

THE “DÉNOUEMENT” OF “EMPIRICISM AND THE PHILOSOPHY OF MIND”

Keith Lehrer and David G. Stern

I. INTRODUCTION

Wilfrid Sellars spoke of section 59 of his classic paper, “Empiricism and the Philosophy of Mind” (Sellars 1997, first published as Sellars 1956), as the *dénouement* of the myth of Jones, the story about “our Rylean ancestors” (1997, p. 90) that occupies the last part of that paper. Sellars’s myth aims to show humans limited to a language with a vocabulary that speaks of “public properties of public objects” (1997, p. 91) could learn to speak of inner episodes and immediate experience. By section 58, Jones, the protagonist of Sellars’s anthropological science fiction, has developed a theoretical framework that allows him to explain the behavior of others in terms of “thoughts”: postulated linguistic entities, modeled on overt utterances, but akin to the particles of physical theory in that they are theoretical, not observational.

Section 59 reads as follows:

Here, then, is the *dénouement*. I have suggested a number of times that although it would be most misleading to say that concepts pertaining to thinking are theoretical concepts, yet their status might be illuminated by means of the contrast between theoretical and non-theoretical discourse. We are now in a position to see exactly why this is so. 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 others’ behaviour, it is but a short step to the use of this language in self-description. Thus when Tom, watching Dick, has behavioral evidence which warrants the use of the sentence (in the language of the theory) “Dick is thinking ‘p’” (or “Dick is thinking that p”), Dick, using the same behavioral evidence, can say, in the language of the theory, “I am thinking ‘p’”

(or "I am 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 behavior. Jones brings this about, roughly, by applauding utterances by Dick of "I am thinking that p" when the behavioral 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 purely theoretical use has gained a reporting role.*

As I see it, this story helps us understand that concepts pertaining to such inner episodes as thoughts are primarily and essentially *intersubjective*, as intersubjective as the concept of a positron, and that the reporting role of these concepts—the fact that each of us has a privileged access to his thoughts—constitutes a dimension of the use of these concepts which is *built on* and *presupposes* this intersubjective status. My myth has shown that the fact that language is essentially an *intersubjective* achievement, and is learned in intersubjective contexts—a fact rightly stressed in modern psychologies of language, thus by B. F. Skinner, and by certain philosophers, e.g., Carnap, Wittgenstein—is compatible with the privacy of "inner episodes." It also makes clear that this privacy is not an "absolute privacy." For if it recognizes that these concepts have a reporting use in which one is not drawing inferences from behavioral evidence, it nevertheless insists that the fact that overt behavior is evidence for these episodes *is built into the very logic of these concepts*, just as the fact that the observable behavior of gases is evidence for molecular episodes is built into the very logic of molecule talk. (Sellars 1997, §59, pp. 106–7)

In the Introduction to the recent Harvard University Press edition of "Empiricism and the Philosophy of Mind" (Sellars 1997), Richard Rorty aptly sums up the "myth of Jones" in sections 48–63 as

a story which explains why we can be naturalists without being behaviorists, why we can accept Wittgenstein's doubts about what Sellars calls 'self-authenticating non-verbal episodes' without sharing Ryle's doubts about the existence of such mental entities as thoughts and sense-impressions. . . . Sellars's account of inner episodes as having originally been postulated, rather than observed, entities, together with his account of how speakers might then come to make introspective reports (sect. 59) of such episodes, made clear how one could be Wittgensteinian without being Rylean. Sellars showed how one could give a non-reductive account of 'mental event' while nevertheless eschewing, with Wittgenstein, the picture of the eye of the mind witnessing these events in a sort of immaterial inner theater. (Sellars 1997, pp. 6–7)

Rorty rightly emphasizes the importance of section 59 for Sellars's overall argument; Sellars himself described that section as the "dénouement of the first episode in the saga of Jones" (Sellars 1997, §58, p. 105). According to the *Oxford English Dictionary*, the term "dénouement" is derived from a French word, "dénouer," to untie, or un-knot, and means "unravelling, the final unravelling of the plot in a drama, etc.; the catastrophe; the issue of a complication, difficulty, or mystery." The knot that Sellars aims to unravel in section 59 concerns the place of knowledge of one's own inner states in the philosophy of mind. In particular, it has to do with the question of the role of the privacy of experience, or "privileged access" to one's own mental states, within the overall argument of that paper.

In "Empiricism and the Philosophy of Mind" (EPM) Sellars advanced a theory of meaning and philosophy of mind in which overt speech is the model for a theory of thought and one component of perception and sensation. Thought is introduced as a theory to explain overt behavior on the basis of a model of overt speech. Perception and those aspects of sensation concerning the appearance and look of things involve a thought, endorsed in varying degrees and ways, about how things are. It is not the purpose of this paper to reiterate this theory, which Sellars lucidly articulated, but to deal with a problem raised by the account of thoughts as theoretical entities. The problem is that we learn to report on our thoughts and perceptions in a reliable way, as Sellars acknowledges, but *how* we do this is left mysterious in Sellars's writings. It just turns out, he says, that we are able to do this. A better explanation is needed within the framework of Sellars's theory, for the ability to report on the existence of our thoughts and perception in a reliable manner is remarkable, and unless an explanation of how this occurs is forthcoming within the framework of Sellars's philosophy, comfort is offered to the arch opponents of Sellars, the defenders of the myth of given. For the latter will argue that the reason we learn to report reliably on our thoughts is that we have immediate knowledge of them, which we only need to learn to articulate in a public language. To meet this objection, a Sellarsian must offer an alternative account of how we come to have this remarkable ability. Sellars does provide a very rapid outline of his response to this objection in section 59 of EPM, but it is so rapid that it has proved hard for his readers to see what he was suggesting, let alone work out how it might be developed into a systematic account of the ability in question.

In Part Two we turn to Sellars's much more detailed exposition of his view about how the ability to report on our thoughts arises, as set out in an extensive correspondence with Hector-Neri Castañeda on the philosophy of mind. The correspondence was never formally published but it was widely circulated by Sellars and Castañeda and is now available on the World Wide Web (Sellars and Castañeda 1961-62). The typescript of the correspondence, which consists of 47 single-spaced pages, begins with a letter from Castañeda to Sellars, dated March 6, 1961. There was a back and forth exchange of eight more letters, four replies from Sellars and four more letters from Castañeda to Sellars. In his first letter, Castañeda says that he considers himself to be in fundamental agreement with Sellars's approach in EPM, but it soon becomes clear that while he may have accepted many of Sellars's arguments, he remained wedded to a much more traditional conception of subjectivity and privileged access. As a result, Castañeda's friendly questions and proposals lead Sellars to set out what he takes to be the obvious consequences of the position set out in EPM §59, effectively providing an extremely informative exposition of its principal claims.

The correspondence is particularly valuable for the way in which it forces Sellars to clarify and enlarge on his conception of privileged access. Sellars's theory of thoughts makes use of materials taken from behavioristic learning theory, on which a person's acquisition of the ability to think about and report on his or her thoughts is conceived of as a matter of training the person to make a non-rational conditioned response. In other words, Sellars treats the initial stages of learning to say "I am thinking that p" when I have the thought that p not as a matter of learning an inference, but rather as being conditioned to say the appropriate words under the right circumstances. Furthermore, he claims that the difference between someone who has merely been conditioned to respond in this way, and someone whose utterance of "I am thinking that p" expresses direct self-knowledge "is *not* that in the latter case the statement *isn't occurring as a conditioned response*. It is. The difference is that in the latter case the conditioning is itself caught up in a conceptual framework" (Sellars to Castañeda, 3 April 1961, p. 7).

While Sellars's detailed replies to Castañeda's questions certainly help to spell out the character of the theory he was proposing in EPM, and also cast light on his thoroughgoing departure from traditional theories of privileged access, the precise nature of the relationship between the causal process of conditioning and the knowledge it yields is left unspecified. In Part Three we build on

the materials provided by Sellars, offering a more detailed explanation of how the conditioned response results in direct noninferential knowledge of our own thoughts. We do this by applying Sellars's own account of the meaning of overt speech to the case of inner thoughts. The basis of the explanation is a process of quotation and disquotation of spoken words and the internal verbal episodes in the language of thought.

II. SELLARS AND CASTAÑEDA ON PRIVACY AND PRIVILEGED ACCESS

In 1956, the year in which "Empiricism and the Philosophy of Mind" was published, Hector-Neri Castañeda heard Sellars give the lectures on which the paper was based. In March 1961, he wrote to Sellars that he had been able to read for the first time the paper that, he said, "so much impressed me when I heard it exactly five years ago." He went on to say that

what you say about theories and models seems to me substantially sound, and your proposal to consider inner episodes as theoretical entities is very insightful. But in the way you propose it seems to me to do injustice to self-knowledge. You regard the basic language, the language of the model, to be behavioral language, so that my own mental states as well as yours are for me theoretical entities. Since this clearly won't do, you correctly suggest that first-person statements about thoughts acquire another role: a reportive one. This you claim on pp. 321 and 327 explains the privacy of the mental. But it is not clear how this claim is justified. . . .

Your whole discussion of how the reportive role (and the privacy) comes about is:

Dick, using the same behavioral evidence, can say, in the language of the theory, "I am thinking that p." And it now turns out that Dick can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe his overt behavior. Jones brings this about, roughly, by applauding utterances by Dick of "I am thinking that p" when the behavioral 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 purely theoretical use has gained a reporting role. (Sellars 1997, §59, pp. 106-7)

What exactly is what Jones reports in the new use of "I am thinking that p"? How is it that he can make correctly such a statement without observing his behavior? It is not easy to see how on your

view these questions can be answered, and if they are not answered it is difficult to see what exactly your view is accomplishing. (Sellars and Castañeda 1961-62, 6 March 1961, pp. 1-2)

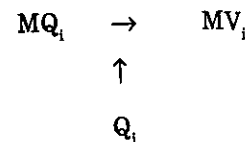
The remainder of Castañeda's letter offered an attempted elaboration of Sellars's account of the reporting use of "I am thinking that p." However, he assumed that the report would have to be based on some aspect of Dick's behavior, and so could not see how to account for first-person knowledge of unreported inner episodes. As a result, Castañeda was led back to thinking that the myth of the given was true. He concluded that Sellars's account of our knowledge of our thoughts was appropriate for an understanding of others' thought and behavior, but not one's own: one could not do without non-theoretical mental entities whose "*occurrence entails my knowledge of them.*" Consequently, "awareness has to be added to the basic language" (Sellars and Castañeda 1961-62, March 6 1961, p. 4. Underlining in the original).

Sellars began his reply by acknowledging that the passage in question was "terse and skeletal in the extreme. It obviously is not an effective way of making the point I wished to make, since you have not understood it. Let me make another try" (Sellars and Castañeda 1961-62, 3 April 1961, p. 5). After recapitulating the story of the emergence of the Jonesean theory of thoughts, he sets out his response as follows:

9. Dick, understanding the [Jonesean] theory [of thoughts], now has two *inferential* routes to the idea that someone (it may be himself) is thinking "There is a caribou ahead; I shall slow down and take cover."

- (1) The non-verbal behavior characteristic of that stage of a hunt.
- (2) The overt verbal behavior which is the natural culmination of the thought.

10. The stage is now set for the interpretation of the passage of EPM which you quote at length and puzzle over. The important thing to note is that the *core* of Dick's learning to report what he is thinking is a matter of his acquiring a tendency (*ceteris paribus*) to respond to his thought that-p by saying "I am thinking that-p." Everything hinges on the force of the word "respond" in this connection. *It is being used as a technical term borrowed from learning theory.* The following diagram will help clarify matters:



where Q_i is a thought that-p, MQ_i is a meta-thought "I am thinking that-p" and MV_i is a meta-statement "I am thinking that-p."

11. The connection between Q_i and MQ_i is *in the first instance a conditioning* and not an *inference*. As such it presupposes neither an awareness on Dick's part that he is thinking that-p nor any recognition on Dick's part that the circumstances are such as would usually involve his thinking that-p. It requires only that the reinforcer (applauder), in this case Jones, correctly infer that Dick is thinking that-p and, given that Dick *happens* to say "I am thinking that-p", applauds. . . . the decisive feature is that the connection between Q_i and $(MQ_i - MV_i)$ is a direct non-rational S-R connection. Certainly this S-R connection exists within a rich conceptual context, but unless it existed as an S-R connection, there could be no direct non-inferential self-knowledge. . . .
13. Now the important difference between a person who has merely been conditioned to respond to his thought that-p by saying "I have the thought that-p" and a person whose statement "I have the thought that-p" *expresses direct self-knowledge* is *not* that in the latter case the statement *isn't occurring as a conditioned response*. It is. The difference is that in the latter case the conditioning is itself caught up in a conceptual framework. Compare the case of the child who has *merely* been conditioned to respond to green objects (in standard conditions) by "This is green," as contrasted with a child whose utterance of "This is green" expresses *direct non-inferential knowledge*. Indeed the key to the account I have given of the direct non-inferential knowledge of inner episodes is the apparatus I developed in discussing the status of non-inferential perceptual knowledge. (See the discussion of "epistemic authority" in the section entitled "Does Empirical Knowledge have a Foundation?" [Sellars 1997, Part VIII, §§32-38].) I simply assumed that anyone who worked through the latter would see how the relevant distinctions applied to direct self-knowledge. . . .
15. The above type of account explains the "privileged access" a person has to his own inner episodes. For (although worlds are conceivable in which this is not the case) only the person who

has a thought that-p can respond to it (in the manner discussed above) with the thought that he has the thought that-p. (Sellars and Castañeda 1961-62, 3 April 1961, pp. 6-8)

Sellars concludes by saying that he would be interested in learning more about what Castañeda has in mind by "awareness": "If what you mean is direct non-inferential knowledge of particular matter of fact, then my theory *insists on awareness and offers an account of it*. But perhaps my attempt to give an account which avoids 'the myth of the given' has not persuaded you" (Sellars and Castañeda 1961-62, April 3 1961, p. 8).

Castañeda's response, written ten days later, begins by thanking Sellars for his extensive discussion, saying that "some of the things you say are (it seems to me) worth adding to your famous essay" (Sellars and Castañeda 1961-62, April 13 1961, p. 9). However, some problems remained: "1. I am [not]¹ sure I understand your proposition that self-knowledge involves a me-a-thought which is a conditioned response. And I would like to have an argument to show that that is so" (Sellars and Castañeda 1961-62, April 13 1961, p. 9). After raising a number of queries about the relationship between the various terms in Sellars's diagram, he turns to his principal concern: "5. But what perplexes me most is the same I mentioned in my first letter. . . . What does it mean to say that Dick has learned to react by MQ to Q?" (Sellars and Castañeda 1961-62, April 13 1961, p. 10). Castañeda sets out his perplexity by way of a story: imagine that Dick shows all the signs of a person with a specific infection in his left kidney. He is taught the theory of viruses so that he can infer that he has such a colony. "Now, what does it mean to say that Dick can be conditioned to react to the colony of viruses . . . ?" According to Castañeda, it would be a matter of uttering the appropriate sentence while perceiving some of the signs in question.

The important point here is that S-R connection, if you want to say that, is not colony-utterance, but *signs of the colony-utterance*. . . . Now, if that is what happens in the case of the theoretical entity Q and relations, we would have to say that Dick is conditioned to utter "I am thinking that-p" . . . on inspection of his behavior and circumstances. But that is just what we do not want. . . . In other words, as I said in my previous letter, it seems to me that your analogy with theoretical entities will do only for another person's thoughts, etc. (Sellars and Castañeda 1961-62, April 13 1961, pp. 10-11)

In his reply, dated November 14, Sellars responds to the discussion of the virus story by focusing on the claim at the beginning of the passage I have just quoted. He writes:

This is simply a *petitio*. You are assuming that we can't train Dick to say "I have a colony of viruses," when he has one, in a way which is not mediated by *his* noticing "signs of the colony." And you are assuming this because you are assuming a general principle (of which your example is intended to be an illustrative instance) to the effect that

(P) To train a subject, S, to say, "I am in state \emptyset " when he is in state \emptyset , S must either observe that he is in state \emptyset (where this is possible) or observe that he is in a state which is a sign of state \emptyset .

But, Sellars replies, this is not always true: "for some states \emptyset (but by no means all) we can bring about a connection between *being in state \emptyset* and saying 'I am in state \emptyset '" (Sellars and Castañeda 1961-62, November 14 1961, p. 14).

9. I believe that I have put my finger on the fundamental sources of your puzzlement. If I am right about this, you should now see exactly what I had in mind in EPM when I concluded that the correct contrast between other people's mental states and our own is that between *theoretical entities* and *theoretical entities plus*, rather than, as you suggest, between *theoretical entities* and *non-theoretical entities*. It must be remembered that while we have *direct non-inferential knowledge* of our mental states, we do not *observe* them. We *have* them, and we *know that we have them*. Needless to say, the conditioning discussed in this letter, while it is a *necessary* component in the ability to know what is going on in one's own mind, is not *sufficient* to account for it. For a discussion of the more that is involved, I refer you to section 13 of my letter [quoted above in full, p. 207], and to the sections of EPM to which it refers. (Sellars and Castañeda 1961-62, November 14 1961, p. 15)

Sellars's discussion of "the more that is involved" had emphasized that the difference between someone who has merely been conditioned to respond to his thought that-p by saying "I have the thought that-p" and someone who expresses direct self-knowledge by uttering those words is that "in the latter case, the conditioning is itself caught up in a conceptual framework." But what, precisely, is the framework in question, and how is the conditioning caught up in it? Four years later, David Rosenthal, who was writing a

dissertation on Sellars's views on intentionality at Princeton, wrote a letter to Sellars in which he asked him what he had meant by that very expression in his correspondence with Castañeda² (Rosenthal and Sellars 1972, p. 462).

Sellars's initial reply did little more than reiterate what he had said in the original letter, namely to refer back to the account of observational knowledge he had given in Part VIII of "Empiricism and the Philosophy of Mind": "the point I had in mind was intended to be the exact counterpart of the additional condition I lay down for 'Lo! this is green' to count as observational knowledge" (Rosenberg and Sellars 1972, September 3, 1965). Sellars's account in "Empiricism and the Philosophy of Mind" of the conditions under which "This is green" counts as observational knowledge requires that (1) "it must be a *symptom* or *sign* of the presence of a green object in standard conditions" and (2) "the perceiver must know that tokens of 'This is green' are symptoms of the presence of green objects in conditions which are standard for visual perception" (Sellars 1997, §35). Rosenthal wrote back to ask whether these two conditions—that the speaker's utterance must be reliable, and the speaker knows that it is reliable—were really all that Sellars meant by saying that "the conditioning is itself caught up in a conceptual framework"³ (Rosenthal and Sellars 1972, p. 471).

In his reply, Sellars made it clear that rather more is involved:

As for your concluding questions, you are again on the right track with respect to my views on direct self-knowledge. The only point I would want to add is that the avowal "I have a thought that-p" (1) asserts the occurrence of a thought that-p, (2) is a reliable symptom of the occurrence of such a thought, and (3) gives overt expression to a metathought (inner episode) which is an "I have a thought that-p", exactly as the candid assertion "it is raining" gives overt expression to a non-metathought which is an "it is raining". In other words, we post-Jonesians explain the connection between thoughts that it is raining and reports that one has a thought that it is raining, in terms of a connection between thoughts that it is raining and apperceptive metathoughts that one has a thought that it is raining.⁴ (Sellars and Rosenthal 1972, p. 480)

In this last answer to Rosenthal's persistent questions, Sellars acknowledged the importance of an explicit appeal to a meta-thought, a thought that one has the thought in question, in the account of direct self-knowledge he had given earlier. In the next section of this paper, we explore this idea in greater detail.

III. SELLARS ON DOT QUOTES IN THOUGHT AND SENSATION

What we still need to add to the account that Sellars himself provided is an explanation of the relationship between the internal theoretical state Q_i and the meta-thought MQ_i . As we have seen, Sellars does say that MQ_i is a meta-thought "I am thinking that-p" that leads to the meta-statement MV_i "I am thinking that-p," which does not assume that the subject observes or perceives Q_i . The question arises exactly why MQ_i and MV_i are something at the meta-level. Moreover, one wonders how one might explain the relationship between Q_i and MQ_i beyond simply saying that it is a matter of conditioning, for there is a more direct connection manifested by the fact that Q_i occurs in MQ_i . The answer is the loop of quotation and disquotation. When Dick is conditioned to respond to Q_i he is being conditioned to respond to an internal episode of linguistic tokening. To be sure the tokening is a theoretical state, but theoretical states are real enough, and one may be conditioned to respond to it. However, the conditioning in this case is a conditioning that uses the state in a special way, namely, by reacting to it as if one were placing it in quotation marks at the metalevel and then disquoting it to the dot quoted form.

We can fill in this account on a Sellarsian approach to mind and language in which speech is the model of thought, and the meaning of the former is the model of the content of the latter. So let us consider briefly the theory of meaning of speech Sellars offered us. Meaning sentences such as

"Rot" in German means red

or

"Das ist rot" in German means that is red

are intentional in form since the words following the word "means" in both sentences are used rather than mentioned, but the use is a special one. They are used to exhibit a role in the language of the speaker. In fact, the words "red" and "that is red" are used as exemplars of words playing a linguistic role in the language of the meaning sentences, in this case, English. It is not difficult to see that Sellars is really a closet extensionalist when one recalls that the role an expression plays in a language can be characterized extensionally in terms of its function in language entry transitions, language to language transitions, and language exit transitions. Transitions are extensional, and meaning is the role of the word

analyzed in terms of these transitions. The meaning sentence only uses the words describing the role as exemplars of a class of words playing the role.

This theory of meaning becomes more explicit with the introduction of dot quotes in later works, though the idea is fully contained in "Empiricism and the Philosophy of Mind." The idea is that you can use a word to describe a role played by the word in a language by saying, in effect, that the role of the word is the role played by this word. Dot quotes are introduced to pick out the role, so that •red• stands for the role played by the word "red" in the language being used. Thus, the sentences above might be made more explicit when rewritten as

"rot" in German is a •red•

and

"Das ist rot" in German is a •that is red•.

This shows that the function of meaning sentences is to classify an expression as belonging to a class of sentences playing the role played by the dot quoted sentence in the language used. It is crucial to notice, however, that dot quoted expressions are quoted expressions, though not only quoted sentences, exhibited as exemplars of sentences playing the role that they play.

In effect, therefore, meaning sentences and the dot quoted clarification of them involves a loop of quotation and disquotation to describe the linguistic role of the sentence. This becomes obvious when the meaning sentence remains within the home language. The sentences

"red" in English means red

and

"That is red" in English means that is red

which become

"red" in English is a •red•

and

"That is red" in English means •That is red•

reveal the way in which the sentence is quoted, disquoted, and dot quoted to describe the role of the English expressions in English. Lehrer (1996; 1997, pp. 153–183; forthcoming) has called the process *exemplarization*. It is the process of using something as an exemplar to represent itself as well as other things of the same

sort. The process of quotation and disquotation used to describe the linguistic role is extensional as is the role described. Moreover, once the meaning rubric is mastered, quotation and disquotation suffices for a minimal account of meaning. That is the secret behind the explanation of how we can learn to reliably report the meaning of overt speech, which is now apparent, and the content of thought and perception, which now requires elucidation.

To offer the needed explanation of how we can learn to reliably report our own thoughts and perceptions, we need only take quotation and disquotation inside. We shall find internal exemplarization sufficient to explain the reliability of our reports of internal states. Sellars, in his myth of Jones, which is more than a myth, asks us to suppose the postulation of a theory of thoughts based on the model of overt speech. Thoughts are theoretical verbal episodes based on the model of overt speech episodes, which can be used to explain behavior in a way analogous to the way in which overt speech explains behavior. There are many refinements of this postulation that, though important to the viability of Sellars's proposal, we shall ignore in order to deal with the problem of how we can so easily learn to report the content of our own thought. The explanation is that all we need to do is to learn to apply the same process of quotation and disquotation to our internal verbal episodes as we do to the external ones, and exemplarization explains the rest.

Suppose that there is an internal verbal thought episode "That is red." How can I report on the content of the verbal episode? I must learn to quote it and then the loop of exemplarization suffices. Consider the sentence

"That is red" in my thought has the content that is red
which in the dot quoted version is

"That is red" in my thought is a •That is red•.

There is nothing more to understanding the sentence than to understand quotation and disquotation yielding exemplarization. Exemplarization enables me to report the minimal content of my thought, namely, "That is red." I must learn to quote my internal verbal episodes, which it turns out I can do. All that remains for me to report the content of my thought, however, is disquotation and exemplarization. These I already understand from my understanding of exemplarization in overt speech. Exemplarization and disquotation explain how I can reliably report the content of thought once I have learned how to apply quotation to the internal episodes. Exemplarization is all there is to it.

IV. CONCLUSION

Philosophers often draw a sharp distinction between "doing philosophy" and "doing the history of philosophy." Sellars himself had little respect for this distinction, believing that the "history of philosophy is the *lingua franca* which makes communication between philosophers, at least of different points of view, possible. Philosophy without the history of philosophy, if not empty or blind, is at least dumb" (Sellars 1967, p. 1). While "Empiricism and the Philosophy of Mind" was primarily a contribution to contemporary epistemology and philosophy of mind, the paper begins with an appeal to Kantian and Hegelian precedents, and the psychological nominalism Sellars defends there can be seen as a restatement of Kant's insight that "intuitions without concepts are blind" in the idiom of the logical empiricism of the 1950s. In this paper, we have followed in Sellars's footsteps, drawing on both the historical record of Sellars's exposition of his own ideas and our Sellarsian development of his argument in order to unravel the mystery of the tightly knotted web of argument in section 59 of "Empiricism and the Philosophy of Mind."

University of Arizona and University of Iowa

NOTES

This paper is based on two papers that the co-authors wrote independently for a conference on Wilfrid Sellars's "Empiricism and the Philosophy of Mind" held in Dunabogdány, Hungary, in October 1996. We would like to thank Professor J. C. Nyiri for organizing an extremely successful series of discussions, and the conference participants for their comments on our papers. Stern would also like to thank the Alexander von Humboldt Foundation and the Department of Philosophy at the University of Bielefeld for fellowship support that made possible the completion of this paper. Section III was written by Lehrer; the rest of the paper, except for the summary of Section III provided in Section I, was written by Stern.

1. We have added the "not" here in order to make sense of what Castañeda is saying, as the rest of the letter sets out why Castañeda does not understand Sellars's proposition. It is possible that the word was overlooked when the letter was transcribed.

2. Rosenthal put his question as follows:

A point at which I was conscious of not understanding your views as well as I would like to involves the nature of the so-called reporting

role which uses of sentences in the language of Jones' theory come to have. I have gone through, in particular, the letter of yours to Castañeda (April 3, 1961, #10) in which you discuss this, and the passages you refer to, but I find myself unclear on what, for the language of thoughts, is meant by your statement that "the conditioning is itself caught up in a conceptual framework" (#13). I would be immensely grateful for any help you could give me with this. (Rosenthal and Sellars 1972, p. 462)

3. Actually, Rosenthal phrased the question much more precisely:

If I understand your account, it runs as follows. If (a) "I have a thought that-p" expresses direct self-knowledge, then both (b) the uttering was the result of a conditioned response and (c) "the conditioning is itself caught up in a conceptual framework" (Castañeda Correspondence, p. 6 [April 3, 1961, #13]). But if (c), then (d) the person who utters "I have a thought that p" must recognize the speech act as a report (EPM [pp. 75/298 ff. §§35 ff.]). Finally if (d), then (e) the speaker recognizes that such reports are reliable symptoms of the person who reports having the thought that-p (EPM [pp. 75-76/298]). If this is correct so far, then my question is whether (e) is a sufficient condition for (c), and whether (c) and (b) are jointly sufficient for (a). Put differently, is any more involved in what you call meta-thinking (Castañeda Correspondence, p. 17 [Dec. 8, 1961]) than conditions (b) and (e)? (Rosenthal and Sellars 1972, p. 471).

4. Sellars went on to make it clear that this account went beyond the terms set out in Rosenthal's previous letter: "Thus I cannot agree that nothing more 'is . . . involved in what [I] call meta-thinking . . . than conditions (b) and (e)'" (Sellars and Rosenthal 1972, p. 480).

LIST OF REFERENCES

- Lehrer, Keith. 1996. "Skepticism, Lucid Content and the Metamental Loop." In *Philosophy and Cognitive Science: Categories, Consciousness and Reasoning*, ed. Andy Clark, Jesús Ezquerro, and Jesús M. Larrazabal. Dordrecht and Boston: Kluwer Academic Publishers, pp. 73-94.
- Lehrer, Keith. 1997. *Self-Trust: A Study of Reason, Knowledge and Autonomy*. Oxford and New York: Oxford University Press.
- Lehrer, Keith. forthcoming. "Meaning, Exemplarization, and Metarepresentation." In *Metarepresentation*, ed. Dan Sperber. Oxford and New York: Oxford University Press.
- Rosenthal, David M. and Wilfrid Sellars. 1972. "The Rosenthal-Sellars Correspondence on Intentionality." In *Intentionality, Mind,*

- and *Language*, ed. Ausonio Marras. Urbana: University of Illinois Press, pp. 461-503.
- Sellars, Wilfrid. 1956. "Empiricism and the Philosophy of Mind." In *Minnesota Studies in the Philosophy of Science*, vol. 1, ed. Herbert Feigl and Michael Scriven. Minneapolis: University of Minnesota Press, pp. 253-329.
- Sellars, Wilfrid. 1967. *Science and Metaphysics: Variations on Kantian Themes*. London: Routledge and Kegan Paul. Reprinted in 1992 by Ridgeview, Atascadero, Calif.
- Sellars, Wilfrid. 1997. *Empiricism and the Philosophy of Mind*. Boston: Harvard University Press.
- Sellars, Wilfrid and Hector-Neri Castañeda. 1961-62. "Correspondence between Hector Castaneda and Wilfrid Sellars on Philosophy of Mind." Unpublished typescript distributed by Hector-Neri Castañeda and Wilfrid Sellars. Quotations will give the date of the letter and the page number in the typescript. An edited transcript of the correspondence is available from the following web address, where it is part of Andrew Chrucky's "Problems From Wilfrid Sellars" page: <<http://csmacslab-www.cs.uchicago.edu/philosophyProject/sellars/corr.html>>

HISTORY OF PHILOSOPHY QUARTERLY



POLICY

The *History of Philosophy Quarterly* welcomes articles in English on any aspect of the history of philosophy. However, only self-sufficient articles will be published, and not news items, book reviews, critical notices, or "discussion notes" (short or long). The journal plans to specialize in papers of a particular sort, those manifesting a strong interaction between contemporary and historical concerns. Either historical material should be exploited to deal with matters on the agenda of current discussion, or present-day concepts, methods, distinctions, arguments, etc., should be used to illuminate historical questions.

MANUSCRIPTS

Contributions may be as short as 3,000 words or as long as 8,000. All manuscripts should be typewritten with wide margins, and at least double spacing between lines. Footnotes should be used sparingly and should be numbered consecutively, typed with wide margins and double spacing, placed at the end of the paper. Submissions should always be made in duplicate. The journal will consider only papers whose authors certify that, while under consideration with us, their papers will not be submitted elsewhere for consideration.

COMMUNICATIONS

Articles for publication and all other editorial communications and enquiries should be addressed to: Professor Catherine Wilson, Editor, *History of Philosophy Quarterly*, Department of Philosophy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada. Other correspondence not dealing with subscriptions should be addressed to Ms. Dorothy Henle, Operations Manager, *History of Philosophy Quarterly*, North American Philosophical Publications, Inc., 1151 Freeport Rd. (No. 154), Fox Chapel, PA 15238. (For subscriptions, see below.)

REPRINTING AND COPYING

The *History of Philosophy Quarterly* regrets that it cannot make author's reprints available. However, authors have the journal's permission to reproduce limited numbers of their contributions for the use of colleagues and students. (Two copies of the relevant issue will be provided gratis for this purpose.) While the journal holds the copyright on materials published in its pages, it routinely accords contributors permission to reprint in books or anthologies authored or coauthored by themselves. All other requests to reprint should be addressed to the editor. The journal gives teachers and their institution blanket permission for reproducing individual articles for class use in limited numbers (up to 100 copies) at a fee of \$1 per copy, by check payable to the journal.

SUBSCRIPTIONS

The journal is published four times a year in January, April, July, and October. The subscription price for 1999 is \$185 to institutions, \$40 to individuals (back issues: institutions \$44, individuals \$20). All correspondence regarding subscriptions, renewals, back orders, and related matters regarding the distribution of the journal should be addressed to the subscription agent: Philosophy Documentation Center, Bowling Green State University, Bowling Green, OH 43403-0189 USA; or e-mail to pkswope@bgnet.bgsu.edu.