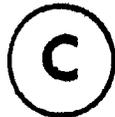


Sensa in Sellars' Theory
of Perception



Peter K. Dauphinee

Submitted to the Faculty of Graduate Studies and
Research in partial fulfillment of the requirements
for the degree of Master of Arts.

Department of Philosophy
McGill University

Montreal, Quebec
September, 1979

SENSA IN SELLARS' THEORY OF PERCEPTION

ABSTRACT

Wilfrid Sellars claims that a scientific account of human beings will find it necessary to postulate a new type of basic particular, which he calls *sensa*. It is not entirely clear what are the reasons for which Sellars makes this claim. This thesis seeks to answer this question, and subsequently to evaluate what it finds to be Sellars' argument for *sensa*. Three possible arguments are distinguished. The argument from the homogeneity of colours is taken to be Sellars' principal argument. A key premise in this argument is that colours, because of their ultimate homogeneity, cannot be reduced to complexes of scientific properties and relations. This premise involves a principle of reducibility which close scrutiny reveals to be more problematic than it initially appears. The conclusion is that the argument for *sensa* fails because it relies on an unjustified application of the principle.

Peter K. Dauphinee

Department of Philosophy
McGill University

Montreal, Quebec
September, 1979

SENSA IN SELLARS' THEORY OF PERCEPTION

RESUME

Wilfrid Sellars affirme qu'une explication scientifique de l'être humain va nécessiter le besoin de postuler une nouvelle sorte d'individu élémentaire, qu'il appelle "sensa." Ce n'est pas très clair pour quelles raisons Sellars fait cette déclaration. Cette thèse cherche à répondre à cette question et subséquemment à évaluer ce qu'il trouve à être les arguments de Sellars pour "sensa." Trois arguments possibles peuvent être discernés. L'argument principal de Sellars est l'homogénéité des couleurs. Un principal facteur de cet argument est que les couleurs, à cause de leur homogénéité ultime, ne peuvent être réduites à des complexes des propriétés et relations scientifiques. Ce facteur comprend un principe de réductibilité qui, après une étude attentive se dévoile être un plus grand problème qu'initialement prévu. La conclusion nous dit que l'argument pour "sensa" échoue parce qu'il se base trop sur une application injustifiée du principe.

Peter K. Dauphinee

Département de Philosophie
Université McGill

Montréal, Québec

Septembre, 1979

TABLE OF CONTENTS

Chapter 1	Introduction	1
Chapter 2	Meaning and Truth	35
Chapter 3	The Manifest Image	52
Chapter 4	Theoretical Explanation	74
Chapter 5	Conclusion	90
Bibliography		106

Chapter 1: Introduction

Wilfrid Sellars argues that in the final scientific account of the world, the logical space of the sensuous features of the manifest world such as colours will appear transposed as features of a type of basic particular called *sensa*. The argument is emergentist; *sensa*, although basic particulars, are constituents only of animate, perceiving beings--they do not appear in a scientific account which is necessary and sufficient for a description of inanimate objects. Of course, the claim that secondary qualities are really aspects of perceiving beings is not surprising; what is, however, is Sellars' claim that they are aspects of basic particulars--particulars which are not reducible to those particulars which are adequate for a description of the inanimate world (that is roughly, the basic particulars currently postulated by science).

Serious critiques of the argument have been published by J. W. Cornman and C. A. Hooker. Each seeks to disprove the argument and each fails, I think, largely because of certain difficulties involved in the rather peculiar way Sellars presents his argument. The argument appears to have many strands not all of which appear in any one of Sellars' works. Thus, drawing from various papers, Cornman and Hooker have separately attempted to reconstruct then criticize a single, coherent argument for *sensa* from Sellarsian resources. I do not intend

to examine in detail these construals of Sellars argument since I believe they tend to conceal rather than reveal whatever force the argument for *sensa* has. I propose instead a detailed examination and discussion of the relevant texts. Of particular interest are three passages which provide what may be considered three independent reasons for supposing *sensa* to be basic. These appear in "Philosophy and the Scientific Image of Man," "Phenomenalism," and "The Identity Approach to the Mind Body Problem." The problem is that no one of these arguments appears by itself to provide a sufficient argument for the basicness or primitiveness of *sensa*. Thus we must attempt to understand how they are related to each other and whether together they provide a convincing argument for *sensa*.

Before pressing on with this strategy, however, I wish to look briefly at a statement of the argument for *sensa* which is provided by Hooker. Sellars divides our multi-faceted ways of viewing the world into two broad conceptual frameworks--the manifest image and the scientific image. The manifest image is a view of the world which relies heavily on common sense. The entities and concepts of this image are either those encountered in everyday experience, or may otherwise be theoretical, although such theoretical entities and concepts are conceived in direct analogy with non-theoretical concepts. Thus, thoughts and sense impressions, though theoretically postulated even in the manifest image, are conceived as being directly analogous with spoken statements (thinking-out-loud) in the first case, and

physical objects in the second. The scientific image is the view of the world we will (or would) have in a complete, unified science. The ultimate objects of the scientific image, theoretically postulated, have a conceptual independence from common sense concepts. The scientific image is incompatible with the manifest image which it will ultimately replace as the true view of the world because of its greater explanatory power and, perhaps, coherence. The scientific image, however, must, or will, contain successor concepts to those of the manifest image which will capture the "logical spaces" of the succeeded concepts.

Among the "logical spaces" of the manifest image which must be captured in the scientific image are those of the sensuous qualities of the world conceived in the manifest image along naive realist lines, as purely occurrent properties inhering in physical objects. In the manifest image sense impressions are construed adverbially as states of perceivers. Sense impressions have properties which are analogous to those of the physical objects they represent. The successors of sense impressions in the scientific image are again states of persons, construed now, however, as neurophysiological systems. As part of these systems, science will find it necessary to postulate *sensa*, a type of particular not found among those necessary and sufficient for a scientific description of non-living matter, which will have properties which capture the logical space of sensuous properties. These new particulars will be required because none

4

of the entities currently postulated by science have, or could have properties of the right type. In particular, none of the entities of current theory have properties which capture the "ultimate homogeneity" of our manifest concept of colour, and therefore, "coloured" sense impressions. In the final scientific account of human behaviour, *sensa* will be the theoretical entities to which successor concepts of manifest colour predicates which preserve logical features such as "ultimate homogeneity" will be applicable.

Although the manifest image is prior in the order of knowing, the scientific image is prior in the order of being. Thus, scientific objects are what "really exist;" manifest objects, on the other hand, don't "really exist." Sellars is not a materialist in the strong sense in which the only predicates necessary to describe the world are those found in a theory minimally adequate to describe inanimate matter (physical₂), but in a weaker sense in that these predicates are integrated into a single extensional space-time-causal scheme (physical₁).

Note: this account of Sellars is essentially that of Hooker.¹⁷

The argument involves the notion of two alternate conceptual frameworks--the one we have, the manifest or common sense framework, and the developing scientific framework which will, at some ideal time, replace it as the true view of the world. Sellars holds two theses concerning these frameworks, which play a crucial role in his argument: one concerns a "successor" relation which holds between those concepts defined in the common

sense framework and those defined in the scientific framework; the other is a principle of reducibility relating the attributes of wholes to the attributes of their parts which applies within each framework separately.

Now the argument for *sensa* derives its force from a subtle interplay of these two themes--an interplay which is extremely difficult to unravel. In my opinion neither Cornman nor Hooker has succeeded at this task; in fact, the formalised statements of Sellars' argument which they have contrived tend to obscure the complexity of this interplay by focusing attention on one principle or the other. Thus Cornman early in his discussion, disposes of the problem of why and how colour terms must appear in the scientific image, and concentrates on attacking the principle of reducibility. Hooker, on the other hand, takes the opposite tack, attacking Sellars' account of colour concepts in the manifest image, and his requirements for their successors in the scientific image.

Neither of these themes, or principles, are expounded or defended in detail by Sellars. However, from what he does say about them, they appear not implausible. The notion of counterparts or successor concepts is captured by Hooker as follows:

Principle of Framework Transformation Adequacy:
For any two conceptual schemes s , s^1 , if s^1 is to be an adequate successor to s then s^1 must represent in itself the logical structures ('logical spaces') of each of the concepts of s .

Thus the scientific image will contain counterpart or successor concepts of the concepts of the common sense framework. In

general these successor concepts will be "enriched:" thus the successor concept of a pink ice cube will be a complex of imperceptible particles. This enrichment reflects the fact that the scientific image provides a more detailed description of the world. The counterpart relation is one of identity in only a very weak sense--Leibniz Law does not hold.

Hooker does not dispute this principle, but he believes that Sellars uses it improperly in his argument for *sensa*. According to Hooker, the principle requires that the scientific image contain *sensa* whose

...properties reproduce the 'logical spaces' of the secondary qualities. To do this latter is...to preserve the truth of all true higher order claims about the original concepts. In the present context it is to reproduce all of the conceptual relationships among the original concepts, for example, determinate/determinable structure, colour compatibility and so on, as well as reproducing the other logical characteristics of the concepts, for example, simplicity and homogeneity.³

However, Hooker later claims that the principle by itself will not suffice to carry the argument for *sensa* as basic particulars--the principle must be strengthened to insure the simplicity and homogeneity of colours:

Thus we find appended to the general argument for (the principle) a special argument operative for just the perceptible qualities, a special argument to the effect that the 'logical space' of the perceptible qualities and relations of physical things and processes on which that of the attributes and relations of raw feels is modeled is, in an important sense, closed. Perceptible qualities and relations are...pure occurrent qualities and relations....The intrinsic structure of their 'closed' logical space requires ...that they be re-located in the Scientific Image, rather than reductively identified with other properties within it.⁴

Hooker has taken arguments from two different papers by Sellars and combined them as if they constituted a single argument. But it is not at all clear that he is justified in doing so; indeed I believe he is not. His analysis of the application of the principle of framework transformation adequacy to the argument from homogeneity fails mainly because Hooker does not recognize that the principle of reducibility also plays an essential role. Thus I shall consider these to be two arguments for the primitiveness of *sensa*, if for no other reason than that they appear in different places. The third appears in "Phenomenalism" and argues essentially that since colours are primitive in the manifest image, they must also be primitive in the scientific image--thus perhaps the principle of framework adequacy requires that primitives in one framework be primitive in another.

The other side of the argument for the basicness of *sensa* is the principle of reducibility which is an ontological condition concerning the structure of conceptual frameworks. It is expressed by Sellars as follows:

If an object is in a strict sense a system of objects, then every property of the object must consist in the fact that its constituents have such and such qualities and stand in such and such relations or, roughly, every property of a system of objects consists of properties of and relations between, its constituents.

Alternatively it may be said that the properties of conglomerates of particles are defined in terms of the properties and relations of the basic particulars which constitute the whole.

The principal arguments for *sensa* appear in "Philosophy and the Scientific Image of Man" and "Phenomenalism." The arguments are similar but importantly different. In both cases Sellars first considers the problem of accommodating inanimate physical objects of the manifest image in the scientific image. The problem is to reconcile the occurrently coloured objects of the former with systems of imperceptible objects of the latter. Since an instrumentalist interpretation of micro-theoretical entities is ruled out by Sellars two possibilities remain:

- a) Manifest objects are identical with systems of imperceptible particles in that simple sense in which a forest is identical with a number of trees.
- b) Manifest objects are 'appearances' to human mind of a reality which is constituted by systems of imperceptible particles.

In "Philosophy and the Scientific Image of Man" the argument depends on the homogeneity of colours. Alternative (a) envisages that manifest objects and systems of imperceptible objects both exist in a single coherent conceptual framework--one in which therefore, the principle of reducibility applies. For example, a ladder is a system of pieces of wood, none of which itself is a ladder. The argument proceeds:

But the case of a pink ice cube, it would seem clear, cannot be treated in this way. It does not seem plausible to say that for a system of particles to be a pink ice cube is for them to have such and such imperceptible qualities, and to be so related to one another as to make up an approximate cube. Pink does not seem to made up of imperceptible qualities in the way in which being a ladder is made up of being cylindrical

(the rungs), rectangular (the frame), wooden, etc. The manifest ice cube presents itself to us as something which is pink through and through, as a pink continuum, all the regions of which, however small, are pink. It presents itself to us as ultimately homogeneous; and an ice cube variegated in colour is, though not homogeneous in its specific colour, 'ultimately homogeneous', in the sense to which I am calling attention, with respect to the generic trait of being coloured.⁷

This argument is taken to rule out (a). Its force seems to be that the concept of colour we do have is not that of a complex of qualities and relations; therefore it is a simple or primitive property. The conclusion is that manifest objects do not exist; they are 'appearances' to perceivers.

In "Phenomenalism" the identity expressed by (a) is also rejected but this time by an argument that begins with the primitiveness of colour predicates:

The attempt to melt together Eddington's two tables does violence to both and justice to neither. It requires one to say that one and the same thing is both the single logical subject of which an undefined descriptive predicate (e.g. 'red') is true, and a set of logical subjects none of which is truly characterized by this predicate, thus raising all the logical puzzles of 'emergence.' And if, as is often done, 'red' as predicable of physical objects is tacitly shifted from the category of primitive descriptive predicates (where it properly belongs) to the category of defined descriptive predicates by being given the sense of 'power to cause normal observers to have impressions of red,' then the very stuffing has been knocked out of the framework of physical objects, leaving not enough to permit the formulation of the very laws which are implied by the existence of these powers, and which are pre-supposed by the micro-theory which might be invoked to explain them.⁸

The conclusion again, of course, is that the physical objects of the manifest image do not exist--or, alternatively, that the framework of perceptible physical objects is false. The question arises, are these supposed to be two separate arguments for the non-reality of the common sense framework of physical objects? Indeed, is "ultimate homogeneity" the same property as "simplicity" as applied to colours?

I think the answer to both these questions is no. Sellars is not very explicit about the concept of ultimate homogeneity which suggests that it means just what it seems to mean; that is, it is a property that we detect by observation. Thus Sellars says that the manifest ice cube "presents itself to us as ultimately homogeneous." But is the primitiveness or simplicity of colour also something we detect by observation? I think not. In the passage quoted above, Sellars argues that colours are not defined or definable as powers because he holds that the manifest image requires physical objects to be occurrently coloured. But this does not rule out a reductive definition of colours. The concept of ultimate homogeneity is required to rule this out. Sellars tells us that:

...the concept of ultimate homogeneity is closely related to the traditional concept of a simple quality. It differs primarily by relating the latter to the logic of whole and part. Applied to my example it says that the pinkness of a whole (the pink ice cube) does not consist in a relationship of non-pink parts.⁹

Perhaps we can detect by observation that manifest coloured objects are non-complex i.e. that colour terms are not defined

terms in the manifest image. But Sellars' argument requires an additional premise that colours are not definable in terms of the properties and relations of the theoretical entities of the scientific image. And it is specifically the ultimate homogeneity of colour which cannot be so defined--and which therefore is essential to the argument. On the other hand, the claim that physical objects must be occurrently coloured--i.e. that physical colours are not to be analysed as powers to cause sense impressions--is also an essential part of the argument to this point. Thus the arguments in "Philosophy and the Scientific Image of Man" and "Phenomenalism" are not separate and independent arguments but rather must be considered two aspects of the same argument--two aspects which are jointly necessary and mutually supporting.

The conclusion is that the coloured objects of the manifest image simply do not exist. This move is possible since colours analogically conceived with their logical space intact already exist as features of sense impressions. Now the argument(s) for accommodating sense impressions into the scientific image follows a similar pattern to that of the argument(s) just examined. Both thoughts and sense impressions are theoretically conceived inner episodes of persons. Thoughts are conceived by analogy with overt speech, sense impressions, by analogy with the occurrently coloured physical objects which are their cause. Now with thoughts, there is no problem in principle in accommodating them in the scientific image, which construes persons as

neurophysiological systems. Since Sellars accepts a functional analysis of language, the analogy concerns only the role of whatever tokens are involved:

...thoughts in the manifest image are conceived not in terms of their 'quality,' but rather as inner 'goings-on' which are analogous to speech...

The point is an important one, for if the concept of a thought is the concept of an inner state analogous to speech, this leaves open the possibility that the inner state conceived in terms of this analogy is in its qualitative character a neurophysiological process.¹⁰

The case with sense impressions is not so straightforward, however, since the analogy concerns the intrinsic character of the things involved:

Thus a 'blue and triangular sensation' is conceived by analogy with the blue and triangular (facing) surface of a physical object which, when looked at in daylight, is its cause. The crucial issue then is this: can we define, in the framework of neurophysiology, states which are sufficiently analogous in their intrinsic character to sensations to make identification plausible? The answer seems clearly to be 'no.' This is not to say that neurophysiological states cannot be defined (in principle) which have a high degree of analogy to the sensations of the manifest image. That this can be done is an elementary fact in psycho-physics. The trouble is, rather, that the feature which we referred to as 'ultimate homogeneity,' and which characterizes the perceptible qualities of things, e.g. their colour, seems to be essentially lacking in the domain of the definable states of nerves and their interactions.¹¹

At this point, Sellars poses the following antinomy:

- (a) the neurophysiological image is incomplete, i.e., and must be supplemented by new objects ('sense fields') which do have ultimate homogeneity, and which somehow make their presence felt in the activity of the visual cortex as a system of physical particles;

(b) the neurophysiological image is complete and the ultimate homogeneity of the sense qualities (and, hence, the sense qualities, themselves) is mere appearance in the very radical sense of not existing in the spatio-temporal world at all.¹²

The solution, of course, is that there is something more basic than the particulate foundation of physics:

...although for many purposes the central nervous system can be construed without loss as a complex system of physical particles, when it comes to an adequate understanding of the relation of sensory consciousness to neurophysiological process, we must penetrate to the non-particulate foundation of the particulate image, and recognize that in this non-particulate image the qualities of sense are a dimension of natural process which occurs only in connection with those complex physical processes which, when 'cut up' into particles in terms of those features which are the least common denominators of physical process--present in inorganic as well as organic processes alike--become the complex system of particles which, in the current scientific image, is the central nervous system.¹³

Again the argument is presented somewhat differently in "Phenomenalism" but with similar results. Sense impressions of the manifest image are adverbial states of single logical subjects, persons. But identifying persons with multiplicities of logical subjects undermines the logic of sense impressions:

For whether these parts be construed as material particles or as nerve cells, the fact that they are a plurality precludes them from serving either jointly or separately as the subjects of the verb 'to sense red-triangle-wise.' We must therefore ...introduce a new category of entity ('phantasms' or 'sensa' we might call them) with predicates the logical space of which is modelled on that of visual impressions, as the latter was modelled on that of coloured and shaped physical objects.¹⁴

At this point Sellars is more specific about what he conceives

sensa to be: the basic ontological category is not things but events. The questions asked by a science of perception becomes:

How are we to conceive the relationship between the sequence of micro-physical 'events' which constitute a brain's being in the physical state appropriate to the occurrence of a red and triangular sensum, and the sequence of 'events' which is the sensum?¹⁵

We are faced with the following facts:

a) Sellars is convinced that colours are so basic to our account of reality that they must appear transposed but with their logical space intact in the scientific image. For whatever reason, this includes that they be primitive i.e. predictable of ontologically basic entities in the scientific as well as the manifest framework.

b) Sellars evidently thinks this is possible only if we adopt a basic ontology of "events" rather than "things." It may be that Sellars has reasons independent of the argument for sensa for this position. My concern is whether this conclusion can be drawn from the particular considerations advanced in the argument for sensa. Everything hinges on the reasons for which Sellars holds colours must be primitive. My conclusion will be that Sellars has not shown that colours need be primitive in the ontological sense required for his argument. At best, they must be regarded as epistemological primitives. However, there is some difficulty in determining Sellars' reason, or reasons for holding colours to be primitive. I discern the following three possibilities:

(1) In "Phenomenalism" the argument might be taken to be that the primitiveness of colours in the manifest image is an essential part of their logical space and so must be retained in the scientific image. I have already suggested that this argument requires additional premises. In Chapter 4 I show why it is insufficient, and therefore requires supplementation by further arguments such as:

(2) The ultimate homogeneity of the manifest image is part of the logical space of colour concepts in the manifest image and must appear in the scientific image. Since we concluded from this ultimate homogeneity in the manifest image that colours were primitive predicates, a similar argument forces us to conclude that colours transposed into the scientific image must also be primitive, i.e. simplicity necessarily follows from homogeneity. I shall argue, in the last chapter, that this is the most persuasive argument Sellars advances for the thesis that colours are ontologically primitive in the scientific image. However, there are some very complicated issues involving both the principle of framework transformation adequacy and the principle of reducibility and it is in sorting these issues out that the weakness in Sellars' argument becomes clear. It is the principle of framework adequacy which requires that the logical space of colour be reconstructed in the scientific image. This principle requires that certain (manifest) concepts appear in the scientific framework in analogous form, and others in identical form--the latter being, in general, formal second order attributes which are used to specify the former, merely analogous,

concepts. How is the concept of "ultimate homogeneity" to appear in the scientific image--as identical to the manifest concept of homogeneity, or merely as transposed into an analogous concept?

That it is the latter is, I think, beyond argument. The manifest concept of ultimate homogeneity involves physical objects and their geographic parts: i.e. the concept is (what I shall call) a spatial one. The scientific image concept of ultimate homogeneity, in which we are interested, involves sense impressions, which are analogically conceived states of perceivers. But the fact that these concepts e.g. "coloured" sense impressions are analogical does not necessarily mean that "homogeneity" must be analogically construed. However, since it is clear that nothing will have the spatial homogeneity of physical objects, the scientist will attempt (and is so justified) to define or reconstruct an analogical concept of homogeneity from the primitives that exist (e.g. neurons).

An analogical concept of homogeneity would be one which preserved certain essential second order properties of the model concept--spatial homogeneity. Roughly, the relevant properties would be the mathematical features of the continuum. The science of analogue computer processes shows, I believe, that a neurophysiological process could be defined with the right properties. Sense impressions would be identified with these processes, and the spatial homogeneity of coloured objects would "map" onto certain temporal (perhaps) features of these

processes--i.e. we would have an analogous, "temporal" concept of homogeneity.

Now I think Sellars would agree that it is an analogous, not identical concept of homogeneity which applies to coloured sense impressions. His argument, by speaking of "processes" and "events," opens the way for a temporal concept of homogeneity. But now it is not as clear that the primitiveness of colours of sense impression is entailed by this new, scientific image, concept of homogeneity, as (I suggested) the primitiveness of colour in the manifest image is entailed by their spatial homogeneity. In the latter case, the principle of reducibility was invoked to conclude that the occurrently coloured physical objects of the manifest image did not exist in the scientific image. The reason given was that the scientific successors of manifest physical objects were systems of imperceptible particles. In the logical space of these imperceptible particles it seems clear that nothing analogous to the spatial homogeneity of physical colour can be defined. The conclusion is that in the final analysis, physical objects construed scientifically as systems of particles, are not occurrently coloured--i.e. in a sense, those physical objects which were occurrently coloured, do not exist.

When we move to the problem of putting the colours of sense impressions into the scientific image, Sellars attempts a similar treatment, reaching of course, a different conclusion. Manifest sense impressions are states of persons, construed as single

logical subjects. Can states be defined of neurophysiological systems which are sufficiently analogous to warrant identity? We saw above (pp. 12) that Sellars thinks not. His reason is specifically that "the feature which we referred to as 'ultimate homogeneity,' and which characterizes the perceptible qualities of things...seems to be essentially lacking in the domain of the definable states of nerves and their interactions."¹⁶

In my view, this argument is clearly acceptable only if Sellars is using the spatial notion of homogeneity. If it is accepted that an analogous "temporal" notion is sufficient, the question of whether or not neurophysiological states or, more appropriately, sequences of states, i.e. processes can be defined which can be identified with sense impressions is more difficult to decide. According to the present account, just as perceptible physical objects are the model for both manifest and scientific sense impressions, the homogeneity of the colours of these objects is the model for whatever homogeneity is applicable to scientific sense impressions. When it comes to the point of considering in what terms this scientific concept of homogeneity may possibly be defined (i.e. in terms of what objects' sense impressions may be defined) we must not mistake the model for the thing itself. In any such case of modeling, there are certain properties shared by the model and the thing modeled, and other properties not shared. My point is that the science of perception is not sufficiently advanced that we can specify which properties are held by both (the positive analogy) and

which are not (the negative analogy). In such circumstances, only confusion can result in an argument which makes an uncritical use of the model. My view is that Sellars' argument does tend to rely too heavily on the model concepts themselves. Scientific sense impressions will turn out to be extremely complex constructs of neurophysiological primitives. An argument such as Sellars gives to the effect that sense impressions could not be defined by the primitives currently postulated by neurophysiology would have to unravel some of these extremely complex details. Indeed we are faced by a great deal of ignorance on both sides of the disputed reduction--of what exact nature the homogeneity, on one hand, and on the other, the unknown possibilities of definable neurophysiological states--modeled on computer states. This difficulty is acknowledged, ironically, by Sellars in a different context:

Over and above this all we need is to recognize the force of Spinoza's statement: 'No one has thus far determined what the body can do nor no one has yet been taught by experience what the body can do merely, by the laws of nature insofar as nature is considered merely as corporeal and extended.' (Ethics, Part Three, Prop. II (note)).

(3) Faced with the question of why colour must be primitive in the scientific image, however, Sellars gives a very different answer. The question arises in a reply Sellars gives to Cornman concerning the primitiveness of the predicate "senses redly." Although it is primitive in the manifest image, it will be a defined predicate in the scientific image because it is not true of any of the many logical subjects which constitute persons in the scientific image:

The conclusion... is that ~~this successor predicate~~ must be a defined predicate, but that the primitive character of "senses redly," which has its origin in the primitive character of "red" as a predicate of physical objects, will be preserved if this successor predicate were defined in such a way that its definiens includes a reference not only to such scientific objects as are involved in non-living matter, but also to a new domain of particulars (sensa) to which a new primitive predicate "red" applies.

If it be asked why the successor in the Scientific Image to sensing redly and, ultimately, to red as an occurrent quality of the physical objects of the Manifest Image, must preserve the primitiveness or irreducibility of color... the answer lies in the distinctive character of the explanandum which called for the introduction of sense impressions in the first plane.¹⁸

But what is this "distinctive character?" Surely it must be more than the fact of being itself primitive in its own conceptual framework. Indeed the passage itself admits that sense predicates which are primitive in the manifest image i.e. "senses redly" are defined in the scientific image. Thus the "distinctive character" of something's being red must be more than that the predicate is primitive, or that the thing is a single logical subject. The paragraph, quoted above is accompanied by a footnote which refers us to two other places in Sellars' work. One reference is to an earlier part of the same paper in which Sellars does indeed discuss the explanandum in question, which is human behaviour. Part of the behaviour which science must explain is the fact that humans have perceptual propositional attitudes i.e. are sometimes disposed to say, in certain circumstances "Lo! Here is a red and rectangular object." Sellars' point in this passage is that he takes seriously, as Quine does not

...the claims of an intermediate stage of explanation--intermediate, that is, between explanation in terms of behavior and propensities to behave, and explanation in terms of neurophysiology and central states. This intermediate explanatory framework is roughly (but only roughly), that of thoughts and sense impressions...as construed in classical theories of mind.¹⁹

Sellars complains about the failure of certain of his colleagues

to pay serious attention to the problem of specifying the conditions which an adequate scientific account of human behavior must meet. This problem calls for just that careful examination of what we already know about human behavior in terms of existing categories, for which, with few exceptions, they have no patience. This failure leads to a reliance on overly simple and inadequate paradigms of what will count as a "scientific object" or "bodily state" or "neuro-physiological process" in this anticipated scientific account.²⁰

Despite these suggestive passages, however, Sellars does not go into any more detail at this point (or, as far as I can see, in this paper) to spell out what it is about "knowledge of human behavior in terms of existing categories" which would persuade us to require that a scientific account of colours preserve their primitive character.

The other reference is to the concluding sections of "The Identity Approach to the Mind-Body Problem." At this point Sellars does discuss his belief that sensa are primitive:

...instead of the primitive predicates of the reduced theory ending up as defined predicates in the unified theory...these primitive predicates [i.e. applying to sense impressions] could perfectly end up as primitive predicates once more in the unified theory....The to-be-discovered sense-impression universals would be no more complex than the sense-impression universals expressed by current sense impression predicates; they would have a different categorical framework, and be nomologically related.

to (but not analysable into) universals expressed by other primitive predicates in the to-be-achieved unified sense impression, brain state theory.... (The logical space of sense impressions) would not... have become internally more complex in the way in which the logical space of chemical properties becomes internally more complex by virtue of their identification with micro-physical properties. That is to say there would be no increase in complexity with respect to what might be called the factual content of sense impression universals. Such increased complexity as occurred would be of a logical character.²¹

Again, the question is why Sellars holds this thesis and there appears to be an answer in the last paragraph of the paper. The logical space of the attributes of sense impressions is modelled on that of physical objects which is in an important sense closed:

Perceptible qualities and relations are...pure occurrent qualities and relations. They are neither dispositional nor mongrel states....Now it is not the 'internal structure' of the families of occurrent perceptible qualities and relations which generate the demand for theoretical explanation, but rather the nomological structure of the changes and interactions of the physical things and processes to which these qualities and relations belong. Roughly, it is not such facts, expounded in a 'phenomenology' of sensible qualities and relations, as that to be orange is to be between red and yellow in color which demand scientific explanation, but rather such nomological facts as that black objects sink further into snow than white objects when the sun is shining.²²

Thus it seems we have an argument or reason for believing that in the final scientific image, colour terms will be primitive just as they are in the manifest image. This discussion is reminiscent of Sellars' argument in "Phenomenalism" that the framework of physical objects is unreal. That argument, I claimed, was not independent of the argument from homogeneity

because the fact that colours are pure occurrents does not necessarily mean they are primitive. The fact that colours of physical objects are pure occurrents in the manifest image rules out the possibility that they are, or can be defined as powers to cause sense impressions. The significance of this fact is that there does not seem to be any other way the colours of physical objects might be defined. The homogeneity of such colours is explicitly and exclusively spatial and is obviously lacking in the definable states of imperceptible particles.

Now, however, the case is quite different for we are talking about the "colours" of sense impressions. We can grant that, perhaps for similar reasons as before, the "colours" of sense impression must be pure occurrents. This fact has quite different implications than it did before. The fact that the attributes of sense impressions are pure occurrents rules out the possibility of defining them as dispositions. Now, however, we cannot conclude from the occurrence of such attributes to their primitiveness because there may be definable states of neurophysiology which are sufficiently analogous to colours of physical objects to warrant identifying them with the (scientific counterparts of the) "colours" of sense impressions."

We saw earlier that the argument that the logical space of colour is closed because purely occurrent etc. was taken by Hooker to be Sellars' principal argument for the primitiveness of sensa. Indeed, some of the things Sellars says strongly gives this impression. However, I will argue in Chapter 3 that this

argument from the purely occurrent nature of colours is not sufficient to conclude they must be primitive in the scientific image. At most, this provides a motivation for so arguing. The most that can be concluded from the fact that colours are (and, one must add, must remain) pure occurrents, is that the logical space of colour does not require scientific explanation in the same sense as does, for example, certain dispositional properties of coloured objects--the fact that black objects sink further into the snow than white objects when the sun is shining. But this fact about the logical space of colours does not rule out the possibility that colours may be ontologically reduced--in other words, the fact that the logical space of colours is closed in the manifest image is not by itself sufficient to require that it not be "enriched" in the scientific image.

Nonetheless it is clear that Sellars does believe the logical space of colours, in the scientific image, will not be more complex than in the manifest image. While the argument that colours, being pure occurrents, do not require scientific explanation serves, perhaps, to motivate and support this belief, I think the argument from homogeneity is essential. I agree that something must appear in the scientific image which preserves the logical space of colours, but dispute the claim that the predicates which stand for sensuous properties, in particular colour predicates, need be primitive. The preservation of the logical space of colours is necessary to account for our

forming conceptual, and particularly, perceptual, pictures of the world. It is only within the framework of the logical spaces of our various sensory faculties that the stimuli impinging on our receptors provides the information which constitutes perceptual representation. As part of the explanation of how we come to know the world, the essential role of the logical spaces of sensuous properties would explain their purely occurrent character. However, neither their role in interpreting stimulus information, nor their occurrent character requires that they preserve the irreducibility they enjoy in the manifest image.

An essential and pervasive aspect of Sellars' style of argument is his use of model and analogy. Indeed this technique itself is drawn supposedly from the model of normal science-- and assumes the force of a normative principle of rational enquiry. Yet Sellars is somewhat more straightforward in his explicit use of model and analogy in the analysis of concept formation, and rational inquiry in general; its use as a normative force in Sellars' own reasoning is somewhat more subtle. The reason for this is no mystery; the fact is that Sellars has little to say about what justifies the use of model and analogy in philosophical endeavour. In specific cases, when arguing from the familiar to the unfamiliar he would probably be at a loss to tell us why he is justified in stipulating that certain features of the analogy hold while others don't. But if we are to come to definite conclusions, if we are to gain new knowledge instead

of illuminating the old, it is precisely the extent and limit of a particular analogy that must be justified. Of course, Sellars has given us some idea of how such rational inquiry is to be justified--the key notion is explanatory coherence. Yet Sellars gives us little detail about this notion--again, presumably, it is more or less intuitively understood on the basis of our understanding of the process of rational inquiry--specifically science--to date.

This pervasive use of models explains why Sellars often tends to be long on insight and suggestive ideas but short on specific, concrete convincing conclusions, as the argument for *sensa* purports to be. Prima facie, the argument can easily be dismissed as absurd or eccentric--as may be the attitude of the great number of philosophers who have said nothing about it. Even those who seem to have taken it seriously enough to attempt to refute it, Cornman and Hooker, have been forced to supply so many details on their own initiative, that the arguments for *sensa* they come up with are weak and easily disposed of. My opinion is that if the smoke could be cleared from these analogies, the convincing details distilled out, we would find either an obviously false argument for a radical ontological thesis, or a convincing argument for a less remarkable conclusion. Yet such an endeavour would entail plunging into the vast complex of detail which Sellars' technique seeks to outflank, and is quite beyond the scope of this paper. Thus I must content myself with pointing out certain key areas where models function in the argument and indicating certain related problems.

The use of model and analogy as a normative principle of rational inquiry purports to extrapolate validly from the familiar to the unfamiliar. Thus, on a global scale the manifest image is the known and provides the models for comprehending the structure and content of the developing scientific framework. Yet any picture which sees the manifest image as a paradigm conceptual framework, which supplies paradigm principles and concepts as a sort of metaphysical/metalinguistic foundation of knowledge is certain to run into difficulties as being too simplistic, for two reasons. In the first place, the sort of justification which pertains exclusively to the manifest image [e.g. the foundational justifications of traditional empiricism] will not serve as a paradigm for a justification which rejects one conceptual framework (the manifest) for another, (the scientific). For this, the notion of explanatory coherence must be introduced as a novel principle of justification. It is novel because it deals specifically with a unique relationship--between the manifest and scientific images. Speaking more practically, this difficulty surfaces when we ask from where comes the principle of framework adequacy. The model for the application of this principle to the problems of perception appears to be less problematic cases of scientific explanation--the "replacement" of chemical theory by physical theory, or the replacement of Newtonian physical theory by Einsteinian. The point is that the development of theoretical explanation was a unique episode in the history of knowledge and cannot be

understood solely within the context of the manifest image. The principle of framework adequacy is a case of extrapolation from the familiar and accepted to the unfamiliar--but not thereby, from the manifest or common sense to the scientific or "real."

The significance of this point is that it is an illustration of a broader problem, which is the second reason why the manifest image can only with difficulty be regarded as the paradigm. The paradigmatic role of the manifest image presupposes, of course, a correct analysis of the manifest image. Unfortunately, careful consideration of the manifest image reveals it not to be a static, independent conceptual framework, prior and alternative to the scientific image; rather it itself is liable to change in response to theoretical development. The distinction Sellars draws between the two is apparently straightforward, but on examination, extremely problematic. The point of the distinction is to define the limits of each framework, with the goal of extracting by analysis from the manifest, certain principles which will be normative to the development of the scientific. The key distinction, it seems, is that manifest concepts do not include those which are, properly speaking, theoretical, whereas the scientific does. Yet the contrast between the theoretical and the non-theoretical is difficult to clarify--even in Sellars' manifest image, there are "objects" e.g. sense impressions which are very similar to theoretical. In a similar sense, as I will argue in Chapter 3, the manifest

image could be construed along indirect realist lines--which could be seen as a possible development of Sellars's manifest image, an intermediate between the manifest and the scientific. Sellars' defense against this line of thought is that this is not the framework we in fact possess. But what justifies us in extrapolating from a framework we in fact possess, granted more adequate frameworks are possible?

Sellars, in "Philosophy and the Scientific Image of Man," mentions earlier stages of the manifest image--i.e. predecessor frameworks at one time accepted as real. Yet no attempt is made to correlate that conceptual evolution with the projected one. Why did man cease to see trees as persons? Why were certain principles retained, and other rejected? In retrospect, we can speculate that certain forces, among them, developing, primitive science guided man toward an acceptance of more efficient, because more adequate, ways of viewing the world. But in trying to project forward, as we are, this has little relevance. There is no understood method of picking which aspects of our framework are to be retained and which rejected--i.e. which course of action will lead to a more adequate framework.

The reason I make these remarks is not so much to cast aspersion on Sellars' method as to illuminate the very real, but unacknowledged, difficulties inherent in such an approach. In groping for a more adequate conceptual framework we have no alternative of course, to extrapolating from the accepted one. The problem is, to repeat, to pick out those aspects of the present

framework which are justifiably maintained as the positive analogy. My view is that if we are to contemplate the possibility of radical conceptual change we must be as bold as necessary in doing so, conversely, as cautious as possible in projecting familiar elements of the manifest image. We must not allow our models to get in our way. But this ~~is precisely what~~ Sellars does. The key issue is the principle of reducibility. How is this to be understood? Sellars is not very clear about this but, rather, suggestive. Obviously the principle is drawn from the manifest image: the illustrating example is that the property of being a ladder is made up of being cylindrical (the rungs), rectangular (the frame), wooden, etc. Expressions of the principle seem to use the language of whole and part, "consists of," but in a curious way:

"every property...must consist in the fact...."

"every property...consists of properties of...."

Not that Sellars is unaware of the difficulties involved in the principle; at one point he tells us that "A defence of this principle would take one right to the very heart of the philosophy of logic, relating, as it does, the functional calculus to the calculus of individuals."²³ What this suggests is that the notion of "consists of" which we find in discussions of property reduction is merely a rough model to enable us to understand a still mysterious relation. Another model encountered is the notion of definition; thus we may say the property of being a tree is defined as being of a certain size, being woody, green

etc. But it cannot be overemphasized that these are merely models and cannot be taken too seriously in scientific extension. Although the notion of a certain property being defined by another is fairly clear in the manifest image, it is decidedly not in the scientific. The reason is simple: in the manifest image, we can construct, or define concepts out of well understood concepts--e.g. colour terms. In the scientific image the primitive concepts are not well understood. Rather it is more likely that the concepts of the definiendum are as well or better understood than those of the definiens. Scientific primitives are known indirectly, defined only implicitly and partially. Part of their definition involves what higher level constructs they can serve to define--i.e. their definition is their overall role in the theory. Both models of the principle of reducibility are just that--models. How they extend to the scientific framework is a question that will probably only be decidable when the scientists have largely constructed the appropriate theories and left them to the philosophers to analyse.

Thus my chief criticism of Sellars' argument is that he seeks to convince by placing certain paradigms before us and inviting us to draw conclusions from them. Not that Sellars is unaware of the difficulties, but that he suppresses them for the purposes of the argument. In the argument from homogeneity, this tendency is especially evident. In that argument, in "Philosophy and the Scientific Image of Man" Sellars argues first

from the homogeneity of coloured objects in the manifest image, to their unreality. This is not part of the argument for *sensa* but serves as the model for it. There are two models involved: the concept of spatial homogeneity which serves as the model for whatever analogous concept of homogeneity science will postulate, or define for sense impressions. The second model is a particular application of the principle of reducibility: the fact that pink is not made up of, does not consist of, imperceptible qualities. Whatever we may think of this use of the principle, it should be clear that it cannot straightforwardly be extended to the homogeneity of *sensa* impressions. Yet not only is the possible disanalogy not mentioned by Sellars, his statements seem contrived to suggest that the principle applies literally:

Putting it crudely, colour expanses in the manifest world consist of regions which are themselves colour expanses, and these consist in their turn of regions which are colour expanses, and so on; whereas the state of a group of neurons, though it has regions which are also states of groups of neurons, has ultimate regions which are not states of groups of neurons but rather states of single neurons.²⁴

FOOTNOTES

¹C. A. Hooker, "Sellars' Argument for the Inevitability of the Secondary Qualities," Philosophical Studies 32 (1977), 335-348.

²Ibid., p. 339.

³Ibid., pp. 336-337.

⁴Ibid., p. 345. Hooker is quoting from Wilfrid Sellars, "The Identity Approach to the Mind-Body Problem," Boston Studies in the Philosophy of Science, Volume Two (New York: Humanities Press, 1965), edited by Robert Cohen and Marx Wartofsky, pp. 74-75.

⁵Wilfrid Sellars, Science, Perception and Reality (London: Routledge and Kegan Paul, 1963), p. 27.

⁶Ibid., p. 26.

⁷Ibid.

⁸Ibid., p. 98.

⁹Wilfrid Sellars, "Science, Sense Impressions, and Senses: A Reply to Cornman," Review of Metaphysics 25 (1971), 407.

¹⁰Sellars (1963), pp. 32-33.

¹¹Ibid., p. 35.

¹²Ibid., p. 36.

¹³Ibid., p. 37.

¹⁴Ibid., p. 101.

¹⁵Ibid., p. 102.

¹⁶Ibid., p. 35.

¹⁷Ibid., p. 33.

¹⁸Sellars (1971), p. 424.

¹⁹Ibid., p. 398.

²⁰Ibid., p. 399.

²¹Sellars (1965), p. 72.

²²Ibid., p. 74.

²³Sellars (1971), p. 393.

²⁴Sellars (1963), p. 35.



Chapter 2: Meaning and Truth

It is impossible to appreciate the force of Sellars' argument for sense--and the role of the principles of reducibility and of framework transformation adequacy--without understanding the importance and nature of the manifest image. But this in turn presupposes an understanding of the role, in Sellars' philosophy, of the concept of a conceptual framework. It is to a consideration of this latter subject that I turn for the moment.

For Sellars, the primary unit of meaning is the conceptual framework. This represents a break with traditional empiricism which held that some words derive their meaning solely by means of a direct ostensive link with the world. These empiricists held that the learning of language consists initially at least, of the formation of a battery of concepts, derived by abstraction from direct acquaintance with the world and linked to words by association. This battery of concepts formed a foundation for the construction of the great edifice of human knowledge. Thus the meanings of highly complex concepts ultimately reduced to these directly known concepts. Furthermore certain beliefs about the world were said to be basic. These beliefs involved those concepts which were directly known, and were thought either not to be in need of justification, or to be self-justifying. In other words, they could be known independently

of all other knowledge, to be true. Other beliefs, those involving more complex concepts, could be justified because they were logically related to those beliefs which were basic. Necessarily involved in such an account, of course, is a logical atomist view of language.

Now Sellars does not accept all of the theses of the traditional empiricist although his approach is similar at times. Chief among his differences with traditional empiricism is his rejection of the notion of a "given." The empiricist account of concept formation is radically mistaken, according to Sellars, because it presupposes that we have a preconceptual cognitive awareness of the world. In "Empiricism and the Philosophy of Mind" Sellars attempts to show that the foundationalist account of knowledge is false. As a part of Sellars' general critique of the "given," is an alternate account of how we come to have knowledge of the world. It is this account which I am about to look at, for it is an essential part of Sellars' views about scientific explanation.

Among conceptual frameworks, the manifest image occupies a special position, for it is a direct descendant of the conceptual framework which Sellars calls the "original image"

...in terms of which man came to be aware of himself as man-in-the-world. It is the framework in terms of which, to use an existentialist turn of phrase, man first encountered himself-- which is, of course, when he came to be man.... I want to highlight...the paradox of man's encounter with himself...that man couldn't be man until he encountered himself....Its central theme is the idea that anything which can properly be called conceptual thinking can occur only within a framework of conceptual thinking in terms of

which it can be criticized, supported, refuted, in short, evaluated....(A) conceptual framework is a whole which, however sketchy, is prior to its parts, and cannot be construed as a coming together of parts which are already conceptual in character. The conclusion is difficult to avoid that the transition from pre-conceptual patterns of behaviour to conceptual thinking was a holistic one, a jump to a level of awareness which is irreducibly new, a jump which was the coming into being of man.¹

The parts of a conceptual framework are words and the uniformities they follow in the speech of users of that framework. Thus even the most basic terms of a language, or framework, e.g. colour terms, do not gain their meaning solely by means of ostensive connection with sensation. The meaning also consists of their function in sentences and the rôle of those sentences in the language. Sentences are governed by rules which are both rules of meaning and rules of truth. I turn, then, to Sellars' account of meaning and truth, as it appears in Science and Metaphysics.

Meaning

Sellars' theory of meaning involves a radical rejection of much of traditional empiricism. In rejecting the notion of the "given" he also rejects the possibility of analysing the meaning of linguistic expressions in terms of the intentionality of thought. Sellars shows how semantic notions can be analysed in terms which require no reference to mental episodes. Meanings of words, or senses as he calls them, are to be analysed in terms of the role they play in our language. Sellars introduces

dot-quotes which form abstract singular terms. Thus 'not' applies to any expression which plays an analogous role in the language in which it exists, as 'not' does in our language (Sellars abstracts from the various uses of languages and considers them only in their epistemic use). Sellars uses this technique to explain our apparent reference to abstract objects. Thus he forms distributive singular terms as follows:

Triangularity = the 'triangular'

That $2+2$ equals 4 = the ' $2+2$ equals 4 '

The (individual sense) Socrates = the 'Socrates'

With this device, Sellars construes paradigmatic semantic statements such as

'Sage' (in F) stands for wisdom

as

the 'sage' (in F) stands for the 'wise'

otherwise as

'Sage's (in F) stand for 'wise'

and finally

'Sage's (in F) are 'wise's

indicating that "stands for" is to be interpreted as a specialized form of the copula and that semantic statements essentially classify expressions by function.

Sellars takes seriously the analogy of a language with a game. Just as chess may be played with a variety of different types of tokens, provided the tokens are distinguishable in ways that parallel standard chess pieces, so a language game may

be played with different tokens. What is important for the identity of a game is that the tokens are used according to the rules. For a language to maintain its identity, its tokens must participate in a set of semantic uniformities, described by rules of criticism. It is by functioning according to these rules that linguistic expressions can be said to be meaningful. Sellars provides a list of the rules for an idealized language as described in the Tractatus (PM):

I. Intra-linguistic Uniformities:

- (a) Formative (formation rules).
- (b) Consequential (transformation rules).
 - (α) Logical rules in the narrow sense (L-rules).
 - (β) Law-like statements construed as principles of inference (P-rules).
 - (γ) Consequence rules relating names, demonstratives and the language of Space and Time.

II. World \rightarrow Language Uniformities:

- (a) Language-entry (responsive):
 - (α) Demonstrative: responding to situations of different kinds with sentence tokens of the kind 'this is ---'.
 - (β) Sortal: responding to different kinds of situation with tokens of correspondingly different demonstrative sentences--that is, since PM is a subject-predicate language, sentences with correspondingly different predicates.
- (b) Mediated by combination of II(a) with I(b)(β) and I(b)(γ).

III. Language \rightarrow World Uniformities:²

Such a set of rules defines a conceptual framework. Different frameworks can be specified by dropping some rules, or types of rules. The above is, I presume, to be considered an approximation to the rules of our conceptual structure--a structure consisting of many conceptual frameworks related in different ways.

Truth

Statements such as

(The individual sense) Socrates is real

Wisdom is exemplified

That Socrates is wise obtains

are to be understood in terms of the concept of truth. Thus the last is construed as

The 'Socrates is wise' is true
or equivalently

That Socrates is wise is true.

Sellars defines truth in the following words:

...for a proposition to be true is for it to be...correctly assertible...in accordance with the relevant semantical rules, and on the basis of such additional, though unspecified, information as these rules may require....'True,' then, means semantically assertible ('S-assertible') and the varieties of truth correspond to the relevant varieties of semantical rule.³

Thus statements asserting the truth of a proposition are to be regarded as authorizing a performance--namely, asserting the proposition.

A statement such as

Wisdom is exemplified by Socrates

is to be construed roughly as

The result of replacing the 'x' in an 'x is wise' by a 'Socrates' is true, i.e. s-assertible.

The importance of this formulation is that although

Wisdom is exemplified by Socrates

appears to assert a relation between two objects, namely wisdom and Socrates, the statement is correctly to be analysed as a

metalinguistic one in which "Socrates" is mentioned, not used. This is because abstract entities such as wisdom and triangularity are not objects in the normal sense; rather they are what Sellars calls distributive objects, other examples of which are "the pawn" and "the lion" as in

The pawn captures en passant

and

The lion is tawny.

Wisdom and triangularity are linguistic distributive objects and are analyzed as

the ·wise·

and

the ·triangular·.

And just as we may "reduce" the above sentences as follows

Pawns capture en passant.

Lions are tawny.

so we may replace the ·triangular· by ·triangular's. The conclusion is that universals are not extra-linguistic objects.

The more general result is that semantical statements of the Tarski-Carnap variety do not have the form

(Linguistic item) R (non-linguistic item).

The connection of linguistic objects with the world is rather more indirect; it involves participating in semantical uniformities with the appropriate extra-linguistic objects, according to the relevant rules. Thus as we have seen statements such as

'Dreieckig's (in German) stand for triangularity
are classificatory statements of the form

'Dreieckig's (in German) are 'triangular's.

Denotation is analyzed using the notion of material equivalence
and a variable that ranges over senses, formed by dot quoting.

For example,

"Rational animal" (in E) denotes 'featherless bipeds'.

is analysed as

For some S, 'Rational animal's (in E) are S's,
and S is materially equivalent to 'featherless
biped's.

Where 'rational animal' is materially equivalent to 'feather-
less biped' if and only if

$$(x) x \text{ is a rational animal} \equiv x \text{ is a featherless biped.}^4$$

Picturing

Truth as S-assertibility is the generic meaning of truth;
it applies to all forms of truth--factual, mathematical, logi-
cal etc. As was noted in the definition "the varieties of
truth correspond to the relevant varieties of semantical rule."
In the domain of factual truth, the relevant rules are P-rules--
the material generalisations of a language--and the specific
concept of truth is picturing. It is primarily atomic state-
ments which make up "linguistic pictures" of the world. An
atomic statement is true (S-assertible) if the picture it makes
of the world is correct in terms of the semantic rules of the
language, false if incorrect. Molecular statements, on the
other hand, pick out sets of pictures:

Basic factual predicates come in families of competing predicates, one or other of which must be satisfied by every object which can satisfy a predicate of that family. If a is not f₁ it must be f₂ or f₃....

The S-assertibility of molecular statements ... is a function of the syntactical moves which connect them with disjunctions of conjunctions of non-negative and non-competing atomic statements, and of the S-assertibility of these conjunctive complex qua complex pictures.

Thus, if two families are $\langle \text{'f}_1, \text{'f}_2, \text{'f}_3 \rangle$ and $\langle \text{'g}_1, \text{'g}_2, \text{'g}_3 \rangle$

$$\sim(\text{f}_1\text{a} \cdot \sim\text{g}_1\text{b})$$

becomes the disjunction of conjunctions

$$(\text{f}_1\text{a} \cdot \text{g}_1\text{b}) \vee (\text{f}_2\text{a} \cdot \text{g}_1\text{b}) \vee (\text{f}_2\text{a} \cdot \text{g}_2\text{b}) \vee (\text{f}_2\text{a} \cdot \text{g}_3\text{b}) \\ \vee (\text{f}_3\text{a} \cdot \text{g}_1\text{b}) \vee (\text{f}_3\text{a} \cdot \text{g}_2\text{b}) \vee (\text{f}_3\text{a} \cdot \text{g}_3\text{b})$$

and is S-assertible if any of the disjuncts constitutes an S-assertible picture.⁵

Picturing, unlike truth as S-assertibility, is a relation between linguistic objects and the world. Pictures can be regarded as projections of the world on language according to the semantic rules of the language.⁶ On the other hand, atomic statements, in their role as pictures, are to be regarded as objects in rerum natura.⁷ It is because linguistic tokens function according to certain uniformities, in particular $\langle \text{World} \rightarrow \text{Language} \rangle$ and $\langle \text{Language} \rightarrow \text{World} \rangle$ uniformities that they can be said to picture extralinguistic states of affairs. The key notion is correlation:

'fa's (in L) correctly picture O as ϕ ... tells us that (in L) utterances consisting of an 'f' concatenated with an 'a' are correlated with O, which is ϕ , in accordance with the semantic uniformities

which correlate utterances of lower-case letters of the alphabet with objects such as O, and which correlate utterances of lower-case letters of the alphabet which are concatenated with an 'f' with objects which are .⁸

In addition the linguistic tokens must participate in uniformities among themselves which parallel the uniformities which hold among the extra-linguistic objects with which they are correlated. Picturing, then, is a complex relation of two relational structures.

Gilbert Harman, in his review of Science and Metaphysics,⁹ believes that Sellars has confused a theory of evidence with a theory of truth. As he points out, the two are not identical; something can be true without being evident; something can be evident without being true. Harman correctly sees that Sellars wishes to identify the semantical rules of meaning of a language with its rules of truth. But Harman believes that some of the rules Sellars mentions are actually rules of evidence. The rules in question are (as Harman calls them) the semantical rules connecting observation and thought:

Everything depends on the exact nature of the rules connecting observation and thought. Evidential rules would correlate stimulations (or how it looks to one) with specific thoughts. Truth rules might correlate the actual (and not just apparent) observation of something with a thought of that thing.¹⁰

The problem arises because of Sellars' logical atomism--atomic statements picture the world, molecular statements do not correspond directly with the world in the Tarski-Carnap sense--

and because at the same time he asserts that the scientific image can, in principle, replace the manifest image as our accepted view of the world.

To identify the rules of truth with those of the language game of thought is...to identify rules which indicate how the truth of certain representations depends on whether they correspond to or picture the facts with rules that permit one to have certain thoughts given certain observations.¹¹

Otherwise of course the latter rule will just be a principle of evidence. Harman takes this to imply that "picturing must always be associated with observationality," a principle Sellars denies since singular microphysical statements picture microphysical events but are not involved in language entry transitions. Harman quotes Sellars' reply to this sort of objection:

This objection assumes...that statements which are basic as the constituents of pictures must also be epistemically basic in the sense that they formulate observable states of affairs. It is, indeed, true of the common-sense framework that statements which are basic in one sense are also basic in the other. Yet the two senses of "basic" are different, and a transcendental philosophy which rises to a level of abstraction which distinguishes the generic character of epistemic concepts (e.g. language entry transition, conceptual picture, object) from the specific forms they take in common-sense discourse will not assume that the basic constituents of conceptual pictures must be statements of the kind which occur as conceptual responses to sensory stimulation.

Harman concludes that this shows that

a transcendental philosophy must not identify the rules of truth with the rules of the language game of thought, if these latter rules must include rules for "language entry transitions" expressing appropriate responses to observation.¹²

Harman's position consists of the following propositions:

- (1) Picturing is not always associated with observationality.
- (2) If rules of meaning are to be identified with rules of truth, "language entry transitions" would represent picturing rather than responses to observation.
- (3) The rules of the (language) game that determine which responses are appropriate are epistemic rather than truth rules.

All three are false because Harman has misinterpreted Sellars at some key points. Harman asserts (1) on the ground that singular microphysical statements cannot be language entry transitions. However, this does not rule out that language entry transitions are associated with picturing. The transitions involved "would have to be complex, and enormously so."¹³

But the fact that molecular statements do not, as such, form pictures, does not mean they are not associated with picturing, for Sellars is quite clear that molecular statements "pick out sets of pictures...and are true if the set of pictures they pick out includes the correct picture...."¹⁴ Thus language entry transitions of the sort in question can picture, although indirectly. That this is what Sellars has in mind is indicated when he says "it makes...sense to speak of basic singular statements in the framework of micro-physical pictures, according to a complicated manner of projection, of microphysical objects...."¹⁵

The manner of projection is complicated precisely because the language entry transitions involve highly complex responses which, nonetheless, pick out sets of pictures of the world. Thus it is false to say picturing is not associated with language entry transitions, although obviously the association is not

as direct as in the manifest image.

But I think Harman's error in this was caused by a radical misinterpretation of language entry transitions as responses to observation. Earlier, we saw, Harman distinguished two ways language entry rules could be interpreted; the second of these, appropriate to a theory of truth rather than evidence, could be taken as a formulation of Sellars' view of language entry rules: responses to observation in the weak sense of "response to objects noticed." But given the contrast he is drawing in (2), it is obvious he takes observation to be something else. Harman seems to be construing observation in precisely the sense Sellars emphatically rejects--as a sort of cognitive but non-linguistic awareness that something is the case. This becomes clear at the end of Harman's paper where he says:

(Sellars) could say that semantical rules... are truth rules. Then truth would be semantic assertibility; but "language entry transitions" would represent picturing rather than responses to observation. Furthermore, it is somewhat less clear in what sense the semantical rules... would be rules a person follows, and can be observed to follow, in playing the language game of thought; for this second sort of rule is a rule of the game only in the sense that it can be associated with the game.¹⁶

Sellars is not as careful in Science and Metaphysics as he is elsewhere to distinguish these two sorts of rules, but it is clear in "Some Reflections on Language Games" that initially at least, linguistic rules are not followed. Rather they are

norms, ought-to-be's, and it is the job of our language teachers to teach us to conform to them, although we have no notion of conforming to rules. This is especially true in the case of language entry rules for, on Sellars' account, we have no pre-conceptual awareness of extralinguistic objects, and could in no sense follow a rule in responding to sensory stimulation. On the other hand, as conceptually aware language users, we can train ourselves to respond differently to stimulation--that is, change our language by changing its rules--and this is the move envisaged in adopting the language of microphysics. It is in this way that our language, indirectly and by "a complicated manner of projection" involving observationality, will be able to picture the singular statements of microphysics.

Sentence (2) then is false because although "language entry transitions" do not represent picturing, they are involved with it, and this is sufficient to reject Harman's objection. Furthermore, "language entry transitions" are responses to observation although only in a specific sense of "observation." Sentence (3) taken as a reference to Sellars' "language entry rules" is false; on the other hand, it can be interpreted as a true statement, thanks to the inclusion of the word "appropriate" which puts it in the framework of epistemic evaluation. Of these sorts of rules though, epistemic principles, Science and Metaphysics says little.

It is essential to an understanding of the role of the notion of a conceptual framework in human knowledge, to appreciate how the ostensive element is involved in the meaning of certain terms. In learning a language we first of all are trained to respond to, for example, green objects by uttering tokens of "this is green." However, on Sellars' account such utterances do not, by themselves, constitute having the concept green--a great deal else must be known as well.

Thus, all that the view I am defending requires is that no tokening by S now of 'This is green' is to count as 'expressing observational knowledge' unless it is also correct to say of S that he now knows the appropriate fact of the form X is a reliable symptom of Y, namely that (and again I oversimplify) utterances of 'This is green' are reliable indicators of the presence of green objects in standard conditions of perception. And while the correctness of this statement about Jones requires that Jones could now cite particular facts as evidence for the idea that these utterances are reliable indicators, it requires only that it is correct to say that Jones now knows, thus remembers,* that these particular facts did obtain. It does not require that...at the time these facts did obtain he then knew them to obtain....

*(Sellars' footnote added 1963) My thought was that one can have direct (non-inferential) knowledge of a past fact which one did not or even (as in the case envisaged) could not conceptualize at the time it was present.¹⁷

This, then, is what is behind Sellars' statement, quoted earlier that "a conceptual framework is a whole which, however sketchy, is prior to its parts, and cannot be construed as a coming together of parts which are already conceptual in character." There is no knowledge independent of a conceptual

framework. The manifest image is that framework in which,
historically, man has known himself, and the world.

FOOTNOTES

¹Wilfrid Sellars, Science, Perception and Reality (London: Routledge and Kegan Paul, 1963), p. 6.

²Wilfrid Sellars, Science and Metaphysics: Variations on Kantian Themes (London: Routledge and Kegan Paul, 1968), p. 114.

³Ibid., p. 101.

⁴The discussion of the last few pages is drawn from Sellars (1968), pp. 82-87.

⁵Ibid., pp. 119-120.

⁶Ibid., p. 136, para. 57.

⁷Ibid., p. 137, para. 59.

⁸Ibid., p. 136.

⁹Gilbert Harman, "Sellars' Semantics," Philosophical Review 79 (1970), 404-419.

¹⁰Ibid., pp. 410-411.

¹¹Ibid., p. 417.

¹²Ibid., p. 418.

¹³Sellars (1968), p. 147.

¹⁴Ibid., p. 119.

¹⁵Ibid., p. 145.

¹⁶Harman, p. 419.

¹⁷Sellars (1963), p. 169.

Chapter 3: The Manifest Image

The manifest image is the framework in terms of which we see the world. It is not a static framework but one that has evolved as man's knowledge has grown. Although it has no absolute validity, it has a certain pragmatic warrant, as the framework which has enabled man to find his way about in the world with a certain degree of efficacy. The manifest image is distinguished from the scientific image chiefly by means of its basic categories--in the manifest image the elements which are basic in the epistemic sense are also the basic constituents of the conceptual pictures of the world whereas in the scientific image, the two do not coincide.¹ The two images are not independent despite the fact that each claims to be a complete picture of the world. In the first place the scientific image is still in a rudimentary stage of development and does not present a unified coherent view of the world. Methodologically it is parasitical on the manifest image. However, the manifest image, although providing a coherent and unified world view, remains sketchy, and tends to rely on the developing scientific image for its own development.

Now the difficulty is this: on one hand, we have seen, the manifest image lacks the absolute validity which the empiricists who embraced "givenness" were prepared to grant it; on the other hand, however, it has some validity as the only coherent and

reasonably complete account we have of the world, an account on which the development of a more adequate account logically depends. How are we to envisage the replacement of one of these frameworks by the other, a replacement which preserves (and explains) the knowledge we have in the manifest image?

There are many ways of characterizing the manifest image. In a first approximation it is our common sense way of looking at the world as opposed to the view we get through theoretical science. In "Scientific Realism or Ironic Instrumentalism" Sellars argues against Feyerabend's irreverent attitude toward common sense. Feyerabend speaks of common sense as a theory, and as such, liable to replacement by other theories which better do the job of describing the world, or in Feyerabend's pragmatic account of language, allow us to more effectively orient ourselves with respect to our sense impressions. Sellars pointed out however, that if there is a sense in which common sense is a theory on a par with other theories, this is not the classical notion of a theory. In the classical account of theories, we distinguish between internal and external subject matter. In the kinetic theory of gases, for example, "molecules and their behavior would be the 'internal' subject matter of the theory, and gases as empirical constructs defined without reference to molecules its 'external' subject matter."² If common sense is a theory, it is not a theory of anything, in this sense, for it is false to say it is a theory of sense impressions, for neither Sellars nor Feyerabend accept accounts of experience according

to which sense impressions are the primary objects of knowledge.

In another sense, however, we can speak of common sense as a theory with respect to sense impressions. This is the sense Feyerabend intends. The framework of physical objects provides the terms in which we respond to sense impressions, and this is a language incommensurable with that which we use to describe sense impressions. There is no reason, Feyerabend thinks, why scientists should not train themselves to respond to sense impressions with yet another incommensurable language, that of microphysical theory.

This view results from a rejection of the empirical notion of givenness: "that empirical knowledge rests on an absolute foundation of knowledge pertaining to sense data, and that the content of genuine descriptive concepts is derived from sense data."³

From the pragmatic picture of language as a technique of behavioral orientation, the conclusion might be drawn that we are free to replace, segment by segment, the framework in terms of which we perceive the world, by scientifically contrived structures which enable a more subtle orientation and fewer surprises. Thus, one who is at home in the microphysics of his day would be free to train himself to respond to his environment in terms which, though they externally resemble the vocabulary of his fellow man, have as their descriptive conceptual content highly derived constructs in this theoretical framework. We are free, so to speak, to pour new conceptual content into old bottles.⁴

In this view predicates derive meaningfulness not by being

defined in terms of a "given" observation framework. An observation framework is simply the one that the relevant language speakers use to orient themselves with respect to their sense impressions. A predicate gains meaningfulness by having a place in one of these behaviour orienting conceptual frameworks.

Now, of course Sellars agrees that it is in principle possible to replace our common sense predicates with scientific ones for he accepts the following propositions:

- (α) Micro-physical entities do not have the second class existence of mere "conceptual devices."
- (β) The framework of common sense is radically false (i.e., there really are no such things as the physical objects and processes of the common sense framework).
- (γ) Proposition (α) (β) are to be clarified in terms of the concept of its being reasonable at some stage to abandon the framework of common sense and use only the framework of theoretical science, suitably enriched by the dimension of practical discourse.

However, Sellars does not think this replacement is possible now. Although he rejects the given, he accepts the notion of a rock bottom level of observation predicates. These concepts are the primitives of the manifest image. Any methodological gain which would result in trying to incorporate theoretical predicates piecemeal into our observation base can be achieved by maintaining the classical account of theories with correspondence rules correlating theoretical constructs with rock bottom observation predicates. An attempt to abandon the common sense framework, before a total structure is available to

do the job better, would result in "serious methodological and conceptual loss."

In characterizing the framework of common sense, Sellars distinguishes between common sense beliefs and common sense principles. The former are not binding on the scientist.

However

...not all subject-matter dependent universal propositions to which common sense is committed are properly characterized as beliefs. There are many principles about physical objects and the perception of them ("categorical principles" they might be called) which are constitutive of the very concepts in terms of which we experience the world.⁶

It is the rock bottom concepts and principles of common sense which are binding until a total structure which can do the job better is actually at hand--rather than a "regulative ideal."⁷

Why does Sellars hold that we must retain the common sense or manifest framework until a complete alternative is at hand? In other words what is this conceptual and methodological loss? The problem is, of course, what to do with those sensuous qualities of the physical things characteristic of the manifest image. In the first place Sellars argues that there is no alternative in the manifest image to conceiving physical objects as occurrently coloured, for physical objects must have "content qualities." It is not plausible to suggest that we primarily conceive of sense impressions as coloured, and physical objects as analogically coloured, for our conception of sense impressions is derivative on that of physical objects. But on the other hand, the colour of physical objects is ontologically grounded

in that of sense impressions--as Sellars puts it: "visual impressions are prior in the order of being to concepts pertaining to physical color, whereas the latter are prior in the order of knowing to concepts pertaining to visual impressions."⁸

Sellars illustrates the "methodological and conceptual loss" in the following two points:

1. The abandonment by scientists of the conceptual framework of common sense physical objects would involve either the abandonment of the conceptual space of color tout court, or the retention of this conceptual space as it reappears in its analogical offshoot, the conceptual space of sense impressions. The latter would be cut off from its foundation and left to wither on the vine. In either case, the conceptual space of the qualities of sense ("secondary qualities" in one use of this phrase) would disappear from the public observation base of science. It would enter science only in linguistics, in the study of the structure of the language of non-scientists --and of scientists only to the extent that their sense impression talk continued to reflect the pre-revolutionary framework of common sense physical objects.
2. Only when the conceptual space of sense impressions has acquired a status which is not parasitical on the framework of common sense physical objects--in other words, only with the development of an adequate scientific theory of the sensory capacities of the central nervous system--could the framework of common sense be abandoned without losing conceptual contact with a key dimension of the world.⁹

Thus Sellars believes that if we were to attempt a piecemeal replacement of the manifest image--the replacement of the framework of manifest physical objects--by a fragment of the developing scientific image, the logical space of colour would be lost except as a relic of bygone times. I wish to take issue with this claim, or at least with its use as an argument against

such piecemeal replacement. I disagree that "serious methodological and conceptual loss" necessarily need result. Could we not give an indirect realist account of the manifest image in which colour terms were acknowledged to apply primarily to sense impressions, i.e. sense impressions were both epistemically and ontologically prior to physical objects which were postulated as being analogous to sense impressions? Even if this could not be regarded as an analysis of our common sense framework, could it be a possible categorical refinement of it? In other words, could a more sophisticated version of the manifest image emerge, evolved in response to the pressure of scientific development, although distinguished from the scientific framework as is the (Sellarsian) manifest image? The model for such a development could be the categorical evolution of what Sellars terms the "original image" which, one could argue, evolved in response to primitive science. If such an account were possible, the way would be opened for a piecemeal replacement of the manifest image without risking the loss of the logical space of colours.

Sellars, of course, rejects this possibility:

As Berkley, Kant, and Whitehead, among others, have pointed out, physical objects cannot have primary qualities only--for structural and mathematical properties presuppose what might be called "content qualities." And unless one falls into the trap of thinking of the framework of physical objects as a common sense theory evolved with unconscious wisdom to explain the manner in which sense data occur, it will scarcely do to say that the content qualities of physical objects are conceived, by a common sense use of analogy, to be the physical counterparts of the qualities of data (i.e. to play in the realm of

physical things the content-role played in sense data by sense qualities). For if the conceptual space of common sense physical objects is underived, their content qualities must be directly rather than analogically conceived, for it is only in terms of perceived, and therefore conceptualized, qualitative difference that form and structure can be distinguished.

Thus the rejection of a phenomenistic analysis of the framework of common sense requires that the physical objects of this framework have perceptible qualitative content. And once one realizes this, one sees that there is no alternative to construing these physical objects as colored in the literal occurrent sense. One might wish to say that this framework--which has as its central constituents items which are in this sense colored through and through--is, from the standpoint of theoretically-oriented science, false, although enabling a behavioral adjustment of sufficient accuracy for the practical purposes of life. But false or not, such is the framework of common sense.¹⁰

There are two elements to this argument--1) that physical objects must have qualitative content and 2) that if a phenomenistic (or indirect realist) analysis is rejected, this qualitative content must be colour. But granted Sellars may be correct in his analysis of our common sense framework as being direct realist and, therefore, that in this particular framework there is no alternative to construing physical objects as occurrently coloured, surely an argument is needed to the effect that we could not function in the world if we had an indirect realist conceptual framework. That is, a stronger argument is required against the possibility of a categorical refinement of the manifest image along the lines we are considering.

Sellars does indeed provide such an argument against phenomenism in general. He argues in "Phenomenism" that classical

phenomenalism cannot get off the ground because any attempt to analyse our knowledge in terms of sense contents fails. Our knowledge is constituted of uniformities which cannot be formulated in terms of sense contents alone--the framework of things is presupposed:

...the very principles in terms of which the uniformities are selected carry with them the knowledge that these uniformities are dependent uniformities which will continue only as long as these particular objects constitute one's environment, and hence preclude the credibility of the generalization in sense content terms which abstract consideration might lead us to think of as instantially confirmed by the past uniformities.

The fact that the noticing of complex uniformities within the course of one's sense history presupposes the conceptual picture of oneself as a person having a body and living in a particular environment of physical things will turn out, at a later stage of the argument, to be but a special case of the logical dependence of the framework of private sense contents on the public, inter-subjective, logical space of persons and physical things.¹¹

The argument that physical objects are occurrently coloured is taken to be the general case:

The point I have in mind is essentially the same as that on which our critique of classical phenomenalism was based. For to suppose that the qualities of physical things are powers is to overlook the fact that the occurrent properties of physical objects are presupposed by the laws which authorize both the ascription to 'circumstances' of powers to manifest themselves in the sense contents of percipients (stressed by power phenomenalism) and the assertion of subjunctive conditionals about the sense contents which would eventuate for a perceiver were such and such (phenomenal) conditions to be satisfied (stressed by classical phenomenalism).¹²

But again I think this argument will not rule out the possibility of the sort of refinement I am proposing. It seems to rely on the claim (which we have accepted) that things must have content properties, and assumes that colours are the only candidates for the content properties of physical objects. The argument fails because it is directed against analysis of our common sense framework which, I concede, may be as Sellars claims. As such the claim that the content properties of physical objects are colours is merely a contingent one, pertaining only to our actual common sense framework. This fact about the manifest image, if it is a fact, is part of an explanation by Sellars of how we come to have knowledge--how we form empirical generalizations and concepts by analogy. How we come to have knowledge is both a matter of history, and a matter of science. From the historical point of view, the fact that man has gained knowledge by means of a false framework is inconsequential--his knowledge is nonetheless knowledge and has allowed him (or perhaps, will allow him, some day) to explain how it is possible to gain actual knowledge in a false framework. However, a scientific explanation cannot have man's acquisition of knowledge based necessarily on a false conceptual framework--at least not one as radically false as Sellars' manifest image turns out to be. Indeed Sellars exaggerates if not the degree to which it is false, then the degree to which we are capable of rejecting it. Sellars claims that the common sense framework of physical objects will be rejected, that the objects will be found not

to exist. The truth is, of course, that the framework is merely replaced by a ~~better~~ one, and that the physical objects of the manifest image are merely conceived of differently, hopefully more adequately.

The point of a more adequate successor framework is that it more perspicuously pictures the way things really are. It is clear that in the scientific image physical objects, construed, conceived and perceived as systems of individually imperceptible particles, will not be occurrently coloured. Yet, on the other hand, it seems clear that Sellars envisages perceivers as still perceiving via coloured sense impressions, scientifically construed. My argument is simply this: if it is possible to give a scientific explanation of concept formation and knowledge acquisition as it will take place in the scientific image, apparently without occurrently coloured physical objects (i.e. systems of imperceptible particles which lack occurrent perceptible qualities), surely a similar account could be given in a derivative of the manifest image in which physical objects have occurrent properties but only analogically conceived. My view is that either Sellars must admit that a scientific explanation of how we obtain knowledge using the common sense framework need not necessarily analyse physical objects in this framework as occurrently coloured; or he must give up his account of the in principle possible replacement of the manifest framework by the scientific.

But to push this a bit further it is necessary to imagine

what Sellars might reply to the objection. What distinguishes, in his eyes, the explanation of concept formation in the two cases, so that one is acceptable only if physical objects are occurrently coloured, whereas the other is acceptable without this condition? The key to the answer is the passage quoted above (p. 57) in which Sellars describes the "methodological and conceptual loss" that would result if piecemeal replacement of the manifest image were attempted. Sellars believes that we would lose contact with the conceptual space of colour and, in effect, lose colour as a dimension of the world. Now obviously it would be undesirable to do so since colour plays such an important role in any explanation of concept formation. Nonetheless it is difficult to see why the move envisaged, would necessarily cut the conceptual space of sense impressions off from its roots. Although I admit that in certain historical stages of the manifest image (including perhaps, the present) sense impressions may have been conceived as analogous to physical objects, why could we not train ourselves to consider sense impressions as primarily coloured and physical objects analogously coloured? The situation I have in mind is one in which we have replaced the framework of manifest physical objects by their scientific counterparts although a scientific account of perception e.g. a neurophysiological account was not yet available. I simply disagree with Sellars' view that in such a situation the logical space of colours would wither on the vine, whereas if a complete neurophysiological account

were also available, including *sensa*, we would retain contact with the logical space of colour. Although sense impressions are still unreduced, they are now prior in the order of conceiving as well as ontologically prior. In this framework, language learners are trained to attach colour terms to sense impressions rather than to objects, although objects may be coloured in a different, non-occurrent sense. The important point, however, is that as far as I can see there is no reason why colours of sense impressions should not retain all the features they had when, in the manifest image, they were analogically conceived. They would thus retain their ultimate homogeneity, although no one would suppose this homogeneity to be of a spatial sort. The argument for *sensa* would be no less viable in that situation than it is in the present. I fail to see what difference a scientific explanation or reduction of sense impressions would make in preserving this logical space. In other words, my view is that if Sellars thinks epistemological contact will be lost with colours in one case, he cannot avoid the same conclusion in the other, for in neither, on his construal, is a model available on which colours are patterned. I hold that in both cases the model for colours becomes sense impressions--that our language teachers now train us to describe sense impressions in terms that preserve the logical space of colour.

We seem to have exhausted the supply of reasons Sellars gives for rejecting a partial replacement of the manifest image.

Yet still I do not find any of the reasons individually or together, sufficient. I am forced then to speculate what may be the motives behind the position. Now for an account of concept formation Sellars requires that colours, suitably transposed, end up as occurrent properties of some aspect of human neurophysiology. Yet to be simply occurrent properties of sense impressions is not satisfactory. Partly of course, this may be because sense impressions, in any case, are states, not particulars. But what seems more important to Sellars is that sense impressions and, therefore, their colours are reducible. Sellars seems to be supposing at this point that pure occurrents must also be irreducible, and I believe I can explain why. When Sellars talks about replacing the manifest image, he means quite literally that we shall respond in the language of science in the most perspicuous way possible. The theoretical terms themselves would enter into the observation use of language. For example, instead of responding to red objects with "this is red" we would respond (to the effect): "this is a system of particles emitting such and such electromagnetic radiation." Evidently, on such an account, colour terms will not enter into the observation language. What about sense impression language? Would we say

I am having a sensation of a system of particles emitting such and such electromagnetic radiation.

or would we still say

I am having a red sensation?

Both seem unacceptable--the first, involves the loss of the conceptual space of colour. On the other hand, the second, which preserves it, is divorced entirely from any use it should have as a non-endorsing observation report, any formal connection with the object which it is "about." Sellars seems to think the only solution is to require a completed scientific image in which sense impressions, like physical objects appear in their most perspicuous form i.e. in the terms of neurophysiology. Since the requirements still stand that (a) colours appear as pure occurrents, and (b) sense impressions appear in their most perspicuous, i.e. most reduced form, colours must be primitive properties of basic particulars. On this view, the logical space of colour appears intact in sense impression discourse, and the logical connection of sense impression discourse with physical object discourse is maintained in the complicated sense that both participate in the same theoretical/explanatory/causal framework.

I have two objections to this solution. First, it presupposes, rather than provides, a cogent argument to the effect that the particulars which contribute to a reductive explanation of sense impressions include *sensa*. In other words, the present discussion concerns not the argument for (the primitiveness of) *sensa* itself, but a possible motive for it. In the second place, I believe it makes a false assumption about the use of language. This is the point I made earlier about supposing that purely occurrent properties were necessarily primitive properties.

This is the logical result of the notion of an ideally perspicuous language. It suggests that all defined terms are eliminated from used language--replaced, in effect by their ideally perspicuous and complete definitions. Nonetheless, of course, the idea is absurd. It is possible that there is no limit to the complexity that might ultimately be involved. Even the absolute primitives of today's science will end up as defined in tomorrow's. The primitives of neurophysiology, neurons, are in the context even of biology, impossibly complex. Thus there seems very little point to suggest that we should replace any of our common sense terms by theoretical terms, since it seems virtually inconceivable that the job could ever be completed.

It is obvious then that this could not be what Sellars means in the passage quoted on page 55 above. The reality of the scientific framework does not depend on the possibility of purging all defined terms from language. And indeed the question is discussed in "The Language of Theories," although left finally unresolved:

...might (a correspondence rule) not be construed as a redefinition of observation terms? Such a redefinition would, of course, be a dead letter unless it were actually carried out in linguistic practice. And it is clear that to be fully carried out in any interesting sense, it would not be enough that sign designs which play the role of observation terms be borrowed for use in the theoretical language as the defined equivalents of theoretical expressions. For this would simply amount to making these sign designs ambiguous. In their new use they would no longer be observation terms. The force of the "redefinition" must be such as to demand not only that the observation-sign design correlated with a given theoretical

expression (be) syntactically interchangeable with the latter, but that the latter be given the perceptual or observational role of the former so that the two expressions become synonymous by mutual readjustment. And to this there is an obvious objection: the meaningful use of theories simply does not require this usurpation of the observation role by theoretical expressions. Correspondence rules thus understood would remain dead letters.

But if the above conception of correspondence rules as 'redefinitions' will not do as it stands, it is nevertheless in the neighbourhood of the truth; for if correspondence rules cannot be regarded as implemented redefinitions, can they not be regarded as statements to the effect that certain redefinitions of observation terms would be in principle acceptable.¹³

Thus there ultimately is no objection to keeping our language as close to its current syntactical form as possible, although accepting new definitions for certain terms. The real catch remains, of course, the colours of physical objects. Now even on Sellars account the "red" of a red object is in a certain sense ambiguous: for while it may mean occurrently red, it also means "power to cause a red sense impression." Thus "red" already has a degree of ambiguity; there is no reason why in the scientific image, objects could not still be "red," now meaning "power to cause red sense impression," with both "red object" and "red sense impression" reducible, or defined expressions. This would pose an alternative to the dilemma posed above of retaining a logical connection between objects and sense impression, on one hand, and also retaining the logical space of colour. Assuming critics of the argument for (the primitiveness of) ~~sense~~ are right and Sellars wrong

is the logical space of colour preserved in a way that is acceptable for an explanation of concept formation?

The problem finally, which may be what is troubling Sellars, is this:

We admit the necessity of the logical space of colour in an explanation of human knowledge. Is it satisfactory that this logical space be defined, assuming it to be possible, in terms of primitives to none of which is the logical space of colour applicable? Now we allowed that the scientific image was the "real" view of the world on the condition that it be in principle (although perhaps not practically) possible to better describe and explain the world in terms of its primitives. We allow, however, for strictly practical reasons, the use of defined concepts in the scientific language game. Yet these defined objects and attributes have no "absolute reality" being nothing but other objects and attributes. Thus in the final analysis, colour terms, although they persist in our used language, are not "real" aspects of the world.

In other words, unless colours turn out to be primitive properties of basic particulars, there is a strong sense in which they are not a real part of the world. They exist only insofar as it is practical to use such predicates. A scientific explanation of concept formation would have to use the vocabulary of colour only because it was in fact, but not necessarily, used in concept formation. But if it were possible in principle to replace the use of the vocabulary of colour in actual concept formation i.e. in teaching language to children, as we have seemed to concede, we would be left with no way to explain how any of these concepts are formed. Colour, as the

basic epistemological unit of concept formation, would have been reduced right out of existence.

The answer to this, I think, lies in the vagueness of the contrast between "in principle" and "in practice" possibility. The sense in which it is in principle possible to replace the manifest framework with the scientific is a very limited sense --meaning, in effect, that it is possible to describe using only theoretical primitives at least as much as can be described in manifest image primitives. But if it is in practice impossible to completely carry out this replacement, as I claim, because perhaps of the cognitive limitations of the human brain, then it is in principle impossible to carry out the replacement in practice. Colours remain a necessary feature of the world in an epistemological sense and for epistemological reasons, but not in an ontological sense. This claim is compatible with the claim that colours are ontologically reducible.

In this chapter I have been trying to refute Sellars' contention that partial replacement of the manifest by the scientific is not acceptable. Yet many philosophers, regarding our ability apparently to incorporate the language of micro-physics into practice have taken this to indicate precisely that partial replacement is possible. Although I am convinced that the argument for sensa is independent of this issue, it is important for two reasons. In the first place, I think the preceding discussion reveals a fundamental difference between Sellars and myself regarding colours. Whereas we agree that

they are important and must be preserved, we disagree on how they must appear in the scientific image. For Sellars, it is not sufficient that they survive in our scientific conceptual scheme merely as epistemological primitives. They must appear as ontological primitives. My view is that it is sufficient that they appear as epistemological primitives; if science tells us that in the final analysis, they do not exist, we must accept the conclusion. Sellars cannot accept a transitory state in which no thing is actually coloured, since this would amount to a loss in descriptive power and decreased adequacy. My view is that since it does not matter that colours turn out to be ultimate constituents of the world, a situation is acceptable in which they appear as epistemologically basic, although not predicable of any existing objects. The considerations of the present chapter have been largely epistemological and therefore, independent of the ontological argument for *sensa*.

The second reason why the issue is important is because I think an account which envisages an intermediate conceptual framework between Sellars' manifest and the scientific is more in line with the style of argument that was discussed in Chapter One. At that time I mentioned that Sellars' method of justification of his positions was indirect. He points to a certain model and attempts to project the principles of the model to new situations. But there is no model for the sort of framework replacement he has in mind. The problem is that theoretical

explanation has created a unique relation between two conceptual frameworks, and there are no historical examples which enable us to understand it. Yet I think there are models which help.

My complaint with Sellars' conception of the replacement of the manifest by the scientific is that he seems to violate his own principle that categorical principles are binding on the scientist. (See above p. 56). It is certainly a categorical principle of Sellars' manifest image that physical objects are occurrently coloured. But in the succeeding scientific image, this principle is ignored, for the successors of physical objects are not occurrently coloured. However, there is a model for the sort of categorical change that seems to be involved here--namely the categorical refinement of the original image. I think the move from occurrently coloured physical objects to a phenomenalist/indirect realist framework could be assimilated to this model of conceptual evolution, while remaining within the limits of manifest-type, or common sensical frameworks. The possibility of such a shift has been one of the focal points of the present chapter.

The model for the second stage of the process, which involves the substitution of the common sense for the theoretical, is examples drawn from science which are less problematic than the case of neurophysiology. Thus we can project the principle drawn from the analysis of, for example, the replacement of the framework of chemical theory by that of physics. It is to this question that the following chapter is devoted.

FOOTNOTES

¹See above p. 45.

²Wilfrid Sellars, Philosophical Perspectives (Springfield, Illinois: Charles C. Thomas, Publisher, 1967), p. 339.

³Ibid., p. 351.

⁴Ibid., p. 352.

⁵Ibid., p. 354.

⁶Ibid., p. 338.

⁷Ibid., p. 355.

⁸Ibid., p. 357.

⁹Ibid., p. 358.

¹⁰Ibid., p. 356.

¹¹Wilfrid Sellars, Science, Perception and Reality (London: Routledge and Kegan Paul, 1963), p. 84.

¹²Ibid., p. 98.

¹³Ibid., p. 125.

Chapter 4: Theoretical Explanation

As I interpret Sellars, our conceptual structure must be regarded as a transitional one between an imaginary situation where a fully elaborated manifest image provided all our knowledge of the world, and an ideal situation where the scientific image provides our ontological knowledge of the world. This transitory state is signaled by the existence of correspondence rules which correlate entities defined in theoretical frameworks with entities defined in the common sense framework. I turn now to an examination of some of the details of Sellars' account of theoretical explanation, in particular the correspondence rules which signal a puzzling relationship between two conceptual frameworks.

The kind of theoretical explanation we are interested in is the sort that postulates imperceptible objects to explain the behaviour of observable objects. Sellars accepts the classical account of theories of this type (or, at least, something like it):

(This account) is built upon a distinction between: (a) the vocabulary, postulates, and theorems of the theory as an uninterpreted calculus; (b) the vocabulary and inductively testable statements of the observation framework; (c) the 'correspondence rules' which correlate, in a way which shows certain analogies to inference and certain analogies to translations, statements in the theoretical vocabulary with statements in the language of observation.¹

Each of these categories requires investigation.

(a) The Theoretical Language

The theoretical language contains, in addition to that part of vocabulary which ostensibly refers to unobserved entities and their properties, (a) logical and mathematical expressions which have their ordinary sense, and (b) the vocabulary of space and time. (Query: Can we say that the latter part of the theoretical vocabulary, too, has its ordinary sense? To use the material mode, are the space and time of kinetic theory the same as the space and time of the observable world, or do they merely 'correspond' to them? In relativity physics it is surely the latter.)²

Sellars takes theoretical expressions to be factually meaningful--their meaning being determined by the role they play in the deductive apparatus of the theory. Theoretical terms are implicitly defined by the postulates in which they occur. Although it is ideally possible to give an exhaustive list of the postulates of a theory, in practice this is not the case because the conceptual texture of theoretical terms in scientific use is far richer and more finely grained than the texture generated by the explicitly listed postulates. This lack of detail is compensated by the use of models and analogies which convey features of the objects which are not captured by the postulates. A domain of familiar objects is pointed to that share certain similar features with the objects of the theory. The similarity in question may be similarity of particulars or similarity of attributes, or what may be called first-order and second-order similarity of particulars. An example of first order

similarity is one in which two particulars share an identical attribute. In second-order similarity, they share a similar attribute. Similar attributes are those which share identical higher order properties. The conceptual framework of a theory is generated by specifying the analogies which are to obtain among the objects of the theory and those of the model. Models are used in theory construction to specify new attributes as the attributes which share certain higher order attributes with attributes belonging to the model, fail to share certain others, and which satisfy the conditions laid down by the relevant correspondence rules.

Clearly, in order for this to work, the attributes, first and higher order, which are specified to be identical, must be purely formal attributes of the two frameworks (that of the theory and that of the model) for they are not merely counterparts but identical, and must be independent of the objects of the frameworks. The example Sellars gives is the similarity between points on a line and moments in a temporal series. Although they share no first order properties they do share second order properties; for example, their respective relations "less than" and "before" are transitive.

According to Sellars, sense impressions, (and thoughts) are conceived in the manifest image as theoretical. Clearly he is using this term in a somewhat informal sense--obviously no attempt is generally made to specify the postulates which implicitly define the logical space of (coloured) sense impressions.

This is a case where the specification is purely by analogy-- the model being, of course, the perceptible colours and shapes of physical objects. Thus the colours of sense impressions and the colours of physical objects are attributes which are similar--they share certain higher order properties but not others. Whereas for the purpose of everyday use this informal specification is entirely adequate, a scientific account must attempt a more careful, more specific characterisation of the similarities and differences that must hold.

Sense impressions are postulated to explain how it is we form conceptual representations of the world around us; thus, the colours of sense impressions must be at least sufficiently similar to the colours of physical objects to explain these conceptual representations--in other words, whatever information about the world physical colours can express or contain, must also be expressible in the logical space of the colours of sense impressions. Hooker suggests this includes

determinate/determinable structure, colour compatibility...as well as...the other logical characteristics of the concepts, for example, simplicity and homogeneity.

The first group are formal second order properties of colour which must be reconstructed exactly in the logical space of the colours of sense impressions. The second group, however, are properties of a different sort. Homogeneity is not a formal second order property for it involves the concept of space--indeed, homogeneity is rather complex involving, as it does,

both colour and space. Nor does the homogeneity of colours by itself contain or express information, or increase the information capacity of the logical space of colour. Nonetheless, it is an essential feature of colour since colours cannot be conceived which are not ultimately homogeneous. I think homogeneity should be considered a rather complex type of property which, like colour, has a logical space which can be reconstructed from different "materials." Anything that can be said to have parts can plausibly be said to be or not to be homogeneous. Yet there are different ways things can be said to have parts, and to each of these different possibilities corresponds a different type of homogeneity. To take the simplest possible example--space consists of spatial parts all of which are also space; space is spatially ultimately homogeneous; time, on the other hand, cannot be said to have spatial parts; rather it has temporal parts, and is in that sense homogeneous. Thus homogeneity is a versatile concept which possesses (and is limited by) a certain logical space which can be specified in purely formal language. In the case of ultimate homogeneity the appropriate model would be the mathematics of the continuum.

In the case of a scientific reduction of colours, my view is that the counterparts of manifest colours, i.e. scientific colours, need only be required to be homogeneous in a similar or analogous sense to the homogeneity of manifest colour. This is obviously a weaker requirement than that scientific colours be homogeneous in the identical sense of the homogeneity of

manifest colour. This would mean that scientific homogeneity of colour would have to share certain higher order formal properties with manifest homogeneity, but not others. To return to the example of space and time: obviously both senses of homogeneity will share certain features of the mathematical theory of continua; however, the theory which specifies the spatial concept of homogeneity will be richer in a sense in that it includes three dimensional continua, whereas the theory specifying the logical space of temporal homogeneity will only include one dimensional continua. Certain properties will be applicable to spatial homogeneity which are not applicable to temporal homogeneity.

Simplicity, or primitiveness, is again different. If it were thought that simplicity were a part of the logical space of colour, which had to be reproduced in the logical space of sense impressions, the argument for *sensa* would be straightforward. However, the matter is not so simple. Simplicity is not a property of anything or any property in any usual sense. It does not seem that the simplicity of colours plays an essential (or any) role in the information capacity of the logical space of colour. Simplicity is what we might call a metalinguistic property--that manifest colours are simple is more a feature of the conceptual framework than of the colours themselves. As such, it is the one property which is in no sense binding when one framework succeeds another. Indeed the fact that the set of primitive properties changes is a sign of a change of framework.

(b) The Non Theoretical Language

The non-theoretical language with which a given theory is connected by means of correspondence rules may itself be a theory with respect to some other framework, in which case it is non-theoretical only in a relative sense. This calls up a picture of levels of theory and suggests that there is a level which can be called non-theoretical in an absolute sense. Let us assume for the moment that there is such a level and that it is the level of the observable things and properties of the everyday world and of the constructs which can be explicitly defined in terms of them. If following Carnap we call the language appropriate to this level the physical-thing language, then the above assumption can be formulated as the thesis that the physical-thing language is a non-theoretical language in an absolute sense. The task of theory is then construed to be that of explaining inductively testable generalizations formulated in the physical-thing language, which task is equated with deriving the latter from the theory by means of the correspondence rules.⁴

We have already seen, however, that Sellars rejects the notion that the physical-thing framework is an absolute level. Sellars locates the plausibility of the levels picture in an overly simplistic account of explanation. The account Sellars rejects claims that singular matters of fact are explained only by empirical generalisations, which in turn are explained by being derived from theories. This view automatically relegates theories to a second class status, in principle dispensable.

For to suppose that particular observable matters of fact are the proper explananda of inductive generalizations in the observation framework and of these only, is to suppose that, even though theoretical considerations may lead us to formulate new hypotheses in the observational framework for inductive testing and may lead us to modify, subject to inductive

confirmation, such generalizations as have already received inductive support, the conceptual framework of the observation level is autonomous and immune from theoretical criticism.

The truth of the matter is that the idea that microtheories are designed to explain empirical laws and explain observational matters of fact only in the derivative sense that they explain explainers of the latter rests on the confusion between explanation and derivation. To avoid this confusion is to see that theories about observable things do not 'explain' empirical laws in the manner described, they explain empirical laws by explaining why observable things obey to the extent they do, these empirical laws;...

Furthermore, theories not only explain why observable things obey certain laws, they also explain why in certain respects their behaviour obeys no inductively confirmable generalization in the observation framework.⁵

(c) Correspondence Rules

Correspondence rules typically connect defined expressions in the theoretical language with definable expressions in the language of observation.

In some sense correspondence rules specify identities, for the basic schema of (micro \rightarrow) theoretical explanation is,

Molar objects of such and such kinds obey (approximately) such and such inductive generalizations because they are configurations of such and such theoretical entities.⁷

But it is clear that this identity is not one of sense. The theoretical framework aspires to replace the observation framework. The observation framework is construed as a poor explanatory framework with a better one available to replace it.

...this is the sort of thing that is done when one theoretical framework is 'reduced' to another, and...the notion of the replaceability of a microframework by a micro-micro-framework is a reasonable explanation of the force of such a statement as

Ions behave as they do because they are such and such configurations of subatomic particles.⁸

Sellars believes that a similar account is possible in the case of the theoretical explanation of observables. Thus correspondence rules are interpreted as statements to the effect that certain redefinitions of the observation terms would be in principle acceptable.

Correspondence rules would appear in the material mode as statements to the effect that the objects of the observational framework do not really exist --there really are no such things. They envisage the abandonment of a sense and its denotation.⁹

Of course, the qualitative aspects of the world cannot be left out. They must appear in the scientific theory somehow:

...the sensible qualities of the common sense world, omitted by the physical theory of material things, might reappear in a new guise in the microtheory of sentient organisms. This claim would appear in the material mode as the claim that the sensible qualities of things really are a dimension of neural activity.¹⁰

This process, i.e. Sellars' account of theoretical explanation, is regarded as a replacement of one framework by another with the condition that the truths and concepts of the preceding framework be reproduced in the succeeding framework. Thus we come to the two principles which are so important to the argument for sense--framework adequacy, which specifies the relation

between counterpart concepts in the two frameworks, and reducibility, which applies to each framework individually. We saw in Chapter One that Hooker thought that the argument for *sensa* required a sort of ad hoc argument that the logical space of colours is closed. It is clear that the case of colours is somewhat different from other properties that are scientifically explained, and therefore reduced; nonetheless, we would expect that the principles of theoretical explanation should apply uniformly to all manifest image properties. Now there are two cases--that which involves observables, and that which doesn't. Consider first the latter case--the "reduction" of one theoretical framework by another: the explanation of the behaviour of ions in terms of subatomic particles. Now initially we have two theories belonging to two sciences--chemistry and physics. Ions are basic entities in chemical theory and subatomic particles are basic entities of particle physics. Now it is seen by scientists that certain systems of subatomic particles can be defined with identical properties to the ions of chemistry--this observation is expressed by correspondence rules which connect the appropriate defined expressions of particulate physics with the appropriate primitive but apparently definable expressions of chemistry. Now these correspondence rules cannot be considered to express identities of sense for the reason that the expressions of either side are defined in different frameworks. The ion expressions are

primitive terms of chemical theory; the appropriate particle physics terms are defined terms of that framework. Thus, they are not identified; the correspondence rules appear in the material mode as statements to the effect that the objects of the framework of chemical theory do not really exist--there really are no such things. What really exists are the defined objects of the physical framework. The framework of chemical theory is replaced by that of physical theory.

The replacement was possible because it was seen that the expressions of chemistry, though not actually explicitly defined terms, could be defined in the terms of physical theory. In other words, the logical spaces of the physical primitives, themselves implicitly defined by the postulates of the theory and by the use of model and analogy, were such that those primitives could be used to define states or entities which themselves possessed the logical space of the chemical primitives--except, of course, their primitiveness or simplicity. Thus in this case, the principle of framework adequacy requires that the successor framework contain defined states or entities which have all the properties of the predecessor framework except those related to primitiveness. Of course, the successor concepts of physics may be "richer" than their counterparts in chemistry--they may have properties which were not known or describable in the chemical theory. These properties may explain properties that were described in the chemical theory--such a situation is likely for, according to Sellars, such "enrichment" of concepts is what makes replacement of frameworks reasonable.

When it comes to explaining the behaviour of observables, there is an additional problem. Not only is there the problem that the physical objects of the common sense framework are basic entities, whereas the systems of objects of the scientific framework are defined, but there is also, of course, the problem of phenomenal properties which are not definable in the terms of the appropriate physical theory. Sellars' solution is again to say that the physical objects of the manifest image do not exist. This represents rather a departure from the previous account--the successors of manifest physical objects lack not only the primitiveness of their predecessors, but also their phenomenal occurrent properties. Indeed, they do have the corresponding causal properties, as do the predecessors. I suggest again, as I did in the previous chapter, that it is surely more reasonable to envisage the shift of the primary locus of colour from objects to perceivers as a move made in the manifest image, thus permitting a unified application of the principle of framework adequacy.

Either way, of course, the argument for *sensa* is not affected, for it arises when we try to put sense impressions into the scientific image. I shall save the details of this for the next chapter and limit myself to a few observations. Sense impressions are construed in the manifest image as adverbial states of persons. Thus the analogy with physical objects is transcategorical--between objects and states. Persons, like physical objects, are basic entities--single logical subjects--

in the manifest image. This fact, and Sellars' logical atomism (and related principle of reducibility) requires that the colours of (manifest image) sense impressions be primitive. My point is that it is not a feature of the analogy which requires the colours of sense impressions to be primitive, but the fact that they are properties of states of basic entities of the manifest image.

Now the scientific counterparts of manifest sense impressions are again adverbial states, but now of systems of particles. Scientific sense impressions are not primitive adverbial states— simply because no single logical subject in the scientific image has them. Nonetheless, Sellars claims that the colours of these sense impressions are primitive. The only point I wish to make at this time is that, if we apply the principle of framework adequacy uniformly, it does not require that colours be primitive. As we saw, the application of the principle requires that the successor framework, in this case, the scientific image (specifically, neurophysiology), reconstruct from its own primitives all the properties of manifest colours except their primitiveness. Thus the argument for the primitiveness of colours, the basicness of *sensa*, requires some additional premises.

A complication arises, however. In the earlier case of the reduction of one theory by another, a key condition of the reduction was that the primitives of the chemical theory be definable in terms of the primitiveness of the physical theory.

There is a certain dimension of free play involved here because of the element of open texture and vagueness associated with the specification of the primitives of both the physical theory and the chemical theory. Indeed the subsequent identification contributed to specifying the logical spaces of the concepts. In the case of colours, the case is exactly parallel. In fact to the degree that the sciences of neurophysiology and human behaviour are less developed than physics and chemistry, the degree of open texture and vagueness are increased. On one hand, on the side of manifest colours, the necessary properties which must be reconstructed, or defined by neurophysiological entities is very vague--specified only by analogy. On the other, the definable states of neurophysiology, modeled largely, we may suppose, on computer science, are still only roughly and very incompletely known.

Despite this vagueness, however, I pointed out earlier in this chapter that among the logical space of colour which must be reconstructed from the primitives of neurophysiology, is some analogue of the spatial homogeneity of physical colour. This requirement leaves open the possibility that the homogeneity, which involves the notion of "consisting of parts which no matter how small, are of the same kind" be either spatial or temporal. Now since it is fairly clear that no appropriate spatial homogeneity is definable among the possible states of presently conceived neurophysiological entities, e.g. neurons,

we are left to attempt to define appropriate states which are homogeneous in a temporal dimension. That Sellars allows that *sensa* are homogeneous in this dimension is clear, for he tells us that "the qualities of sense are a dimension of natural process"¹¹ and speaks of "the sequence of 'events' which is the *sensum*."¹² On this point, it seems, I am in agreement with Sellars; where I do not agree with him is his claim that such *sensa* are basic particulars which preserve the primitiveness of colour predicates. It is to his arguments for this that I now turn.

FOOTNOTES

¹Wilfrid Sellars, Science, Perception and Reality (London: Routledge and Kegan Paul, 1963), p. 107.

²Ibid.

³C. A. Hooker, "Sellars' Argument for the Inevitability of the Secondary Qualities," Philosophical Studies 32 (1977), 337.

⁴Sellars (1963), p. 107.

⁵Ibid., p. 121.

⁶Ibid., p. 107.

⁷Ibid., p. 123.

⁸Ibid., p. 124.

⁹Ibid., p. 126.

¹⁰Ibid.

¹¹Ibid., p. 37.

¹²Ibid., p. 10a.

Chapter 5: Conclusion

I am now in a position to make my concluding comments on the three arguments for *sensa*. It should be clear by now that it is only the argument from the homogeneity of colour that I consider to be Sellars' principal argument that *sensa* are basic particulars of the scientific image. Before turning to this argument, I shall briefly summarize my reasons for rejecting the other two arguments.

The argument in "Phenomenalism" does give the appearance of presenting an alternate argument for the basicness of *sensa*, for it follows the same pattern as the argument from homogeneity but using different terminology, and comes to the same conclusion. Yet I think it can best be regarded as supplementary to the principal argument. It emphasizes issues that are not found in the earlier paper, "Philosophy and the Scientific Image of Man," but does not even mention homogeneity, the central issue of the earlier argument. But if we try to regard the argument in "Phenomenalism" as independent we run into difficulties; for example, the only reason given for claiming manifest objects don't exist, instead of identifying them with their scientific counterparts is:

It requires one to say that one and the same thing is both the single logical subject of which an undefined descriptive predicate (e.g. 'red') is true, and a set of logical subjects none of which is truly characterized by this predicate, thus raising all the logical puzzles of 'emergence.'

Yet in the situation which exists before chemical theory is reduced to physical theory, the same can be said of the primitive property of being a water molecule which is subsequently identified with certain complexes of physical primitives. We can say the original water molecule does not exist, but such a claim in this context is not very interesting. The aspects of the argument that are interesting for the argument for *sensa* are glossed over: the use of the principle of reducibility which is essential to the argument for *sensa* is simply "the logical puzzles of emergence;" the peculiar nature of redness, homogeneity, which makes it not only undefined, but not definable is not even mentioned.

When it comes to the actual argument for *sensa*, accounting for sensation in the scientific image, the argument in "Phenomenalism" is again lacking in necessary detail. For it hinges on the claim that:

... 'identifying' ... a person with a plurality of logical subjects, i.e. the constituent parts of the 'computer,' we have undermined the logic of sense impressions. For whether these parts be construed as material particles or as nerve cells, the fact that they are a plurality precludes them from serving either jointly or separately as the subjects of the verb 'to sense red-triangle-wise.'²

But no elaboration of what the logic of sense impressions is that apparently must be preserved, i.e. what feature specifically of the logical space of colour will be undermined. As far as I can see, on any interpretation of this argument, this point is an essential part of the argument; my conclusion is that this

omission indicates that this is not to be taken as an independent argument for *sensa*. Rather it only purports to present the form of the argument, filling in some of the details that were omitted from the main presentation in "Philosophy and the Scientific Image of Man."

I therefore discount the significance of section VIII of "Phenomenalism" as an independent argument for *sensa*. It would seem to amount to the argument that colours are primitive in the manifest image, i.e. are occurrent and apply to single logical subjects, and must therefore be primitive in the scientific image. But we saw in Chapter 4 that such a strong interpretation of the principle of framework adequacy is not justified --in fact primitiveness is the one logical feature of properties which cannot necessarily be expected to survive a framework replacement. Thus the argument must rely on other features of colours--which features are not mentioned in "Phenomenalism."

I don't think this is an unreasonable interpretation of Sellars' intentions regarding this passage. It is Sellars' method to repeat often his arguments in different papers in greater or lesser detail depending on the focus of that paper. The focus of "Phenomenalism" is phenomenism, not scientific explanation of man. Since "Phenomenalism" first appeared in Science, Perception and Reality and, therefore, in close proximity to "Philosophy and the Scientific Image of Man," it is plausible to suppose that it is to be taken as a supplement to the earlier argument which it presupposes. In this vein, I point out that

while the early paper is much concerned with the argument for sensa, less with describing them, the later paper takes much greater pains to characterize sensa, and to expose "mistaken presuppositions and metaphysical assumptions"³ which might stand in the way of an acceptance of sensa.

It should be no surprise that I turn my attention next to discounting the last sections of "The Identity Approach to the Mind-Body Problem" as an argument that sensa must be basic particulars in the scientific image. I argued in the first chapter that we cannot conclude from the fact that colours are pure occurrents that they are immune to scientific explanation, only that they do not invite it in the same way that dispositional properties do. In a similar vein, we could say that the property of being a water molecule does not require scientific explanation; nonetheless this does not prevent the water molecule being reduced in the scientific image.

However, it seems there must be more than this behind Sellars' claim that "the logical space of the perceptible qualities and relations of physical things and processes...is, in an important sense, closed."⁴ Thus in Chapter 3 I explored certain possible results of colours being pure occurrents. I concluded that a possible motivation behind the argument that colours are scientifically primitive is that otherwise there is a sense in which they will not appear in a description of the constituents of the world, being but defined properties and, therefore, nothing but some other properties. On the other hand, however,

I concluded that if this description of the world includes an explanation of how persons come to know the world, it may turn out that colours necessarily turn up, albeit as defined properties. Being necessarily epistemological primitives is not incompatible with being ontologically defined properties.

There is, of course, some difficulty in disposing of the evidence that Sellars does take this passage to be his argument for *sensa* being primitive--the evidence consisting of Sellars' reference in his "Reply to Cornman"⁵ to the concluding sections of "The Identity Approach to the Mind-Body Problem" as an explanation of why colours must be primitive in the scientific image. On the interpretation of Sellars' argument that I am pushing, he should have referred to "Philosophy and the Scientific Image of Man" for a discussion of "the distinctive character of the explanandum which called for the introduction of sense impressions in the first place."⁶ Yet I think a fairly satisfactory explanation can be given.

In the first place, it is possible that the references given, in elaboration of this phrase, are meant to draw attention not to precisely what the distinctive character is--in my interpretation, homogeneity--but that the explanandum is distinctive. Thus in the first reference to "pp. 399ff" of the same paper ("Reply to Cornman") the emphasis is on the fact that an intermediate level of explanation is needed to explain the human behaviour which includes expressing perceptual propositional attitudes, but not to explain discriminative behaviour. But I think it is clear

that this point is necessary but not sufficient for the argument for the primitiveness of sense--also essential is the distinctive nature of the properties of sense impressions. I assumed, initially, that the second reference given to "the concluding sections" of "The Identity Approach to the Mind-Body Problem," was intended to elaborate on this other aspect of the argument. Yet if we examine carefully the order of the argument in this passage, section VI, it can be argued that the second reference merely makes the same point as the first.

Section VI is obviously concerned with the primitiveness of colour in the scientific image. This discussion takes place in the first paragraphs, numbers 41-49. Yet in these paragraphs, there is no discussion of the specific character of "raw feel universals" which requires that they be primitive in the scientific image--the fact is simply asserted that they will be. The question is discussed in paragraphs 41-47, and the conclusion summed up in paragraph 48:

It is my conviction that a theory which is to explain the properties of core persons will involve a family of families of predicates which would be a categorical transformation, but not substantive reduction, of raw feel predicates, and which would apply only to systems of scientific objects which are the theoretical counterparts at the most fundamental level of empirical brains. In other words I accept the identity theory only in its weak form according...to which raw feels or sense impressions are states of core persons, for as I see it, the logical space of raw feels will reappear transposed but unreduced in a theoretical framework adequate to the job of explaining what core persons can do. In my opinion such a theory is not even yet on the horizon.⁷

But it is only in paragraph 49 that a reason is given for this conclusion, and this reason is precisely the one given in "Reply to Cornman," pp. 399ff.:

The plausibility of the more radical interpretation of the reducibility of neurophysiology to micro-physics rests on the fact that if one thinks of 'sense impressions' or 'raw feels' as theoretical constructs introduced for the purpose of explaining simple 'discriminative behavior' such as is found in white rats, then one would indeed find no reason to suppose that the postulated states might not be conceived of as reducible along the lines described in 46. After all, we can conceive of--and even construct--machines which can perform these discriminations. It is therefore crucial to my thesis to emphasize that sense impressions or raw feels are common sense theoretical constructs introduced to explain the occurrence not of white rat discriminative behavior, but rather of perceptual proposition attitudes, and are therefore bound up with the explanation of why human language contains families of predicates having the logical properties of words for perceptible qualities and relations.

Needless to say, I have already argued against the claim that an explanation of why humans have words for colours need conclude that such colours are primitive (and therefore in this sense, absolute) aspects of the world. My view is that colours can be irreducible primitives in an account of learning, i.e. epistemologically basic, without being ontologically irreducible. Yet on the interpretation I am presently pursuing, Sellars does take this to be a sufficient reason for concluding sense must be ontologically primitive in the scientific image. On this view, it is clear that by the end of paragraph 49, Sellars has completed his argument that raw feel predicates are transposed unreduced into the scientific image.

Yet there remain three paragraphs, and if one is not persuaded by the argument of paragraph 49, one is tempted to look for an argument in paragraphs 50-52. This is indeed what Hooker has done, drawing an argument from these three paragraphs (see above, p. 6). However, a careful examination of the text shows that Hooker has taken key phrases out of context, constructing an argument about *sensa* where none was intended. In the first place, if these paragraphs were part of the argument that *sensa* are primitive, we should expect them to be conjoined somehow with the logical development of the preceding paragraphs. Instead, the last paragraphs are clearly separated from 41-49, for 50 begins with the words: "I shall conclude with a brief mention of other facets of the problem." Secondly, the argument for *sensa* concerns primarily the qualities of sense impressions, yet paragraphs 50-52 are almost exclusively about "the logical space of the perceptible qualities and relations of physical things and processes on which that of the attributes and relations of raw feels is modeled." The discussion of these paragraphs is concerned specifically with the puzzle of these properties in the scientific image--and indeed, if it is thought that physical objects are quite literally homogeneously coloured, this is a problem. In this specific context, the argument that this logical space is closed, and therefore, I suppose, irreducible is merely an extension of the early argument of paragraphs 48 and 49 concerning the logical space of raw feel attributes, for we find a transposed form of the argument of 49:

Roughly, it is not such facts, expounded in a 'phenomenology' of sensible qualities and relations, as that to be orange is to be between red and yellow in color which demand scientific explanation, but rather such nomological facts as that black objects sink further into snow than white objects when the sun is shining.

Again, however, it is my belief that such an argument cannot stand on its feet--it requires an argument which deals specifically with the homogeneity of the colours in question (whether of physical objects, or sense impressions). It seems Sellars may accept a principle to the effect that properties such as colours are so basic to our view of the world that it is implausible to envisage that they don't really exist in some form. Indeed such a sentiment seems to be behind a great deal of argument against reducing colours out of existence. My view of this type of consideration is that if we are going to be scientific realists; we must be prepared to give up even the most fundamental aspects of the manifest world, if that is dictated by our scientific ontology. The fact that it may seem very implausible to suggest colours don't exist in any fundamental sense leaves me unmoved. Our feelings of certainty in considering the existence of colour can be explained in a scientific account of the thought and beliefs of persons. In fact, the impossibility of being totally objective in such a matter dictates that we ought to attempt to form our ontological and scientific principles on the basis of cases further removed from the realm of the conceptual, and then apply them as dispassionately as possible to

an explanation of the conceptual processes of persons. According to this principle, then, we ought to try to explain everything in terms of the scientific/ontological concepts we have evolved to explain non-living matter; only if such an undertaking fails are we justified in postulating new types of particulars. It is according to this principle that I hold that a specific argument concerning the inability of current scientific particulars to explain persons is necessary--i.e. the argument from homogeneity.

In any case, the conclusions of the last three paragraphs of "The Identity Approach to the Mind-Body Problem" are primarily about physical objects--not, as Hooker seems to imply, sense impressions--as is evident in paragraph 52:

Scientific Realism maintains the in principle replaceability of the framework of perceptible things by a framework of scientific objects which contains highly derived counterparts of the inductively established causal properties of the former. But while Scientific Realism grants that the framework of scientific objects also contains highly derived counterparts of the occurrent perceptible qualities of perceptible things, it need not and, if my argument is correct, must not hold that these qualities are reducible to i.e. replaceable by, their counterparts in micro-physical theory--as in the chemistry-physics case. The intrinsic structure of their 'closed' logical space (paragraph 50) requires rather that they be relocated. This relocation involves a simultaneous move on the sense-impression front. For the qualities and relations which are irreducible to their counterparts in the micro-physics of the objects of perception, are reducible, i.e. identifiable with, the qualities and relations which, I have contended (paragraphs 45-48), must be postulated in an adequate theoretical explanation of the nature and function of sense impressions or 'raw feels.'¹⁰

Of course, I have argued against the claim that it is a necessary feature of the manifest image that physical objects be occurrently coloured. Nonetheless, it is clear that if this is an accurate account of how the manifest image actually is, these properties of physical objects cannot be reduced but must be relocated i.e. identified, in a sense, with the logical space of manifest sense impressions. How the subsequent reduction of this logical space is construed is a different question.

I think my insistence that Sellars' argument for sensa is the argument from homogeneity is confirmed in his "Reply to Cornman" where his reconstruction of the argument is as follows:

We must find a place in the world for color in the aesthetically interesting sense with its ultimate homogeneity....Can we rest content with the idea that red in the aesthetically interesting sense is a manner of sensing, that (in traditional terms) its esse is percipi? Is its ontological status given by the context: Person senses-redly?

Now science suggests that persons are systems of scientific objects. But, if we accept the principle of reducibility, for a system of scientific objects to sense-redly must consist in its constituents being in certain states and standing in certain relations to each other. Now sensing-redly as conceived in the Manifest Image does not consist in a relationship of objects in states other than sensings. A sensing can include other sensings, as when we sense a-red-circle-in-a-green-square, but it cannot consist of non-sensings.

...(In the scientific image sensing) will still be a state of a person...taking the principle of reducibility into account, it will be a state which consists in certain states of and relations between these "objects."

Now the successor concept of (visual) sensing... must relocate the "ultimate homogeneity" of the

(colors of the Manifest Image). But it cannot do so if the persons to which this successor concept applies consist of objects to which color concepts, in a sense which preserves the essentials of color space, do not apply.¹¹

Sellars thinks that the only outcome to this argument, if the principle of reducibility is not to be abandoned, is either to accept Cartesian minds or postulate "a new domain of scientific objects to be the subjects of these color predicates."

It should be fairly clear from the preceding, particularly Chapter One, what my criticism of the argument from homogeneity is. The argument works by setting certain paradigms before us and extrapolating conclusions from them. Two concepts interact in the argument: the concept of homogeneity and the principle of reducibility. The model for homogeneity is coloured manifest physical objects; it is explicitly spatial, but I have argued that its analogue, the homogeneity of sense impressions, is most likely to be temporal--somehow the three dimensional structure of a homogeneous pink ice cube will "map" onto a one dimensional, highly complex process. The principle of reducibility is more difficult to get a grip on. Indeed within the manifest image we have a fairly clear idea of it--the model is that the property of being a ladder consists of or is defined by the properties of being cylindrical, wooden etc. Thus we have two paradigms for the principle, both of which are extended with difficulty into the scientific framework. The notion of properties being defined by other properties is probably the most straightforward paradigm. Yet if it is fairly clear as it pertains to

the manifest image, where we are very familiar with the meanings of the terms both of the definiens and definienda, the notion is only with great difficulty extended to situations such as sketchy theories in which our grasp of the meanings, and precise logical spaces of the terms is extremely limited.

I think Sellars is more comfortable, in the argument for *sensa*, to use the paradigm of properties consisting of other properties. I like to think of this paradigm in terms of the material mode, involving the things themselves, as opposed to the formal mode model of definition. As such the former has a certain intuitive appeal, focusing our attention on individuals rather than meanings or roles in a language. Sellars' choice of this paradigm in an unfamiliar situation parallels his analysis of the use of models in science for intuitively understanding concepts when it is not possible to specify their logical space. I have no objection to the notion of consists of as a paradigm of property reduction; however, it must be used with extreme care in situations where more than one ontological type of individual may be relevant. But this is precisely the situation in the argument for *sensa*, where we are attempting to decide between, on one hand, an ontology of enduring particulars (or "things") and states of particulars, and a ontology of "events." To each of these kinds of individual corresponds a different notion of consists of. To the former belongs a spatial notion of consists of; to the latter, a more complex notion involving both space and time.

The point is just this: the manifest image, and the current account of the scientific image, already contain both types of individuals--things which constitute other things, and events which constitute processes. Just as a ladder consists of various pieces, so a hydrogen molecule consists of hydrogen atoms. Similarly, processes consist of other processes and events: making a telephone call consists of lifting the receiver, and turning the dial; the chemical process of oxidation consists of the movement of electrons between subatomic particles. Within this familiar framework the homogeneity of "coloured" sense impressions, involving a temporal notion of homogeneity, can be accommodated. Sellars' argument has force only if one is willing to deny the existence of process and event--and the corresponding notion of consists of--within our current conceptual structure.

Of course, the "events" Sellars has in mind are not the events to which I refer. Presumably Sellarsian "events," unlike our events, do not involve things i.e. there is no thing which is the subject of the "event." My claim is that Sellars has not given sufficient reasons for the introduction of these esoteric individuals. The introduction of such a notion, with a correspondingly novel notion of consists of, unnecessarily complicates an ontology, already both rich and little understood, which contains both things and events. Sellars' argument for *sensa* relies on the question of what can, and what cannot, be defined within our current conceptual scheme; but since he fails

to explore sufficiently carefully the possibilities of definability within this scheme, the argument fails.

FOOTNOTES

¹Wilfrid Sellars, Science, Perception and Reality (London: Routledge and Kegan Paul, 1963), p. 98.

²Ibid., p. 101.

³Ibid., p. 102.

⁴Wilfrid Sellars, "The Identity Approach to the Mind-Body Problem," Boston Studies in the Philosophy of Science, Volume Two (New York: Humanities Press, 1965), p. 74.

⁵Wilfrid Sellars, "Science, Sense Impressions, and Sense: A Reply to Cornman," Review of Metaphysics 25 (1971), p. 424, fn. 25.

⁶Ibid., p. 424.

⁷Sellars (1965), p. 73.

⁸Ibid., pp. 73-74.

⁹Ibid., p. 74.

¹⁰Ibid., p. 75.

¹¹Sellars (1971), pp. 408-409.

BIBLIOGRAPHY

A. Relevant Secondary Sources

Cornman, J. W. Materialism and Sensations. New York: Yale University Press, 1971.

----- "Sellars, Scientific Realism, and Sense,"
Review of Metaphysics 23 (1970), 417-451.

Harman, G. "Sellars' Semantics," Philosophical Review 79 (1970),
404-419.

Hooker, C. A. "Sellars' Argument for the Inevitability of the
Secondary Qualities," Philosophical Studies 32 (1977),
335-348.

B. Relevant Works by Wilfrid Sellars*

"Empiricism and the Philosophy of Mind," Science, Perception
and Reality. London: Routledge and Kegan Paul, 1963.
Pp. 127-196.

"Phenomenalism," Science, Perception and Reality. London:
Routledge and Kegan Paul, 1963. Pp. 60-105.

Philosophical Perspectives. Springfield, Illinois: Charles
C. Thomas, Publisher, 1967.

"Philosophy and the Scientific Image of Man," Science, Perception
and Reality. London: Routledge and Kegan Paul, 1963. Pp.
1-40.

Science and Metaphysics: Variations on Kantian Themes. London:
Routledge and Kegan Paul, 1968.

Science, Perception and Reality. London: Routledge and Kegan
Paul, 1963.

"Science, Sense Impressions, and Senses: A Reply to Cornman,"
Review of Metaphysics 25 (1971), 391-447.

"Scientific Realism or Ironic Instrumentalism," Philosophical
Perspectives. Springfield, Illinois: Charles C. Thomas,
Publisher, 1967. Pp. 337-369.

"Some Reflections on Language Games," Science, Perception and
Reality. London: Routledge and Kegan Paul, 1963. Pp.
321-358.

"The Identity Approach to the Mind-Body Problem," Boston Studies in the Philosophy of Science, Volume Two. Edited by Robert S. Cohen and Marx W. Wartofsky. New York: Humanities Press, 1965. Pp. 55-76.

"The Language of Theories," Science, Perception and Reality. London: Routledge and Kegan Paul, 1963. Pp. 106-126.

"Theoretical Explanation," Philosophical Perspectives. Springfield, Illinois: Charles C. Thomas, Publisher, 1967. Pp. 321-336.

*Articles and books are listed alphabetically by title. Editions of articles listed here are not necessarily the original printing; reference is to edition used in preparing this thesis.