

## **EXTERNALISM & SELF-KNOWLEDGE: A SOLUTION FROM SELLARS\***

**Keya Maitra**

University of North Carolina at Asheville, USA

**ABSTRACT:** The renewed interest in self-knowledge in the 1990s has been motivated by the idea that the Self-knowledge thesis might hold the key to defeating the externalist theory of content determination. If what I am thinking about is determined not simply by what goes on inside my mind but also by factors existing outside my mind, as content externalism argues, then it seems that I need to consult the external world before I can know my mind and its occupants. Thus the possibility of direct and privileged knowledge of the contents of one's own mind seems to be defeated if content externalism is true. A dominant externalist response argues that one knows one's thoughts and other mental states simply by thinking them while exercising second-order, self-ascriptive powers. This dominant externalist account of privilege of self-knowledge claims has been referred to as the 'deflationary' account. Deflationary accounts are faulted for being epistemologically impoverished since in being automatically justified there is no room left for 'epistemological achievement.' The goal of my paper is to argue that a truly viable non-deflationist and externalist account of self-knowledge can be articulated from Sellars's writings. Sellars offers an interesting way of understanding the privilege of self-knowledge claims according to which though the self-knowledge claims enjoy non-inferential privilege, they are nonetheless defeasible and also justifiable. The effectiveness of this account will become evident from comparing its fate against a couple of common worries that have plagued Shoemaker's very prominent deflationist account.

**KEY WORDS:** Self-knowledge, externalism privileged knowledge, Sellars, non-deflationist and externalist account of knowledge.

---

\* Some of the ideas presented in this paper were first developed in my Ph. D. thesis titled, Our Knowledge about our own Mental States: an Externalist account. I would like to thank Ruth Millikan for helpful comments & suggestions.

A curious fact of late 80s and 90s was that almost every major journal issue in analytic philosophy contained at least one article devoted to the problem of 'compatibilism', the problem of making externalism compatible with self-knowledge. Self-knowledge stands for the thesis that a subject knows at least some of her mental states in a direct and privileged way. Thus, while I know that I believe that Halloween is the time for decoration in a direct way, you need to rely on inferring from my behavior including my verbal utterances that I believe that Halloween is the time for decoration. Externalism about mental content argues that the content of a mental state, i.e., what a mental state is about, is determined at least partially by causal relations that the constituents of that state have to things existing outside of the mind. So when I hope to read *The Lowland* by Jhumpa Lahiri, the content of my mental state is not determined by what is going on inside my mind alone but also by the newest novel by Lahiri that exists outside of my mind. Content externalism is often taken to be incompatible with the immediacy and infallibility implied by the self-knowledge thesis. Indeed, it may not be incorrect to argue that the renewed interest in self-knowledge in the 1990s has been motivated by the idea that the Self-knowledge thesis might hold the key to defeating the externalist theory of content determination. If what I am thinking about is determined not simply by what goes on inside my mind but also by factors existing outside my mind, then it seems that I need to consult the external world before I can know my mind and its occupants. Thus the possibility of direct knowledge of the contents of one's own mind seems to be defeated if content externalism is true.

Interestingly Sellars has rarely been seen as offering a solution to the compatibilism problem. Indeed, Sellars's kind of externalism, where a thing's *looking* red is explained in terms of its *being* red, might seem to exemplify the extent of an externalist's difficulty in accounting for self-knowledge. Even when Sellars has been mentioned in this context, his contribution has been taken to be a version of the dominant compatibilist effort from the externalists (see Thomasson, 2003, 240 for example). This popular externalist response was first proposed by Burge (1988) and later adopted by most externalists (Wright, Heil, Shoemaker, & Dretske among others). In this effort the 'privilege' of the self-knowledge claims is explained by means of the 'containment thesis' (Burge 1988, 659). According to this thesis the first order state is not only thought about in the second order state but it is *also* thought in the very act of making the second order claim. Thus, according to this thesis, "One knows one's thought to be what it is simply by thinking it while exercising second-order, self-ascriptive powers" (Burge, 1988, 656). In this model then, the first order and the second order states are not truly distinct in the sense of one providing justification for and room for fallibility of the other. Rather thinking the second order *is* thinking the first order 'in a certain way' and thereby being automatically rendered as privileged. Using a term of Sawyer's (1999), I am going to call this dominant externalist account of privilege of self-knowledge claims the 'deflationary' account. I take Thomasson to be treating Sellars's account as deflationary since she focuses on its similarities with Shoemaker's account, a clearly deflationist response to the compatibilist problem. Reversing this trend the goal of my paper is to argue that a truly viable non-deflationist compatibilist response can be articulated from Sellars's writings on this topic. The effectiveness of this account would be evident from comparing its fate against a couple of common worries that have plagued Shoemaker's very prominent deflationist account. In order to truly appreciate the merits of the Sellarsean account, however, we will need to clarify the notion of self-knowledge with the help of a couple of distinctions. So that is where we will begin.

## Section I: A distinction relating to the epistemic advantage of Self-knowledge

It is useful to introduce a distinction between two different kinds of epistemic advantage, namely, in terms of ‘epistemic access’ and in terms of ‘epistemic authority,’ that a privileged claim might enjoy. The typical understanding of self-knowledge claims is that the subject of such a claim is specially situated in having an *immediate* access in relation to the object of the knowledge claim. However, traditionally the epistemic advantage of self-knowledge claims is also understood in terms of the security or certainty that these claims are taken to enjoy as implied by terms like ‘infallible,’ ‘incorrigible’ and ‘indubitable’ referring to adjectives typically associated with these claims. But having an immediate access to the object of a knowledge claim does not necessarily equal to such a claim being infallible or without the possibility of being doubted. This is especially so if we understand the immediacy of self-knowledge claims in terms of the specific way of making and justifying such claims, i.e., *how* we make such claims. This is quite different from being certain or sure of such knowledge claims. Thus I have argued elsewhere (Maitra, 2005) that the privilege associated with self-knowledge claims is best understood as a kind of epistemic access privilege. If we accept my argument that ‘privileged in access’ and not ‘privileged in authority’ captures the epistemic privilege of self-knowledge claims, then Sellars can be taken to give us an account of such privileged in access.

In fact, ‘privileged in access’ can be further distinguished to refer to two different things, namely, how the self-knowledge claims are made *and* how such claims are justified and validated. Thus we get two further kinds within the privileged in access kind: (a) non-inferential privilege, and (b) non-evidence based privilege or privileged justification. A knowledge claim is non-inferential if one does not rely on any inference, at least conscious, in *making* that claim. Knowledge claims that are non-evidence based, on the other hand, are claims that are not derived in the sense that their validation or justification is not derived from any other claim. The privileged status here consists of the fact that the justification of the kind of claims does not derive from any other kind of knowledge, thereby being epistemically ‘immediate’ (Alston, 1971, 233), or epistemically independent or epistemically ‘given’ (Sellars, 1963). A claim that is non-inferentially made can nonetheless fail to be epistemically basic, as argued by Sellars. Indeed, Sellars offers an interesting way of understanding the privilege of self-knowledge claims according to which though the self-knowledge claims enjoy non-inferential privilege, they are nonetheless defeasible and also justifiable from an independent source. This provides an advantage over the deflationary accounts of self-knowledge. A characteristic claim in any deflationary account is that the epistemic privilege in relation to the self-knowledge claims derives from the psychological and/or cognitive mechanisms that underlie our abilities to make these claims. Thus they do not need to be and in fact cannot be given independent justificatory reasons. However, as Boghossian (1989) has argued, these deflationary accounts are epistemologically impoverished since in being automatically justified there is no room left for what he calls ‘epistemological achievement.’ A corollary of this automatic justification is the fact that all these claims turn out to be infallible. Once again this rules ‘epistemological achievement’ out since getting something right makes sense only when there is the possibility of getting it wrong! Indeed, many self-knowledge claims that we ordinarily make fall far short of being infallible or incorrigible. Rather, these claims are quite

defeasible. Thus the deflationary account, in restricting itself to infallible claims, presents an inadequate model to understand the epistemic status and privilege of self-knowledge claims.

## Section II: Sellarsean Non-Deflationist account

In order to articulate a compatibilist account of the non-inferential privilege characteristic of self-knowledge claims from the Sellarsean perspective we will need to start with his account of perceptual judgments.

### Section II. 1: Perceptual reports and the underlying epistemology

The originality of the Sellarsean account of the non-inferential privilege of self-knowledge claims lies in his likening it to the non-inferentiality involved in perceptual judgments. In order to grasp Sellars's ideas on perceptual reports, it will be useful to recapitulate here how the epistemology of perceptual judgments is understood in the modern philosophical framework. Take for example Descartes' position in the *Meditations*. Descartes argues that perceptual knowledge is best understood by looking at (introspecting) the inner episodes and inner going-ons –thoughts, beliefs, desires and sensations or sense impressions– which the mind is immediately and directly aware of. External perceptual objects are then understood as unobservable causes of these episodes.

Sellars rejects this model of perception where material object's having a certain color is *explicated* in terms of a hypothetical reference to that object's 'looking or appearing' of be of certain color or my sensation of its color. Reversing this direction of explanation, Sellars argues that the concept of a thing's *looking* red has to be explicated in terms of the concept of its *being* red. According to Sellars we begin with our knowledge of the medium sized objects. Then using this knowledge as a model we come to know about *what is going on in our minds* (among many other things) «by extending and refining the conceptual framework or representational system originally applied to ordinary physical objects» (deVries & Triplet, 2000, xxxix). In fact, he takes the concept of a sense impression, for example, a sensation of a pink ice cube, as a 'theoretical concept',<sup>1</sup> modeled on our concepts of physical objects.

Now the question is how do we come to acquire the conceptual framework of ordinary physical objects? That is to say, how does one come to learn to make perceptual report? Sellars answers these questions by noting the fact that concept acquisition and possession track language learning very closely. We acquire the ability to make the perceptual reports by learning to use the concepts involved in those reports in different perceptual situations. The learning process here is that of being trained by one's peers to respond reliably to a given situation. Learning to use an observational concept further requires that one learns to recognize the conditions that one would need to view the thing under in order to ascertain its properties (say, color) by looking. Thus in this Sellarsean

---

<sup>1</sup> A theoretical concept is a concept of an unobservable entity (like quarks) which is 'posited' in order to explain certain observable phenomena. See below for more on this.

framework, not only does one's understanding of 'x is green' comes prior to one's understanding of 'x looks green,' but to have such an understanding is to have "the ability to tell what colours objects have by looking at them – which, in turn, involves knowing in what circumstances to place an object if one wishes to ascertain its colour by looking at it" (Sellars, 1963, 146). Thus learning a concept is acquiring an ability to respond reliably to given perceptual situations by producing perceptual reports. A mastery of this ability would also imply that

one must have some conception of nonstandard conditions. One would probably have to understand, for example, that conditions are not good for accurate color determination of an object outdoors if it is viewed in the early morning or evening in strongly filtered light. One might understand about illusions caused by indoor colored lights, and about illusions sometimes caused deliberately, as in the case of theater lighting. What will count as standard conditions cannot be fixed *a priori* –daylight need not be standard illumination for a space-traveling culture, for instance– but there will always be some cultural agreement about what counts as standard (deVries & Tripplet, 2000, 29).

However, as Sellars also notes, what counts as standard conditions for a given mode of perception will always have the «vagueness and open texture characteristic of ordinary discourse» (Sellars, 1963, 147). To summarize, according to Sellars the non-inferentiality of perceptual knowledge «involves reliably *responding* to physical objects in standard conditions with the appropriate perceptual sentences; ..., where the *reliability* of the response is a function of the way in which language is learned» (Sellars, [1975], 1991, 376).

Let us now turn to the epistemological side of the story. Our focus is on the nature of these reports and how one justifies them. In the Cartesian model the inner episodes are taken to be epistemologically given, i.e., the starting points, since they are epistemologically independent in not being 'based on' anything else and also efficacious in providing the foundation of all other knowledge claims. Sellars rejects this epistemology of given as a 'myth' since according to Sellars there is nothing that can fit the bill of being epistemologically independent as well as efficacious. But one may think that in rejecting the given as a myth, Sellars is altogether ruling out the role of direct knowledge. This is however, quite far from truth. The interesting point about the perceptual reports of Sellars is that he takes them to be paradigmatic examples of non-inferential belief, i.e., claims made *not* on the basis of some other knowledge that one may have. Thus though the learning process requires that one be able to recognize the standard and also nonstandard conditions, it is never argued that such reports are *inferred* from the prevailing conditions.<sup>2</sup> These reports and the resultant direct knowledge play an indispensable role in our acquiring knowledge about the world. However, Sellars argues

---

<sup>2</sup> As Sellars continues further by presenting his ideas about perceptual reports in terms of the Verbal Behavioral model: «Now to say that this visual thinking-out-loud that something is the case [for example, where Jones thinks out loud: 'Lo! Here is a red apple'] is epistemically *justified* or *reasonable* or has authority is clearly *not* to say that Jones has correctly inferred from certain premises, which he has good reason to believe, that there is a red apple in front of him. For we are dealing with a *paradigm* case of non-inferential belief. *The authority of the thinking accrues to it in quite a different way. It can be traced to the fact that Jones has learned how to use the relevant words in perceptual situation*» (Sellars, [1975], 1991, 376, emphasis in original).

that what is known directly may not be a full-fledged given in the sense of having «no *justificational* intermediate step» (deVries & Triplet, 2000, xviii). No epistemological independence where one does not require to derive the justification from any other belief is possible, according to Sellars, even in relation to these non-inferential perceptual reports. Thus the justification of these reports is finally provided by an appeal to the prevailing conditions and thus being inferred from certain other knowledge that the subject might have. What provides the justificational basis for these reports are «fundamental associations [that are] acquired via complex learning activities usually mediated by social direction and intervention» (deVries & Triplet, 2000, xxxv-vi). As Sellars writes of a person making the perceptual report that there is a red apple in front of him, «He did not originally *infer* that there is red apple in front of him. *Now*, [while justifying the claim], however, he is inferring from the character and context of his experience that it is veridical and that there is good reason to believe that there is indeed a red apple in front of him» (Sellars, [1975], 1991, 376). Thus clarifying the ambiguity that often mars the classical understanding of ‘given’ as well as ‘direct’ knowledge, Sellars introduces a distinction between two senses of ‘direct’ one meaning non-inferential report while the other meaning involving no justificational intermediary. He then argues that while there is direct knowledge in the sense of first, there is no such knowledge in the sense of second. With this stage setting I will now turn to Sellars’ conception of self-knowledge.

## **Section II. 2: Verbal Behavioral (VB) model of thought and the underlying epistemology**

Like the concepts of sense-impression invoked in explicating our observational reports or perceptual judgments, Sellars argues that the concepts of our various propositional attitude states are like our concepts of theoretical entities<sup>3</sup> in the sense that like electrons they are posited in order to explain the behavior of some observable entities. According to Sellars, human thoughts and other propositional attitude states are modeled on concepts of overt verbal behaviors. So Sellars needs to tell us a story about our conception of overt verbal behavior in such a way that they do not rely on our prior conceptions of thoughts. Sellars provides such a story by telling us his ‘anthropological’ science fiction involving our ancestral community of Ryleans which culminates in the myth of Jones. The language spoken in this community is ‘Rylean’ since its «fundamental descriptive vocabulary speaks [only] of public properties of public objects [therefore non-theoretical] located in Space and enduring through Time» (Sellars, 1963, 178). The list of public objects also includes each others’ observable verbal and non-verbal behavior. The question that Sellars raises is: «What resources would have to be added to the Rylean language of these talking animals in order that they might come to recognize each other and themselves as animals that *think*, *observe*, and have *feelings* and *sensations*, as we use these terms?» (1963, 179) Sellars’s answer consists of noting

---

<sup>3</sup> Here by focusing on the concept of the states and likening them to concepts of theoretical entities, Sellars avoids, at least for the time, the difficult question about the ontological status of these states. Sellars clarifies further by claiming that «concepts of thoughts are like theoretical concepts, not that they are theoretical concepts» (Sellars, 1963, 188).

that such a resource would lie in the brilliant discovery of Jones, a member of this Rylean clan, of a «*theory* according to which overt utterances are nothing but culmination of a process which begins with certain inner episodes» like belief and desire etc. (Sellars, 1963, 186). Jones' theory thus posits concepts of unobservable entities like belief and desires in explaining and predicting the behavior of other people by ascribing these various theoretical inner episodes –mental states– to them. The word of Jones' invention gets around and soon all the other members learn (or get trained by Jones) to posit these theoretical entities in predicting and explaining the verbal and non-verbal behaviors of others. They also soon discover and likewise learn to or get trained in using these theoretical entities in explaining and predicting their own behavior.<sup>4</sup> Eventually members of this Rylean clan get so good at these self-ascriptions that they do not have to rely on observing (or inferring from) their own behavior in making these self-ascriptive claims. In other words, they come to *report* their various inner episodes non-inferentially. But how to get from our Rylean ancestors to the present language user and belief ascriber?

As mentioned earlier Sellars models his concept of belief and other propositional attitudes on his concept of overt behavior itself. Thus he offers his Verbal Behaviorist or VB model of thought that models thought and other propositional inner episodes as internalized speech. Learning a language or speech for Sellars is to be trained to respond reliably to a situation with a sentence or a word under standard conditions. Learning to think likewise is to learn to respond to a situation with a short-term propensity to *say* a sentence or a word without saying it loud. In this model 'thinking that-p' is saying or having a short-term proximate propensity to say 'p'. It is, as it were, a 'think-out-loud' model. Thinking is in essence nothing more than non-verbalized speech. Thus having a thought of a thought or a second order mental state typical of self-knowledge claims depends on one's learning to respond reliably to one's own short term verbal propensities with a certain *not-spoken-out-loud* words. As Sellars writes, «Can we not, as children, be trained by those who know us intimately (our parents), and who therefore know (*ceteris paribus*) what our short-term verbal propensities are (i.e., what we are thinking), to *respond* reliably to our own short-term propensities to say that-*p*, as well as to respond to our actual sayings of '*p*'?»<sup>5</sup> (Sellars, [1975], 1991, 377)

---

<sup>4</sup> Sellars' claim here about the order in which we learn to ascribe mental states to others and to ourselves is diametrically opposed to that of the Cartesian order. Interestingly, Sellars here seems also to be denying the popular idea that we need two different concepts of belief in order to explain the first person cases and the third person cases of belief ascriptions.

<sup>5</sup> Here is a long passage that represents the heart of Sellarsian ideas well.

For once our fictitious ancestor, Jones, has developed the theory that overt verbal behaviour is the expression of thoughts, and taught his compatriots to make use of the theory in interpreting each other's behaviour, it is but a short step to the use of this language in self-description. Thus, when Tom, watching Dick, has behavioural evidence which warrants the use of the sentence (in the language of the theory) 'Dick is thinking «p»' (or 'Dick is thinking that p'). And it now turns out –need it have? – that Dick can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe his overt behaviour. Jones brings this about, roughly, by applauding utterances by Dick of 'I am thinking that p' when the behavioural evidence strongly supports the theoretical statement 'Dick is thinking that p'; and by frowning on utterances of 'I am thinking that p', when the evidence does not support this theoretical statement. Our ancestors begin to speak of the privileged access each of us has to his own thoughts. What began as a language with a purely theoretical use has gained a reporting role (Sellars, 1963, 188-189).

Thus in this model «we can know what we think, in the primary sense, by literally hearing ourselves think. When we hear ourselves say (in a candid frame of mind) ‘I’ve just missed my bus,’ we are literally hearing the thought occur to us that we have just missed the bus» (Sellars, [1975], 1991, 375). This model also helps us know what we think when we are *not* candidly thinking out loud (Sellars, [1975], 1991, 375) because the learning process involved here train one to respond reliably to one’s actual sayings of ‘p’ as well as to one’s short-term propensities to say that-p. What is interesting for us in this Sellarsean model is its non-Cartesianism. Instead of conceptualizing in terms of ‘pure’ inner episodes detectable in inward glances, the concepts of beliefs and other mental states are conceived as theoretical entities understood in terms of non-verbalized speech which in turn is understood in terms of our ‘propensities to respond reliably.’ Thus Sellarsean naturalism here consists of the fact that human mind and privacy are treated as items within the natural, material world. Humans are natural organisms that learn to regulate their behavior by constructing even more adequate representation of their world and especially of themselves. Even though it is our self-knowledge that makes humans special (in comparison with other animals), no intrinsic epistemological priority is granted to self-knowledge in Sellars’ picture (deVries & Tripplet, 2000, xxxix). It is this point that underlies Sellars’ externalism as well. Thus rejecting the Cartesian model that takes the knowledge as ‘growing from inside out,’ the Sellarsean externalist model takes knowledge as ‘growing from the outside in.’ It is here that the strength of the Sellarsean model becomes the clearest. I will come back to this point.

Presently I want to draw our attention to the striking similarity of Sellars’ views on self-knowledge claims to his view about the non-inferentiality of the perceptual reports. In both cases the non-inferentiality comes about as a result of learning to recognize the standard conditions from which to infer these reports reliably. As Sellars asks, «would not the fact that such responses are *reliable* constitute the core of the explanation of non-inferential knowledge of what one is thinking (in the proximate propensity sense), as the existence of reliable verbal responses to perceptible things in the core of the explication of non-inferential perceptual knowledge?» (Sellars, [1975], 1991, 377) Thus, as in the perceptual case, making these reports does not involve inferring these reports and herein consists the privileged access and also the epistemic asymmetry characteristic of self-knowledge claims. The asymmetry here consists of the fact while *I* can make these claims about my *own* mental states non-inferentially, *you* will have to infer these states from my behavior. But in order for these claims to be knowledge, according to Sellars, I have to be able to justify these claims by referring to my verbal behavior. This is because, as Sellars points out, in order to know something we must be able to enter the ‘space of reasons.’

Let us now turn to Sellars’ distinction between the ‘space of reasons’ and the ‘space of description.’ Sellars writes, «The essential point is that in characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says» (Sellars, 1963, 169). This is because there is an «inescapable evaluative and normative dimension to knowledge» (deVries & Tripplet, 2000, 91), according to Sellars. To describe an episode as a knowing is to *endorse* it for ‘know’ according to Sellars is an achievement verb like ‘see.’ Calling a state a state of knowing does not merely involve providing an empirical description of it. To know is to do something, or to play one’s role in the ‘game’ of knowing that, in a Wittgensteinian



spirit, stands for a ‘complex, inter-subjective and norm-governed’ form of activity (deVries & Tripplet, 2000, 94). Justifying, within this context also stands for one’s ability to explain or give reasons for one’s move in the context of that game. Thus, acquiring the ability to know is like acquiring the ability to play a game by mastering the rule-governed complex activity. Success and failure in knowing is like failure and success in playing a certain game that requires knowledge and application of certain rules. In fact, it is this aspect that turns a simple movement of a wooden piece into a ‘move in the game.’ not the mere descriptive similarity that this movement may enjoy to a real move in a game. Thus, within the space of reasons, *any* claim is justified in being inferred from certain other claims because that is the very nature of the enterprise of reason-giving or justifying. In this regard it is useful to compare this Sellarsean idea of justification with the traditional foundationalist idea of justification.

One of the worries that motivated Sellars’ attack on the given is the misleading nature of the metaphor of ‘foundation.’ As Sellars remarks, «I do wish to insist that the metaphor of ‘foundation’ is misleading in that it keeps us from seeing that if there is a logical dimension in which other empirical propositions rest on observational reports, there is another logical dimension in which the latter rest on the former» (1963, 170). Now, what are these two dimensions and how exactly are they different? One is the dimension of general propositions of the form ‘Xs, when viewed under standard conditions, serve as reliable signs for the presence of Ys’. The other is the dimension of particular proposition of the form ‘X is Y.’ «Sellars claims that knowledge of *particular* facts, such as the fact that the crow outside Jones’ window is black, presupposes knowledge of certain *general* facts, such as the fact that reports of black things are, in standard conditions, reliable signs of the presence of black things. But how could Jones come to know such general facts?» (deVries & Tripplet, 2000, 89) Sellars here relies on the general inductive way of acquiring these general propositions. Is there a circularity involved? Because, particular propositions depend for their justification on knowledge of certain general propositions while general propositions themselves rely on the knowledge of some particular propositions for their justification. There would be such a circularity, indeed a vicious one, if the act of justification is understood in terms of temporal starting points. There is no such circularity in Sellars’ system, however, since the ability to know and justify is not based on any temporal assumption<sup>6</sup> here. This temporal assumption is replaced by a holistic coherence theory of justification. Sellars argues that it is possible to know both general and particular facts, *all at once*, as it were. This justificational coherentism and the resulting holism are backed up by a conceptual coherentism and holism in Sellars. According to this conceptual holism, an individual concept cannot be learned without also learning a battery of other concepts. The merit of the Sellarsean position can be highlighted when compared with the Shoemaker deflationist account.

Sellars rejects the pure externalism about justification in the deflationist account which takes justification to be a pure matter of production by reliable mechanism. Sellars adds, in order to know, «we must also *know* that our response is reliable» in addition to producing the response (deVries & Tripplet, 2000, xxxvii). Further, Shoemaker’s

---

<sup>6</sup> This is basically the assumption that the act of justification happens in a linear progression with starting points representing cases that do not rely on anything else in the system for being justified.

deflationary account seems to be caught in Boghossian's 'spurious trilemma' of either by inference, by observation or by nothing. Since the self-knowledge claims are not justified either by inference, or by observation, according to deflationism, they fail to reflect any epistemological achievements. Boghossian further argued if the justification was inferred then that compromised the directness and non-inferentiality of the self-knowledge claims. But does that really compromise the *non-inferential* nature of these claims? That would be so only if one were making the temporal assumption and also failing to make the very important Sellarsean distinction between the space of reasons and the space of description. As Sellars points out, to think that a knowledge in logically presupposing knowledge of other facts concedes its non-inferentiality «is itself an episode in the myth [of the given]» (Sellars, 1963, 164). Within the space of description we can describe and explain how one comes to produce self-knowledge reports. In this space the claims are made non-inferentially but that does not preclude one's ability to enter the space of reasons while justifying one's claims. From the perspective of the space of reasons, there is *no* epistemic given point a reference to which justifies the rest of the system of beliefs. Rather what Sellars offers is a holistic picture where one claim is justified by its relation to certain other claims. Thus while Sellarsean idea of justification and the related distinction between the space of reason and the space of description provide us with a way of surpassing Boghossian's spurious trilemma, Shoemaker's deflationism seems to be a victim of that trilemma. It is this point that holds the key of the attraction of the Sellarsean compatibilism.

### **Conclusion: A Few Worries**

The first worry that I want to consider is, in requiring that one be able to enter the space of reasons in order to know a claim, is Sellars surrendering to an internalism about epistemic justification that requires us to know all that justifies a claim before knowing that claim? It seems to me that the kind of justification that Sellars requires by entering the space of reasons is different from the kind of justification required in the traditional Cartesian internalist model. This is because unlike the latter, Sellars' conception of justification is 'holistic.' It is also important to remind us of the role that Cartesian internalism played in the shaping of the incompatibility charge. The idea is that, as we noted in the Introduction, a person should know *all* the justificatory conditions *a priori*. This was prompted by the need that all possible errors should be ruled out and in the case of privileged authority of self-knowledge claims, they needed to be ruled out *a priori*. Sellars's insistence that one be able to justify comes from the idea that one be able to play the game. Further, though he insists that one be able to enter the space of reasons in order for her perceptual as well as self-knowledge claims to be knowledge, he never commits himself to the position that one be able to do this justifying *a priori*. Rather, keeping in tune with his basic reliabilist ideas, he understands justification in terms of the ability to draw one's attention to the conditions under which the claim is made in ensuring that an inference to the fact that the reported state-of-affairs is true. It is precisely this point that allows Sellars to satisfy Boghossian's demand that a knowledge proper must involve 'epistemological achievement.'

The second worry that I want to consider is, is there a kind of infallibility that comes about in Sellars's account of self-knowledge? In the «Empiricism and the Philosophy of

Mind» (1963) Sellars insists that for a report of a person to be knowledge, it has to have ‘authority’ in the sense that one can be able to infer the presence of green objects, for example, from the report ‘x is green.’ So it does seem that in order for a claim to be knowledge proper, that claim has to have authority and also has to be true.<sup>7</sup> But this is basically Sellarsean way of adhering to the traditional understanding of knowledge as justified true belief. What is important from our perspective is that even though Sellars insists that for my claim to be knowledge proper I should be able to justify an inference to the presence of the state of affairs reported, such infallibility does not play any role in Sellars’ understanding of the privileged status of self-knowledge claims. Such privilege, as I have argued earlier, consists solely of the fact that I can make a claim (‘report’) about my own propositional attitudes non-inferentially while you will have to infer it from my behavior. However, when it comes to justifying these claims, *both* you and I have to enter the space of reasons, thus inferring the presence of the propositional attitudes from standard conditions. Further, since the reliable conditions and our abilities to recognize them are defeasible, our reliable responses can fail if the conditions are not favorable. As Sellars writes,

...[W]hen Jones candidly says, in response to visual stimulation, «Lo! Here is a red apple,’ it is *likely to be true* that there is a red apple in front of him. I say *likely* to be true, because we all know of various ways in which things can go wrong. Jones is in front of a mirror; the supposed apple is a piece of wax; the illumination is abnormal and the object is really purple; or, there is nothing in front of him, but he has taken LSD, and people have been pounding his ears about red apples. (Sellars, [1975], 1991, 376)

Similarly under unfavorable conditions my non-inferential self-knowledge claim may go wrong. Just as an ice cube may look pink under pink light, I may report a hope as a belief under unfavorable conditions. Thus, it seems to me that Sellars provides us an account in which the non-inferentially privileged self-knowledge claims are not only justifiable but also defeasible. A further merit of this account thus is captured in the fact that it can maintain the distinction privileged in access and privileged in authority. This distinction gives it the defeasible flexibility that allows it to be applied to a wider range of self-knowledge claims.

<sup>7</sup> It should be mentioned here that Sellars (1963) does not differentiate between two factors involved in knowledge that are normally differentiated, namely, the justification and the truth. Sellars seems to refer to both of them by his term ‘authority’. Also Sellars applies his term ‘authority’ to all cases of knowledge without reserving this predicate only for a select group of claims.

**Bibliography**

- ALSTON, W. (1971): «Varieties of Privileged Access» *American Philosophical Quarterly* 8: 223-41.
- BOGHOSSIAN, P. (1989): «Content and Self-Knowledge» *Philosophical Topics* 17: 5-26.
- BURGE, T. (1988): «Individualism and Self-Knowledge» *Journal of Philosophy* 85: 649-63.
- DEVRIES W. & TRIPPLET, W. (2000): *Knowledge, Mind and the Given* Hackett Publishing Co.
- DRETSKE, F. (1997): «Conscious Experience» in N. Block, O. Flanagan & G. Guzeldere (eds.) *The Nature of Consciousness: Philosophical Debates 773-788* Cambridge, MA: MIT Press.
- HEIL, J. (1988): «Privileged Access» *Mind* 97: 238-51.
- MAITRA, K. (2005): «Self-Knowledge: Privileged in Access or Privileged in Authority?» *Southwest Philosophy Review* 21:2: 101-114.
- SAWYER, S. A. (1999): «An Externalist Account of Introspective Knowledge» *Pacific Philosophical Quarterly* 80: 358-378.
- SELLARS, W. «Empiricism and the Philosophy of Mind» in *Science, Perception and Reality* 127-196 California: Ridgeview Publishing Co.
- SELLARS, W. (1991): «The Structure of Knowledge, Lecture II: Minds» in *The Nature of Mind* (ed.) D. M. Rosenthal 372-379 (originally published in 1975).
- SHOEMAKER, S. (1996): *The First-Person Perspective and Other Essays* New York: Cambridge University Press.
- THOMASSON, A. L. (2003): «Introspection and the Phenomenological Method» *Phenomenology and the Cognitive Sciences* 2: 239-54.
- WRIGHT, C. (1998): «Self-Knowledge: The Wittgensteinian Legacy» in C. Wright, B. C. Smith, & C. Macdonald (eds) *Knowing Our Own Minds* 13-45 Oxford: Clarendon Press.