of a long essay, he construes “Concepts as Involving Laws, and Inconceivable without Them.” This is his response to Quine’s implicit challenge in “Two Dogmas of Empiricism” to say what feature of their *use* distinguishes inferences determining conceptual contents from those that simply register matters of fact. Since empirical inquiry is generally required to determine what laws govern concepts such as copper, temperature, and mass, Sellars accepts the consequence that inquiry plays the role not only of determining facts but also of improving our conceptions—of teaching us more about the concepts that articulate those facts by teaching us more about what really follows from what--in a subjunctively robust, counterfactual-supporting sense of "follows from."

On this way of understanding conceptual content, the *modal* concepts that express the lawfulness of connections among concepts and so underwrite subjunctively robust

---

<sup>15</sup> Sellars, "Inference and Meaning" *PPPW* pp. 284/336.

implications—concepts such as law, necessity, and what is expressed by the use of the subjunctive mood—have a different status from those of ordinary empirical descriptive concepts. Rather than in the first instance describing how the world is, they make explicit features of the framework that makes such description possible. Because they play this distinctive framework-explicating role, what they express must be implicitly understood by anyone who can deploy *any* ground-level descriptive concepts. As I would like to put the point, in knowing how to (being able to) use any ordinary empirical descriptive vocabulary, each interlocutor thereby already knows how to do everything she needs to know how to do, in order to be able to deploy the modal locutions that register the subjunctive robustness of the inferences that in turn determine the content of the descriptive concepts that vocabulary expresses. This is what Kant’s idea that the pure concepts of the understanding are knowable *a priori* becomes when transposed into Sellars’s framework.

The two lines of thought that orient Sellars’s treatment of alethic modality, namely semantic inferentialism and a metalinguistic understanding of the expressive role characteristic of modal locutions, are epitomized in an early formulation:

I shall be interpreting our judgments to the effect that A causally necessitates B as the expression of a rule governing our use of the terms 'A' and 'B',<sup>16</sup>

where the rule in question is understood as a rule licensing subjunctively robust inferences. I have been filling in the claim that this overall approach to modality deserves to count as a development of Kant’s notion of categories, pure concepts of the understanding, as concepts that make explicit features of the discursive framework that

makes empirical description possible. Sellars himself, however, does not discuss this aspect of his work under that heading. When he talks about categories he turns instead to his nominalism about abstract entities. The central text here is “Towards a Theory of the Categories” of 1970.<sup>17</sup> The story he tells there begins with Aristotle’s notion of categories (though he waves his hands wistfully at a discussion of its origins in Plato’s *Sophist* that he feels cannot shoehorn into the paper) as ontological *summa genera*. There he opposes an unobjectionable hierarchy

Fido is a dachshund.

Fido is a dog.

Fido is a brute.

Fido is an animal.

Fido is a corporeal substance.

Fido is a substance.

To a potentially problematic one

X is a red.

X is a color.

X is a perceptual quality.

X is a quality.<sup>18</sup>

The next decisive move in understanding the latter hierarchy he attributes to Ockham, whom he reads as transposing the discussion into a metalinguistic key. Ockham’s strategy, he tells us, is to understand

---

<sup>16</sup> Sellars, "Language, Rules, and Behavior" footnote 2 to p. 136/296 in *PPPW*.

<sup>17</sup> In *Experience and Theory*, edited by L. Foster and J.W. Swanson (University of Massachusetts Press, 1970), pp. 55-78. Reprinted in *Essays in Philosophy and its History* (D. Reidel, 1974).

<sup>18</sup> “Towards a Theory of the Categories” (TTC) §10-11.

(A) Man is a species.

as

(B) ·Man· is a sortal mental term.<sup>19</sup>

while construing mental items as “analogous to linguistic expressions in overt speech.”

This sketch sets up the transition to what Sellars makes of Kant’s understanding of categories:

What all this amounts to is that to apply Ockham’s strategy to the theory of categories is to construe categories as classifications of conceptual items. This becomes, in Kant’s hands, the idea that categories are the most generic functional classifications of the elements of judgments.<sup>20</sup>

At the end of this development from Aristotle through Ockham to Kant, he concludes

[I]nstead of being *summa genera* of entities which are objects ‘in the world,’ ...categories are *summa genera* of conceptual items.<sup>21</sup>

The account he goes on to expound in this essay, as well as in his other expositions of his nominalism about terms for qualities or properties, construes such terms metalinguistically, as referring to the inferential roles of the base-level concepts as used in empirical descriptions. I explain how I understand the view and the arguments on this topic in Chapter Seven of *From Empiricism to Expressivism*: “Sellars’s Metalinguistic Expressive Nominalism.” Without going further into that intricate view here, the point I want to make is that although Sellars does not say so, the metaconceptual role he here explicitly puts forward as a successor-concept to Kant’s notion of category is generically

---

<sup>19</sup> TTC §16.

<sup>20</sup> TTC §22.

the same as that I have argued he takes alethic modal locutions to play. It is this capacious conception I want to build upon and develop further.

#### 4. Categories Today

The general conception of pure categorial concepts that I have been attributing to Sellars, based on the commonalities visible in his treatment of alethic modal vocabulary and of abstract ontological vocabulary, develops Kant's idea by treating some vocabularies (and the concepts they express) as "covertly metalinguistic." This Sellarsian conception represents his development of Carnap's classification of some expressions as "quasi-syntactic." The underlying insight is that some important kinds of vocabularies that are not strictly or evidently metalinguistic are used not (only) to describe things, but in ways that (also) depend on the use of *other* vocabularies—paradigmatically, empirical descriptive ones.

The lessons I draw from the strengths and weaknesses of Sellars's successor-conception of the "pure concepts of the Understanding" are four-fold. That is, I think he is pointing towards an expressive role characteristic of some concepts, and the vocabularies expressing them, that has four distinctive features. *First*, these concepts express what I will call "pragmatically mediated semantic relations" between vocabularies. *Second*, these concepts play the expressive role of *making explicit* essential features of the use of some other vocabulary. *Third*, the proper use of these concepts can be systematically *elaborated from* the use of that other vocabulary. *Fourth*, the features

---

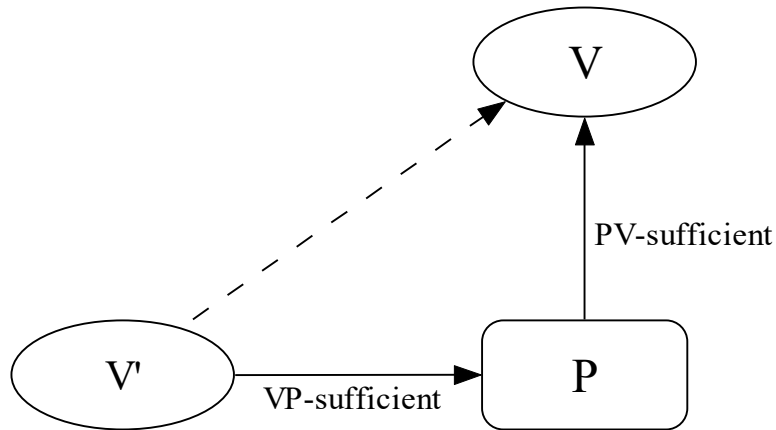
<sup>21</sup> TTC §23.

of vocabulary(concept)-use they explicate are *universal*: they are features of any and every autonomous discursive practice. I think there are concepts that play this distinctive four-fold expressive role, and that a good thing to mean today by the term “category” is metaconcepts that do so.

Carnap and Tarski introduced the expression “metalanguage” for languages that let one talk about languages, with the paradigmatic examples being syntactic and semantic metalanguages. In his earliest writings, Sellars also talks about “pragmatic metalanguages,” meaning languages for talking about the *use* of expressions—rather than the syntactic or semantic properties of expressions. These were to be the languages in which we conduct what he called “pure pragmatics.” During and after Sellars’s most important work in the the *anni mirabiles* of 1954-63, however (possibly influenced by Carnap), he shifts to using the expression “semantics” to cover the essentially the same ground. I think that this was a step backward, and that it is one of the obstacles that prevented him from getting clear about the sense in which he wanted to claim that such locutions as alethic modal vocabulary and singular terms purporting to refer to universals (“circularity”) and their kinds (“property”) are “covertly metalinguistic.” One vocabulary serving as a *pragmatic metavocabulary* for another is the most basic kind of pragmatically mediated semantic relation between vocabularies. It deserves to be called such because the *semantics* of the pragmatic metavocabulary depends on the *use* of the vocabulary for which it is a pragmatic metavocabulary. The relation itself is aptly called a “semantic” relation in the special case where one vocabulary is sufficient to specify

practices or abilities whose exercise is sufficient to confer on another vocabulary the meanings that it expresses.

We could represent such a semantic relation, mediated by the practices of using the second vocabulary that the first vocabulary specifies, like this:<sup>22</sup>



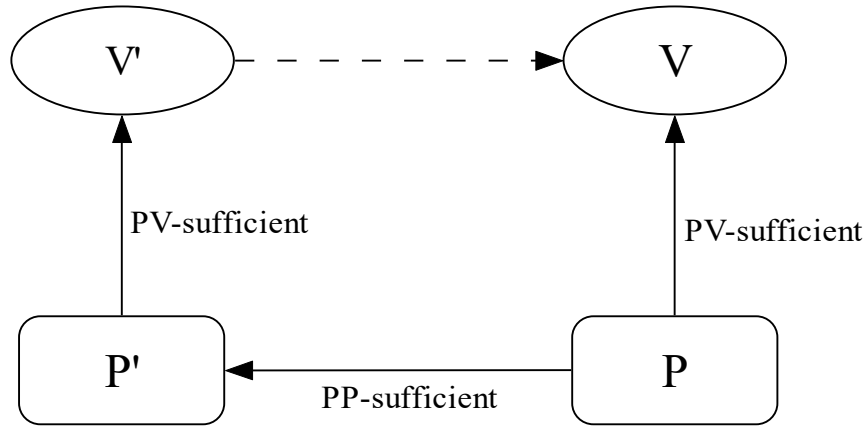
The pragmatically mediated semantic relation between vocabularies V' and V, indicated by the dashed arrow, obtains when vocabulary V' is expressively sufficient to *specify* practices-or-abilities P (that semantic fact about V' with respect to P is here called “VP-sufficiency”) that are sufficient to *deploy* the vocabulary V with the meanings that it expresses when so used. In asserting that this relation between vocabularies obtains, one is claiming that if all the sentences in V' used to specify the practices-or-abilities P are true of P, then anyone engaging in those practices or exercising those abilities as specified in V' is using the expressions of V with their proper meanings. This semantic relation between what is expressible in the two vocabularies is mediated by the practices P that the first specifies and which are the use of the second. This particular pragmatically mediated semantic relation holds when the vocabulary V' allows one to *say*

what one must *do* in order to *say* what can be said in the vocabulary V. In that sense V' makes *explicit* (sayable, claimable) the practices-or-abilities *implicit* in using V. This is the explicative relation I mention as the second component of the complex expressive role that I am offering as a candidate for a contemporary successor-(meta)concept to Kant's (meta)concept of category. There are other pragmatically mediated semantic relations besides being a pragmatic metavocabulary in this sense, and others are involved in the categorial expressive role. The result will still fall under the general rubric that is the first condition: being a pragmatically mediated semantic relation.

One such further pragmatically mediated semantic relations between vocabularies holds when the practices PV-sufficient for deploying one vocabulary, though not themselves PV-sufficient for deploying a second one, can be systematically elaborated into such practices. That is, in being able to deploy the first vocabulary, one already knows how to do everything one needs to know how to do, in principle, to deploy the second. But those abilities must be suitably recruited and recombined. The paradigm here is *algorithmic* elaboration of one set of abilities into another. Thus, in the sense I am after, the capacities to do multiplication and subtraction are algorithmically elaborable into the capacity to do long division. *All* you need to learn how to do is to put together what you already know how to do in the right way—a way that can be specified by an algorithm. The diagram for this sort of pragmatically mediated semantic relation between vocabularies is:

---

<sup>22</sup> I introduce, develop, and apply these “meaning-use diagrams” in *Between Saying and Doing: Towards an Analytic Pragmatism* [Oxford University Press, 2008].



The dotted arrow indicates the semantic relation between vocabularies V' and V. It is the relation that holds when all the relations indicated by solid arrows hold—that is, when the practices-or-abilities sufficient to deploy vocabulary V can be elaborated into practices sufficient to deploy vocabulary V'. In this case, the semantic relation in question is mediated by two sets of practices-or-abilities: those sufficient to deploy the two vocabularies.

A concrete example of vocabularies standing in this pragmatically mediated semantic relation, I claim, is that of *conditionals* in relation to ordinary empirical descriptive (OED) vocabulary. For using such OED vocabulary, I claim (following Sellars following Kant), requires distinguishing in practice between materially good inferences involving descriptive predicates and ones that are not materially good. One need not be either infallible or omniscient in this regard, but unless one makes *some* such distinction, one cannot count as deploying the OED vocabulary in question. But in being able practically to distinguish (however fallibly and incompletely) between materially good and materially bad inferences, one knows how to do everything one needs to know how to do, in principle, to deploy conditionals. For conditionals can be introduced by

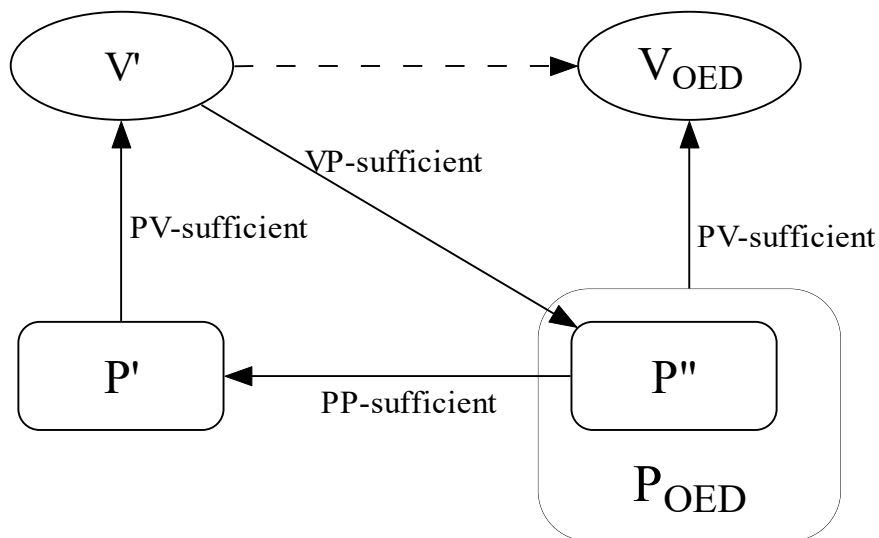
recruiting those abilities in connection with the use of sentences formed from the old vocabulary by using the new vocabulary. On the side of circumstances of application (assertibility conditions), one must acknowledge commitment to the conditional  $p \rightarrow q$  just in case one takes the inference from  $p$  to  $q$  to be a materially good one. And on the side of consequences of application, if one acknowledges commitment to the conditional  $p \rightarrow q$ , then one must take the inference from  $p$  to  $q$  to be a materially good one. These rules constitute an algorithm for elaborating the ability to distinguish materially good from materially bad inference using OED vocabulary (or any other vocabulary, for that matter) into the ability appropriately to use conditionals formed from that vocabulary: to distinguish when such conditionals are assertible, and what the consequences of their assertibility is.

My idea for a successor-concept to what Sellars (with hints from Carnap) made of Kant's metaconception of pure concepts of the Understanding is that they must play *both* of these expressive roles, stand in *both* sorts of pragmatically mediated semantic relations to another vocabulary. It must be possible to *elaborate* their use from the use of the index vocabulary, and they must *explicate* the use of that index vocabulary. Speaking more loosely, we can say that such concepts are both *elaborated from* and *explicative of* the use of other concepts—in short that they are el-ex, or just LX with respect to the index vocabulary.

The fourth condition I imposed above is that the concepts in question must be *universally LX*, by which I mean that they must be LX for every autonomous discursive

practice (ADP)—every language game one could play though one played no other. That is, the practices from which their use can be elaborated and of which their use is explicative must be essential to talking or thinking at all. This universality would distinguish categorial concepts, in the sense being specified, from metaconcepts that were elaborated from and explicative of only some parasitic fragment of discourse—culinary, nautical, or theological vocabulary, for instance. I take it that any autonomous discursive practice must include the use of ordinary empirical descriptive vocabulary. If so, being LX for OED vocabulary would suffice for being *universally* LX, LX for every ADP.

Putting all these conditions together yields the following diagram of the pragmatically mediated semantic relation between vocabularies that obtains when vocabulary  $V'$  plays the expressive role of being universally LX by being elaboratable from and explicative of practices necessary for the deployment of ordinary empirical descriptive vocabulary:



The fact that the rounded rectangle labeled  $P''$ , representing the practices from which vocabulary  $V'$  is elaborated and of which it is explicative, appears inside the rounded

rectangle representing practices sufficient to deploy ordinary empirical descriptive vocabulary indicates that the practices P" are a necessary part of the practices sufficient to deploy OED vocabulary, but need not comprise all such practices. Thus, distinguishing materially good from materially bad inferences involving them is necessary for deploying ordinary empirical descriptive vocabulary (rather than mere labels), but there is a lot more involved in doing so—using such vocabulary observationally, for instance. Different categorial metaconcepts can be LX for different essential features of the use of empirical descriptive vocabulary. Thus alethic modal vocabulary explicates the subjunctive robustness of the inferences explicated by conditionals. “Quasi-syntactic” abstract ontological vocabulary such as ‘property’ and ‘proposition’ explicate structural features of descriptive sentences.

Diagramming the expressive role of being LX for practices necessary to deploy OED vocabulary provides an analysis that breaks down the claim that some vocabulary plays a categorial role into its component sub-claims. To show that alethic modal vocabulary, for instance, stands in this pragmatically mediated semantic relation to ordinary empirical descriptive vocabulary one must show that there is some practices-or-abilities (in this case, to reason subjunctively or counterfactually) that are 1) a necessary component of practices-or-abilities that are 2) (PV)sufficient to deploy OED vocabulary, 3) from which one can elaborate practices-or-abilities that are 4) (PV)sufficient to deploy vocabulary (alethic modal vocabulary) 5) that is (VP)sufficient to explicate or specify the original practices-or-abilities. Although there is by design considerable elasticity in the concepts vocabulary, practices-or-abilities, and the various sufficiency and necessity

relations between them, the fine structure of the distinctive expressive role in question is clearly specified.

What credentials does that expressive role have to pick out a worthy successor metaconcept to what Sellars made of Kant's categories or pure concepts of the Understanding? At the beginning of my story I introduced the idea behind the Kantian categories as the idea that besides the concepts whose principal use is in giving empirical descriptions and explanations, there are concepts whose principal use is in making explicit features of the framework that makes empirical description and explanation possible. The expressive task characteristic of concepts of this latter class is to articulate what Kant called the "transcendental conditions of experience." The concepts expressed by vocabularies that are LX for empirical descriptive vocabulary perform this defining task of concepts that are categories. As explicative of practices necessary for deploying vocabularies performing the complex expressive task of description and explanation (distinguishable only in the context of their complementary relations within a pragmatic and semantic context that necessarily involves both), this kind of vocabulary makes it possible to *say* what practitioners must be able to *do* in order to describe and explain how things empirically are. They do this by providing a pragmatic metavocabulary for describing and explaining. This is a central feature (the 'X' in 'LX') of the complex pragmatically mediated semantic relation between categorial metaconcepts and ordinary empirical descriptive vocabulary.

One feature of the concepts performing this explicative function that Kant emphasizes is that they are “*pure* concepts of the Understanding.” (I take it that the “of” should be understood as expressing both the subjective and objective genitives—as in “Critique of Pure Reason.” These concepts both belong to the Understanding and address it, being both discursive and metaconceptual.) To say that they are pure concepts is to say that they are graspable *a priori*.<sup>23</sup> The feature of the LX model that corresponds to the a prioricity of Kant’s categories is that the use of LX metaconcepts can be elaborated from that of the empirical descriptive vocabularies for which they are LX. As I have put the point, in knowing how to deploy OED vocabulary, one already knows how to do everything one needs to know how to do to deploy vocabulary that is LX for it—such as alethic modal vocabulary, conditionals, and ontological classificatory vocabulary. If we take it, as per Sellars, that grasp of a concept is mastery of the use of a word, then one need not *actually* grasp concepts that are LX for descriptive vocabulary in order to deploy descriptive vocabulary. But in effect, *all* one is missing is the words for them. The circumstances and consequences of application of LX concepts can be formulated by rules that appeal only to abilities one already has in virtue of being able to use OED vocabulary. (Think of the sample rules for conditionals sketched above.) In that sense, the LX concepts are *implicit in* the descriptive concepts. It is not that one must or could grasp these concepts *before* deploying descriptive concepts. It is rather that nothing more is required to grasp them than is required to deploy descriptive concepts, and there are no particular descriptive concepts one must be able to deploy, nor any particular descriptive claims that one must endorse, in order to possess abilities sufficient to deploy the universally LX metaconcepts.

---

<sup>23</sup> Kant does admit also impure *a priori* principles.

The class of concepts that are arguably universally LX (LX for every autonomous discursive practice because LX for OED vocabulary) overlaps Kant's categories in important ways—most notably in the alethic modal concepts that make explicit subjunctively robust consequential relations among descriptive concepts. But the two do not simply coincide. In *Between Saying and Doing* I argue that besides modal vocabulary, logical vocabulary, indexical and demonstrative vocabulary, normative vocabulary, and semantic and intentional vocabulary all should be thought of as LX for OED vocabulary. In spite of this extensional divergence, the fact that vocabulary that is LX for descriptive vocabulary in general principle shares with Kant's categories the two crucial features of being explicative of such vocabulary and being graspable *a priori* makes the idea of universally LX metaconcepts a worthy successor to Kant's breakthrough idea. The fact that Sellars's own development of this idea of Kant's takes such important steps in this direction convinces me that his version of the categories was a progressive step, and a Good Idea.