Received: 16 October 2024 | Accepted: 21 December 2024

DOI: 10.1111/ejop.13039

ORIGINAL ARTICLE

European Journal of Philosophy WILEY

Thought, consciousness, and the given

David Rosenthal

Program in Philosophy, Graduate Center, City University of New York, New York, NY, USA

Correspondence
David Rosenthal,
Email: davidrosenthal1@gmail.com
Program in Philosophy
Graduate Center, City University of New York
Postal Address: 425 Riverside Drive, New York, NY 10025-7731, USA.

Abstract

How do we come to understand the nature of the thoughts that we and others have? And how do we come to have the conceptual resources needed to formulate such understanding? Many would say we understand the nature of thoughts simply by being subjectively aware of our own conscious thoughts. But it is unclear how consciousness could, on its own, provide the conceptual resources required for such understanding. An alternative account holds that we understand the nature of thoughts in a third-person way, by appeal to the speech acts that can express those thoughts. Such an account readily explains how we come to have the required conceptual resources. Wilfrid Sellars, in "Empiricism and the Philosophy of Mind" and elsewhere, developed a view along those lines, appealing to considerations related to and in the spirit of the foregoing concerns. I'll describe and defend that view against two main objections. And I'll argue that any picture of consciousness on which it could reveal the nature of thoughts is independently untenable, and also that such a picture underlies what Sellars denounced as the Myth of the Given. In closing I explain how, given that we understand the nature of thoughts in such a third-person way, some of our thoughts come to be conscious.

KEYWORDS
consciousness, intentionality, speech acts, the given, thought, Wilfrid Sellars

1 | INTRODUCTORY

We initially come to know about physical reality by perceiving and interacting with the physical objects that surround us. But physics then tells us that those objects are congeries of subatomic particles, and tells us other things

that depart even more dramatically from our commonsense picture. It's inviting to see that contrast in terms of Aristotle's distinction between how things are in themselves and how they are best known to us (e.g., *Posterior Analytics* 71b32).

Aristotle's distinction is typically invoked in connection with our grasp of physical reality. But it turns out to be useful also in connection with how we understand the nature of mental states. Thus Wilfrid Sellars has argued that "visual impressions are prior in the order of being to concepts pertaining to physical color," but our concepts of physical colors "are prior in the order of knowing to concepts pertaining to visual impressions" (Sellars 1965: 57). We could not have our concepts of the colors of physical objects without having perceptual sensations of those colors. But we could not, Sellars argues, come to understand what those perceptual sensations are without first having concepts of the physical colors of the objects we see.

In discussing perceptual sensations, Sellars explicitly appeals to the distinction between what is better known to us and how things are in themselves. I have argued elsewhere (Rosenthal 2010, 2024) that doing so is important for understanding the nature of perceptual sensations, though I have also argued that the way Sellars applies that distinction to perceptual sensations is not quite right as it stands, and I have urged an appropriate adjustment (Rosenthal 2016).

But though it is less explicit, Sellars in effect invokes that distinction also in his account of the nature of thoughts and their intentional contents. In rough summary, he holds, with many others, that the intentional content of thinking is prior in the order of being to the speaker's meaning of speech acts that express those intentional states. Speaker's meaning here is what a speaker means in performing a speech act, as against the semantic meaning of the words that occur in the speech act. Sellars' largely novel contribution is to argue forcefully that, for the order of knowing, the speaker's meaning of speech acts is prior to the intentional contents of thoughts. In what follows I argue in support of that account of the relation between thoughts and speech acts.

In section II, I raise a problem about how we come to have the conceptual resources required to represent thoughts and their intentional contents. There is a picture of consciousness on which the subjective appearances of our conscious thoughts might by themselves seem to generate those conceptual resources. But I'll argue that consciousness cannot do that on its own, and that the picture of consciousness on which it seems to do so is indefensible.

In section III, then, I argue that the best account of how we do come to have those conceptual resources is instead that advanced by Sellars in "Empiricism and the Philosophy of Mind" (Sellars 1956; henceforth "EPM"). On that account, we come to have those resources by coming to explain in a proto-scientific way how speech acts occur. And all that is wholly independent of consciousness. Moreover, though Sellars himself said little about consciousness as such and gave no account of it, I'll show in section IV how his account of thought and speech invites explaining consciousness by a higher-order-thought theory (e.g., Rosenthal 2005).

Sellars often castigated a view he called the Myth of the Given (first in Sellars 1956: 267). In section V I argue that what he called the Myth of the Given reflects, and in many cases rests on, the picture of consciousness rejected in section II, on which the consciousness by itself generates our understanding of thoughts.

In section VI I consider two popular objections to an account of the sort Sellars gives, one that appeals to a Gricean account of speaker's meaning, the other a rejection of the content holism that figures in Sellars' functionalist account of both speaker's meanings and thoughts. I argue that both objections rely on the picture of consciousness that figures in the Myth of the Given. I conclude in section VII with an explanation of how it is that thought and other mental states sometimes come to be conscious.

2 | UNDERSTANDING THE NATURE OF THOUGHTS

Every meaningful speech act expresses a thought whose conceptual content corresponds to the semantic content of the speech act. For a speech act to express a thought, the thought must come first; meaningful speech cannot occur without prior thinking that the speech act expresses. And thinking often occurs without speech. Thinking is both temporally prior and basic.

That's our commonsense picture, and it is correct. And because every speech act expresses some prior episode of thinking, it is natural to see thinking as in some way the source of the semantic content of speech acts. As Roderick Chisholm put it in his published correspondence with Sellars, "[t]houghts would be intentional even if there were no linguistic entities, … [but] nothing would be intentional were it not for the fact that thoughts are intentional" (1958: 533).

John Searle echoes Chisholm's remark. Meaning something by saying it cannot "stand on its own in the way that believing" something can (1983, 29). But Searle goes farther, holding also that "the direction of logical analysis is to explain language in terms of [the] Intentionality [of the mental]" (1983: 5). By 'logical analysis' Searle means conceptual analysis; one can understand what it is for speech acts to mean something only by appeal to what it is for thoughts to have intentional content.

Searle does concede the usefulness of "our prior knowledge of language as a heuristic device for explanatory purposes…. The direction of pedagogy is to explain [the] Intentionality [of thought] in terms of language" (1983: 5). But he insists that this is merely heuristic; "speakers' meaning," he maintains, must "be entirely definable in terms of … forms of Intentionality that are not intrinsically linguistic" (1983: 160).

Sellars in effect concurs with the modest claim that meaning cannot stand on its own in the way that believing can. Thoughts are in that way prior to speech. But Sellars rejects the additional claim that we must understand the nature of speech by appeal to the nature of thinking. As I'll explain in section III, Sellars holds, to the contrary, that we initially understand the nature of thoughts by appeal to the nature of speech. That appeal to the nature of speech is not a mere heuristic; it is basic to the way we understand the nature of thinking itself. Searle's inference from the ontological priority of thought to its conceptual priority conflates what is it to be a thought with how we know about thoughts, thereby running afoul of the Aristotelian distinction described above.

To understand the nature of thinking, we must have conceptual resources that enable us to describe and characterize thinking in respect of its content. So a question arises about how we come to have those conceptual resources. If understanding the nature of thinking were prior to understanding the nature of speech, as Searle maintains, it's not obvious how we do come to have those resources. Those resources are required for understanding the nature of thoughts. So if it's unclear how we come to have those resources, it's equally unclear how we come to understand the nature of thinking. We would have an explanatory dead end.

Many hold that there is a straightforward answer to these questions. We understand the nature of all our mental states, thoughts included, by way of consciousness. Our first-person access to our conscious thoughts tells us what thinking is, and tells us what the intentional content of each thought is. And if consciousness does all this, it must also by itself somehow provide the conceptual resources needed to characterize thoughts and their intentional contents. On this view, there is no problem about how we come to understand the nature of thinking and its conceptual content, and so no problem about how we come to have the conceptual resources needed to do so.

This explanation is arguably endorsed by many, though it is not often stated explicitly. We understand the nature of thinking and its intentional contents by the way consciousness reveals our conscious thoughts.

This idea involves two claims, which it is crucial to distinguish. It is plain that consciousness does reveal our thoughts as states of a particular type, and also in respect of their specific intentional contents. So however it is that consciousness operates, it must be able to deploy the conceptual resources required to characterize thoughts in that way. That's one claim.

But the question was how we come to have those conceptual resources in the first place. And the ability of consciousness to use suitable conceptual resources to characterize our thoughts does not address the question of how we come to have those conceptual resources. Because it may not seem obvious how we do come up with those resources, it may be tempting to think that consciousness somehow does so on its own. But that only pushes the explanatory demand back a step. We would still need to explain how consciousness does that. Without explaining that, we would again have an explanatory dead end, in effect simply saying that have those resources, though we cannot say how.

But it is in any case not credible that consciousness generates the needed resources on its own. What could it be about consciousness that generates a conceptual understanding of what it is for a mental state to be a thought? And what could it be about consciousness that enables us to understand the seemingly endless range of intentional contents our thoughts can have?

One who holds that consciousness can generate the needed conceptual resources might urge that consciousness can simply read off the nature of thoughts from those thoughts that are conscious, along with their intentional contents. And that would enable consciousness to come up with the required conceptual resources after all.

There is a picture of consciousness that might encourage holding that consciousness can operate in this way. On that picture, consciousness involves unmediated, intimate access to those of our mental states that are conscious, access sometimes described as a type of acquaintance. And it may be inviting to see this allegedly intimate access as enabling consciousness to read off the mental nature of state that is conscious.

This picture has led some even to hold that for conscious states, there simply is no distinction between how those states subjectively appear in consciousness and how they are in themselves (e.g., Nagel 1974: 444). The subjective appearances of consciousness exhaust the mental nature of the states that are conscious. In that way, consciousness itself would incorporate that mental nature, and thereby the conceptual resources needed to characterize the nature of all conscious states, thinking included.

But this picture of consciousness is difficult to sustain. For one thing, there is no reason to think that consciousness does bear such an intimate, unmediated relation to the states that are conscious. The subjective appearances of consciousness do not reveal any such mediation, but appearances never reveal how they are related to the things that appear. So all we really have to go on is that consciousness subjectively appears to be unmediated, not that it actually is. To determine whether there is any actual mediation would require taking into account factors other than just the subjective appearances themselves. And no other factors help sustain an actual absence of mediation, as against its subjective appearance.

The same holds for the claim that the appearances of consciousness exhaust the relevant mental reality. There is no reason to hold that apart from an insistence that we can appeal only to the subjective appearances to learn about the nature of mental states that are conscious. And that begs the question about whether the subjective appearances exhaust the relevant mentality. This picture of consciousness cannot be sustained (Rosenthal 2022). And without it, we have no basis to see consciousness as incorporating the conceptual resources needed to describe our thoughts, or reading them off from the conscious thoughts themselves. So we have thus far no credible explanation of how we do come to have those conceptual resources.

If one is in a conscious state, it subjectively appears to one that one is in that state. And that subjective appearance is typically reliable about what mental state one is in. But it is not the last word about that; we can investigate an individual's mental states in third-person ways, which are independent of the subjective appearances of consciousness. So it cannot be that the subjective appearances exhaust the relevant mentality. We'll come back to this point in section V.

For any mental state to be conscious, one must be in some way aware of being in that state. If a mental state occurs without one's being in any way aware of being in it, that state does not count as a conscious state (Rosenthal 2005). And first-person access to one's mental states depend on those states being conscious, and accordingly on one's being suitably aware of being in those states.

So for a state of thinking to be conscious, one must be in some way aware of having that thought. Since thinking always involves intentional content, the way one is aware of having a particular thought must always capture the relevant content. And since that intentional content is conceptual content, the awareness of a thought that enables it to be conscious must itself involve conceptual content.

But consciousness cannot on its own provide the required conceptual resources. Since we are subjectively aware of our conscious thoughts in respect of their intentional content, we must come to have those resources in some other way. And we must do so however in a way that makes those resources available to the way consciousness operates.

Sellars does not appeal to the inability of consciousness to deliver the conceptual resources needed to think about thinking in arguing for his view about the nature of thinking and the way we know about it. But because Sellars develops the concept of a thought from exclusively third-person considerations, it might readily regard the inability of consciousness to help as implicit in his argument. In any case, noting that consciousness cannot provide those conceptual resources arguably enhances and strengthens the arguments that Sellars does explicitly give.

3 | SELLARS' THEORY OF THINKING AND SPEECH

One might at this point consider the possibility that the needed conceptual resources are simply innate to the human mind, built somehow into our cognitive machinery. But without some independent support for such nativism, that is not a serious explanation, but rather just a colorful way of saying we have no explanation. And there are other options. In particular, Sellars' theory of thoughts provides an informative and compelling account of how we do actually learn about intentional states and come to have the conceptual resources to describe and to be aware of them.

Sellars argues that we understand the speaker's meaning of speech acts in roughly functionalist terms, that is, by the way speech acts interact with one another, with accompanying nonverbal behavior, and with salient perceptible aspects of the environment. Those considerations provide an account of what it is for a speech act to have some particular speaker's meaning.

And this functionalist account of speaker's meanings does not appeal in any way to the thoughts a speaker has and expresses by various speech acts. We plainly grasp what others mean by their speech acts without independently knowing the thoughts being expressed, since we come to know those thoughts in large part by first knowing the relevant speaker's meanings. Nor does such a functionalist account of speaker's meanings appeal in any way to first-person access to our own thoughts or to consciousness. It proceeds altogether in third-person terms.

Those functionalist considerations enable us to taxonomize speech acts in respect of their speaker's meanings, and thereby to understand what it is for speech acts to have speaker's meanings. This much is reasonably uncontroversial. We do understand the speaker's meaning of others' speech acts, and plainly nothing about first-person access or consciousness figures in how we do so.

Because speech acts express thoughts, it's natural to see speech acts as being caused by the thoughts they express. But taxonomizing and understanding speech acts in functionalist terms does not, by itself, point to any such thoughts. Indeed, being able to taxonomize speech acts in functionalist terms does nothing at all to explain how it is that those speech acts come to occur. So Sellars argues that we could understand speech without knowing anything about how they come to occur, and so without knowing anything about the thoughts that speech acts express.

It is worth noting, for now just in passing, that such a functionalist account does not square with the widely accepted account of speaker's meaning, due mainly to Paul Grice (1957 and 1968), which explains speaker's meaning by appeal to communicative intentions. This issue is addressed in detail in section VI.

It might seem tempting to deny that we can understand speech without knowing about the thoughts that speech acts express. Indeed, how ia first-person case could one fail to know that it is one's own thoughts that one's speech acts express? But one won't know that unless all one's thoughts are conscious. And not all our thoughts are conscious; when a solution to a nagging problem suddenly occurs to one out the blue, unconscious thinking must have led to one's conscious thought of the solution. Since not all our thoughts are conscious, it could have been that none of them are conscious. And that would in no way have prevented us from grasping speaker's meanings by way of functionalist considerations.

We want to explain how we come to understand the thoughts that we and others have, and to have the conceptual resources required to do so. Sellars urges us to consider mythical ancestors who have thoughts, readily express them in speech, understand speaker's meanings in functionalist terms, but are altogether unaware of their thoughts. None of their thoughts are conscious, and they have no conception of anything like a thought. These people taxonomize and understand speech acts, but have no explanation of how their speech acts occur.

Still, these people have a proto-scientific desire to explain why speech acts do occur. The natural, indeed the inevitable, move for these people is to posit, in a folk-theoretical way, states internal to speakers that cause their speech acts. And the only way to describe these posited internal states would be as the causes of speech acts with specific speaker's meanings.

To explain a speech act of asserting that it's raining, these people would posit an internal state that causes that speech act. And they would describe that internal state as an assertoric it's-raining internal state. They would describe the relevant speech act as asserting that it's raining; so they would describe the internal posited state as an assertoric state that it's raining, just as we describe thoughts.

These internal posits would be described as having properties corresponding to the speaker's meaning and illocutionary force of the speech acts that would express them. But they would not be posited as having any other properties corresponding to those of speech acts, such as acoustic or written properties. The only taxonomizing properties of the internal states would be those analogous to the speaker's meaning and illocutionary force of speech acts. In outline, this is the theory Sellars develops in "EPM" about how we learn about and come to have the conceptual resources to describe and think about the thoughts of others and of ourselves.

Some of the posited states tend to cause assertoric speech acts. But others tend to cause desiderative speech acts, such as saying, "Let it be that such-and-such." And such desiderative internal states might cause not only a speech act, but also nonverbal behavior that might bring about the relevant state of affairs. This explains why we can typically say what we are doing when we engage in intentional nonverbal behavior; a single internal state causes both the nonverbal behavior and a speech act that in effect describes the goal of that behavior.

The posited internal states are thoughts, desires, doubtings, and so forth. They have intentional contents and mental attitudes that correspond, respectively, to the speaker's meanings and illocutionary forces of speech acts that could express them. And they interact with behavior and perception in the ways characteristic of intentional states. Because the posited states cause speech acts, they are prior to speech acts in the order of being. But we initially come to know about them by folk-theoretical reasoning designed to explain the occurrence of speech acts, which we understand and taxonomize independently of knowing anything about the posited states. And this accordingly also explains how we come to have conceptual resources needed to describe and think about thoughts.

One might urge that instead of understanding speech acts in functionalist terms and then extrapolating to thoughts, we can simply apply that functionalist apparatus directly to thoughts, understanding their nature by appeal to connections with behavior, perceptual inputs, and other thoughts. But this would leave unexplained how we initially identify any states as thoughts, unless one tacitly assumed either that we do so by way of consciousness or that we initially posit thoughts as Sellars argues. In addition, taking thoughts to be states expressed by speech acts, as Sellars does, turns out to be pivotal for understanding how first-person access to thoughts operates (section IV), and indeed how it arises (section VII).

Extrapolating from speech acts to thoughts in the way Sellars urges may recall Jerry Fodor's well-known language-of-thought hypothesis (1975). And Sellars himself speaks of thoughts as a type of "inner speech" (1964: 656), and seems to have been the first to coin the useful term, 'Mentalese' for such an inner language (1964: 657). But the two views differ in important ways. Fodor maintains that thoughts have subsentential syntactic structures that correspond to sentences. But subsentential syntactic structures do not figure in the positing of thoughts as inner states with contents that correspond to speaker's meanings. Speaker's meaning comes in sentence-sized units; subsentential parts of speech acts have no speaker's meaning on their own. So insofar as the positing of internal states relies on speaker's meanings, subsentential syntactic structure is irrelevant to the states posited in that way. Thoughts so construed constitute inner speech in respect of content and mental attitude, not syntactic structure.

Those who assume that thoughts do have subsentential syntactic structure may see Sellars' positing as failing to do justice to the nature of thoughts. The speech acts that express thoughts use sentences that have subsentential syntactic structure, and we describe the content of thoughts using 'that' clauses and other sentence nominalizations, which also exhibit syntactic structure. And first-person access to our thoughts typically represents them by way of such sentence nominalizations.

But even though we understand thoughts by the speech acts they can express, there is no warrant for imputing subsentential syntactic structure to the thoughts themselves. For one thing, many thoughts can be expressed equally well by speech acts with different syntactic structures, and we can then fully describe those thoughts using sentence nominalizations with different syntax. In these cases it would be wholly arbitrary which structure to impute to the thought. So we cannot in these cases sustain the view that the thought itself exhibits some particular syntactic structure. And if thoughts don't exhibit syntactic structure in these cases, why think they ever do?

Fodor and Zenon Pylyshyn (1988: esp. p. 806), followed by many others, urge that thoughts exhibit what they call systematicity, an example of which is that if one can think that John loves Mary one can also think that Mary loves John. But there is no serious reason to think such regularities hold for the thoughts of nonlinguistic creatures.

These regularities do occur for the thoughts of linguistic creatures, but they result from the syntactic regularities of natural language. Linguistic creatures routinely express thoughts using speech acts. So the syntactic regularities of natural language will induce corresponding regularities in their thoughts. The regularities among thoughts arguably result from the syntactic structures of speech acts, not of the thoughts themselves.

Thoughts do likely exhibit syntactic properties corresponding to various sentential compounds, such as those formed with 'and', 'or', 'if', and the like, at least the thoughts of linguistic creatures (Carey 2023). And there may be considerations independent of the folk-theoretical positing Sellars describes that suggest subsentential syntax for thoughts themselves. Some have urged, for example, that we cannot explain inferential relations among thoughts unless they have syntactic structures, though I'll argue in section VI that this is not so. And Sellars himself holds that the thoughts of English and Japanese speakers likely differ "systematically … in a way which reflects the differences between these two languages" (1964: 663, fn. 14), though he does not specify that such differences are syntactic. Though it is widely held that thoughts themselves have subsentential syntax structure, it is unclear that there is any compelling reason to think so.

In any case, what mainly matters here is just that the folk-theoretical positing of Sellars' account does not, by itself, warrant ascribing subsentential syntactic structure to thoughts, though the account also does not preclude doing so. By contrast, Fodor's language-of-thought hypothesis does require that thoughts have subsentential syntactic structure. In section VI I'll describe another important difference between the two views.

Sellars argues that a full understanding of speech acts in respect of their speaker's meanings involves normativity (1954: §59ff.). Because "linguistic objects are subject to rules and principles," they are, as he memorably put it, "fraught with "ought" (1962: 44, §43.1; on Sellars and normativity see Ferguson & Koons, eds., 2023). So Sellars also held that a full understanding of thoughts involves normativity.

It is arguable that the rules and principles governing speech acts are simply abstractions from the dispositions that govern the occurrence of those speech acts. Speech tends to conform to rules, but not from any choice about whether to follow them. The rules are regularities that encapsulate hugely complex dispositional factors, which could in principle be wholly cashed out in descriptive terms. Indeed, many appeals to normative considerations can in this way be seen as a specific case of the descriptive.

These considerations aside, we need not for present purposes take normativity into consideration. As Sellars himself notes, "we abstract from [the normativity of speech acts] in considering them as objects in the natural order" (1962: 44, §43.1). And though one can debate Sellars' views about normativity, his account of the folk-theoretical positing of thoughts to explain the occurrence of speech acts arguably need not rely on those views. All that's needed is an account of speaker's meaning that makes no appeal to thoughts. And standard functionalist accounts typically satisfy that condition.

To stress the pivotal role of folk-theoretical positing, Sellars casts his account in terms of a mythological theoretical genius he calls Jones, who initially posits internal states to explain speech acts (1956: 314). And I'll often refer to that folk-theoretical reasoning simply as Jones's theory. But Sellars also urges that such folk-psychological theorizing did gradually lead early humans to develop the concept of a thought (1956: p. 328, §63). And it is reasonable also to see a primitive form of such explanatory reasoning as enabling human infants, together with linguistic input from adults, to come to have that concept.

4 | FIRST-PERSON ACCESS

Sellars' account of thoughts as folk-theoretical posits enables us to explain how we come to have the conceptual resources needed to describe and understand thoughts. People initially come to know about thoughts by folk-theoretical reasoning, somewhat along the lines of Daniel Dennett's idea of the intentional stance (1971 and 1987). Such an account is indeed natural for explaining our third-person access to the thoughts of others.

But we also have first-person access to our own thoughts, and that access does not seem subjectively to rely on Jones's folk-theoretical reasoning. So Sellars' account must also explain such first-person access. It must explain how

first-person access operates, and indeed how it is even possible given Sellars' proto-theoretical account of how we initially come to learn about thoughts. And that explanation must help us understand why many take first-person access to be our primary way of knowing about thoughts. Most importantly, Sellars' account must explain how the conceptual resources we get from Jones's theorizing for describing thoughts are available to the way consciousness makes us subjectively aware of our thoughts.

Jones's proto-scientific reasoning applies to one's own thoughts as well as to the thoughts of others. We initially learn about our own thoughts as explanatory posits. In effect, we apply the intentional stance to ourselves. But initially coming to know about something by theoretical reasoning does not preclude subsequently coming to know about it independently of such reasoning. Such cases are common. Genes were first posited as whatever carries genetic information, but subsequently became visible using microscopic examination.

In this way, the internal states that Jones's folk-theoretical posits are initially known only by way of suitable folk-theoretical reasoning, but we subsequently come also to have first-person access to our own internal states independently of any such folk-theorizing. As Sellars writes: "[I]t turns out—need it have?—that [Jones's contemporaries] can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe [their own] overt behavior." In this way our descriptions of thoughts, which "began … with a purely theoretical use, [have] gained a reporting role" (1956: 320).

Sellars' remarks about this process are extremely abbreviated, especially his cryptic and elusive "need it have?". Still, those remarks do point toward a credible account of how such first-person access would come about, and how it would operate. The reporting of a thought is the performing of a speech act. And on Jones's theory every speech act itself expresses a thought whose intentional content matches the relevant speaker's meaning. So a report of some target thought will express a second thought to the effect that one has the target thought.

The relevant reporting, moreover, occurs "without having to observe [one's own] overt behavior." So if one reports being in a particular target thought, the higher-order thought that one's report expresses will itself occur independently of any self-observation. When one has a thought about something as being present, moreover, one is thereby aware of that thing. So if thinks, independently of self-observation, that one has some target thought, one will be aware of oneself as having that target thought in a way that subjectively seems unmediated.

Sellars casts his account of Jones's proto-theorizing as a myth. So one might question whether he sees our concepts of thoughts as genuinely theoretical in nature. Indeed, Sellars recognizes this this question could arise because of our knowing our own thoughts by way of introspection (1956: 278–9). But it is plain that Sellars does regard thoughts and other mental states as theoretical posits, though folk-theoretical in nature. For one thing, Sellars urges that we should understand his myth as actually describing the development of humanity in coming to be able to think about thinking (1956: 328).

Moreover, Sellars explicitly holds that thoughts occur independently or our having first-person access to them, and seems even to countenance the possibility that we might never have come to have such first-person access. First-person access, he argues, occurs only because people can "be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe [their own] overt behavior" (1956: 320).

And introspective access to thoughts is in any case possible only when one has the concept of a thought, which is due to folk-theoretical reasoning. Because our having such subjectively direct access is in effect an add on to the thoughts themselves, Sellars casts the distinction between thoughts as such and thoughts as known by introspection as the "contrast between *theoretical entities* and *theoretical entities plus*" (Sellars & Castañeda, 1961–62: Sellars' letter of November 14, 1961, paragraph 9, Sellars' emphasis).

A thought that one is aware of in a way that subjectively seems unmediated is a conscious thought. So when Sellars envisages that Jones's contemporaries come to report "without having to observe [their own] overt behavior," he is in effect describing conscious thoughts. They are conscious in virtue of one's having a higher-order thought (HOT) about them (Rosenthal 2005). Sellars' account in effect points toward a HOT theory of what it is for a thought to be conscious.

HOTs represent the thoughts they self-ascribe in conceptual terms. So one cannot have HOTs unless the required conceptual resources are already be in place. Sellars' account takes care of that. Our positing of thoughts as

internal states that cause speech acts generates the conceptual resources needed to represent thoughts in respect of their intentional content and mental attitude. And it does so independently of any thoughts' occurring consciously, indeed, in advance of any thoughts occurring consciously. The initial folk-theoretical positing provides the conceptual resources required for thought to become conscious and to have first-person access to them.

The account also explains how those conceptual resources are available to consciousness. Jones's theory enables one to report that one has various thoughts using intentional-stance reasoning based on one's own behavior and independently of consciousness. And each self-report expresses a thought that one has a particular target thought. When one then comes to be able to self-ascribe thoughts independently of observing oneself, those self-ascribing HOTs use the same conceptual resources as the earlier self-ascriptions. They differ only in that the new self-ascribing thoughts seem subjectively to occur independent of any mediation. Their conceptual content is unchanged.

Consciousness reveals one's thoughts in respect of their intentional content, and also as a particular type of internal state. But consciousness does not explicitly say anything about any causal relations one's thoughts may have. The way Jones's theory posits thoughts helps explain these things. A HOT characterizes a target first-order thought by words that could figure in a speech act that expresses that first-order thought. That is sufficient to capture the content of the first-order thought. The HOT does not also need to describe how it is that the relevant speech act comes to have its speaker's meaning. So the HOT need not mention the functionalist machinery that is responsible for that speaker's meaning. So the HOT also needn't mention anything about the causal relations the first-order thought may have.

Sellars does not himself develop this HOT line of reasoning. Indeed, he is not at especially forthcoming about how first-person access operates, nor even why it does arise at all. And he says nothing that directly helps understand the force of his cryptic, parenthetical remark 'need it have?" in the quotation a few paragraphs earlier. And apart from whether it was inevitable that Jones's contemporaries came to report thought thoughts independently of self-observation, it's not obvious why simply becoming adept at applying Jones's theory to oneself in a third-person way would result in that ability, or even facilitate it.

I'll come back to those questions in section VII, where I develop an explanation of why thoughts do ever occur consciously, an explanation that fits comfortably with Sellars' views, and even helps explain what his cryptic "need it have?' likely means. For now, let's simply ask how well Sellars' account of how we initially come to know about thoughts actually squares with our commonsense understanding of these matters.

As noted in section III, we describe thoughts, both our own and those of others, using nominalizations of sentences that could figure in speech acts that express those thoughts. We say, "I think that ..., " where '...' are words that could express the thought we describe. This fits well with Sellars' account, on which thoughts are states whose contents correspond to those of speech acts that could express them. Our actual practice in describing thoughts reflects Jones's theory.

Jones's theorizing explains the intentionality of thinking, the tie between speech and thinking, and how we come to think about thoughts, our own and others'. And that explanation fits our commonsense views about thinking. These are powerful reasons to hold that Sellars' account is correct about how we initially come to know about thoughts.

Still, Sellars' claim that we initially come to know about thoughts as folk-theoretical posits strikes many as so wildly unintuitive as to be not worth taking seriously. Many see it as simply not credible that we initially learn about thoughts theoretically, and insist that we instead initially learn about them by first-person access, independently of any theorizing. On this view, shared by many, we learn first about our own conscious thoughts from such first-person access, and only then infer that others have thoughts as well.

Such a picture may seem inviting given the subjectively unmediated character of our first-person access to our own conscious thoughts. But as argued in section II, consciousness cannot deliver the conceptual resources needed to think about and describe thoughts. Since a picture on which we know about thoughts initially by way of first-person access allows for no explanation of how we come to have the needed conceptual resources, that

consciousness-based picture is untenable as an account of how we do initially learn about thoughts. It is uncontroversial that first-person access reveals our thoughts in respect of their intentional content. But with no account of how we come to have the conceptual resources needed to do that, we cannot even understand how first-person access itself could possibly operate.

In their published correspondence, Chisholm writes to Sellars that "[i]f the people of your myth were to give just a little bit of thought to the semantical statements they make, wouldn't they then see that these semantical statements entail statements about the thoughts of the people whose language is being discussed?" (1958: 537). This objection channels that idea that we primarily know about thoughts by way of first-person access.

But Sellars' account readily explains why first-person access would seem primary in this way. Jones's theory posits thoughts as states that speech acts express. So given Jones's theory, statements about speaker's meanings do entail statements about thoughts. And once Jones's theory has been internalized, as Sellars hypothesizes it now is, that entailment becomes second nature, as Chisholm urges. The internalizing of Jones's theory explains why first-person access now subjectively seems primary, though it seems so only relative to Jones's theory.

But if first-person access were truly primary, it's not obvious how we could come to understand that the conscious thoughts we learn about in that way are states of the same type as the thoughts we understand others to have. First-person access doesn't deliver information about any causal ties thoughts have. But we understand the thoughts of others in terms of functionalist causal connections. If we understood our own thoughts primarily by way of first-person access, how could we figure out that those thoughts are states of the same type as those of others?

It won't work simply to insist that for others to have thoughts is simply for them to be in states like those one knows about in one's own case by first-person access. Though we do assume that others do also have first-person access to their thoughts, we don't conceive of the thoughts of others in those terms. We understand what it is for others to have thoughts by way of the functionalist ties that those thoughts typically exhibit.

And it isn't credible that, having initially learned about thoughts from first-person access, we came to learn about their functionalist connections by observing ourselves. People simply don't engage in such detailed, elaborate self-observation. We come to know about the thoughts of others in functionalist terms. But on the primacy of first-person access, we would not also know about our own thoughts in functionalist terms. So there would be no reason to count our own thoughts as states of the same type as the thoughts of others. The primacy of first-person access cannot be sustained.

5 | THE MYTH OF THE GIVEN

We want to explain how we come to have the conceptual resources required to characterize thoughts. There are broadly speaking two possibilities. We could come to have those conceptual resources in a first-person way, from consciousness itself. Or we could come to have those resources in a third-person way. Sellars' account is arguably the most compelling version of a third-person approach, indeed, perhaps the only version that has been developed in a serious way.

On the picture of consciousness that figures in a first-person approach, consciousness has an intimate, unmediated relation consciousness has to those states that are conscious, so that consciousness reflects or even incorporates the nature of each conscious state. It was argued in section II that there is no support for this picture of consciousness apart from what the subjective appearances themselves seem to tell us. Absent independent support, that picture cannot be sustained.

Nonetheless, many find the pull of the subjective appearances compelling. Appearances never tell us about anything beyond the appearances themselves. But the subjective appearances of consciousness may seem to be appearances of last resort. When an object visually appears red, there is always the further subjective appearance in one is sensing red. But there is no layer of appearance after that. So perhaps the layer of subjective appearance operates differently from the others. Perhaps in this case there is after all no distinction to draw between the appearances

and the reality that appears to us. And then perhaps the subjective appearances would indeed deliver the conceptual resources needed to characterize that mental reality.

Ordinary appearances typically present themselves as being simply given. We have little if any control over how things appear to us in perception. Still, the way appearances do seem to be given in perception is always provisional, since whenever we perceive something we can appeal to factors independent of the appearances to investigate and so determine what it is that we actually perceive.

But if the subjective appearances of consciousness were the ultimate layer of appearance, those subjective appearances would be the last word about the mental states that appear to us in consciousness. There would be no independent way to investigate whether one is actually in the mental states that one subjectively appears to be in. So those appearances would be given to us in a way that is not provisional in the way that the appearances of conscious perceptions are. The subjective appearances of consciousness would then be given in an absolute way. The picture of consciousness on which consciousness could deliver the conceptual resources needed to characterize thoughts underwrites such an absolute given.

Beginning in "EPM" and in many subsequent writings, Sellars is dedicated to disarming and rejecting what he called the Myth of the Given. Indeed, it is that Myth of the Given that stands in the way of accepting his account of thoughts first presented in "EPM." And the picture of consciousness that enables an alternative to Sellars' account also generates an absolute given. All this suggests that this picture of consciousness figures in a central way in what Sellars calls the Myth of the Given.

Sellars initially talks about the Myth of the Given in "EPM" (1956: e.g., 267), but subsequently in many other places, characterizing it in a number of different ways, the connections among which are not always obvious. In a penetrating and compelling discussion, James O'Shea has argued that Sellars' Myth of the Given pertains at bottom to what is sometimes called the categorial given (O'Shea 2021). O'Shea draws in part on Sellars' discussion in the first Carus Lecture. There Sellars writes that "perhaps, the most basic form of what I have castigated as 'The Myth of the Given'" consists in holding that "[i]f a person is directly aware of an item which has categorial status C, then the person is aware of it *as* having categorial status C" (Sellars 1981: 11, §44).

Bracketing for a moment what Sellars has in mind by categorial status, Sellars' remarks here evoke echoes of the picture of consciousness discussed above. Consciousness on that picture is unmediated, and Sellars describes the categorial Myth of the Given as involving direct awareness. And just as consciousness on that picture takes on somehow the nature of mental states from the states themselves, so on the categorial Myth of the Given direct awareness of something is sufficient to be aware of it as having whatever categorial status it actually has.

These similarities are reinforced by Sellars' description of the categorial Given as "*the idea that the categorial structure of the world—if it has a categorial structure—imposes itself on the mind as a seal imposes an image on melted wax*" (Sellars 1981: 12, §45, Sellars' emphasis). As O'Shea stresses (O'Shea 2021: 10554; cf. O'Shea 2021: 10549 and 10553), the Myth of the Given is not, as some have held (e.g., BonJour 1985), a matter of epistemological foundationalism. It is rather a claim about how the mind works. When one is directly aware of something, its nature imposes itself on the mind.

On the categorial Myth of the Given, if one is being directly aware of something as having some categorial status, one cannot come to see it as having any different categorial status (Sellars 1981: 12, §45; O'Shea 2021: 10549, 10553). So the categorial Given, like the picture of consciousness, precludes any theorizing that could compete with or displace the way direct awareness reveals things. And if the relevant categorial structure of something did "*impos[e] itself on the mind as a seal imposes an image on melted wax*," that by itself would generate the conceptual resources needed to understand that thing.

Sellars offers examples of such categorial status in the Carus lectures, but doesn't actually explain what it is. O'Shea helpfully offers a convincing account (e.g., O'Shea 2021: 10552). But because the categorial Given is cast specifically in respect of the categorial nature of what one is directly aware of, it is narrower than the picture of consciousness, which is not confined to the categorial nature of a thing, however one understands that. And the picture of consciousness concerns only the way we are aware of our conscious states, whereas Sellars' categorial Given pertains to direct awareness of nonmental reality as well.

So the picture of consciousness does not, in at least those two ways, match Sellars' categorial Given. Still, the two may well operate in the same way in connection with the categorial nature of thoughts, since that categorial nature might simply consist in what it is for a mental state to be a thought. And the picture of consciousness is arguably the main source of opposition to Sellars' account of thoughts, just as he sees the Myth of the Given itself. And that special case aside, the categorial Given and the picture of consciousness have in common the idea that the subjective appearances of consciousness accurately reveal the nature of what appears in awareness.

6 | TWO OBJECTIONS

The picture of consciousness is doubtless the major source of discomfort with Sellars' argument that we initially come to know about thoughts by way of Jones's proto-scientific reasoning. But there are two more specific objections to Sellars' argument that are important to consider.

On Sellars' account, the speaker's meaning of a speech act is a matter of the intentional content of the thought that speech act expresses. As noted earlier, that conflicts with a highly influential and widely accepted account of speaker's meaning due to Grice (1957 and 1968). Simplifying Grice's complex account slightly, the speaker's meaning of a speech act results from a speaker's intention that a hearer should believe or do something (Grice 1957) or, in a later, slightly revised version, a speaker's intention that a hearer should at least think that the speaker believes something or that the hearer should at least intend to do something (Grice 1968). The speaker's meaning of a speech act derives, on this account, not from its expressing a thought with the same content, but rather from a speaker's intention that the speech act should have a particular effect on a hearer or hearers.

On both accounts, the speaker's meaning of a speech act derives from an intentional mental state. But the difference about what kind of intentional state is responsible for speaker's meaning is pivotal. On Sellars' account, thoughts are posits whose content matches the speaker's meaning of speech acts that could express the thoughts. The content of each thought matches the speaker's meaning of a speech act that could express that thought. So the account locks in the relation that holds between speaker's meanings and intentional states. Speaker's meaning cannot then be determined by a speaker's communicative intentions.

So Grice's view of speaker's meaning if correct would be an objection to Sellars' account. But there is compelling reason to see Sellars' account as correct, and not that of Grice. The intentions Grice describes are plainly crucial to communication. But that by itself doesn't show that speaker's meaning is determined communicative intentions, rather than by the thought or desire a speech act expresses. Communication requires one to express a thought or desire, and speech acts can have speaker's meaning independently of any communicative context. And a speech act will have speaker's meaning even if Grice's communicative intentions, which are highly intricate, go wrong. On the face of it, the intentional state a speech act expresses is basic to speaker's meaning, and communicative intentions simply flesh out the communicative context. All that is in addition to the various types of counterexample that have been leveled against a Gricean account (e.g., Pavese & Radulescu 2023).

And an important test case for the two accounts occurs in connection with insincere speech. When one speaks insincerely, one intends one's utterance to get somebody to believe something that one does not believe. On a communicative-intention view, that intention straightforwardly provides one's insincere utterance with speaker's meaning. The speaker's meaning is whatever one intended to get a hearer to believe by making that utterance (Grice 1957), or at least whatever one intended the hearer to think that one believed (Grice 1968). Indeed, a communicative-intention account is altogether indifferent to what the speaker actually believes. So such an account cannot on their own differentiate sincere from insincere utterances, and must give the same account of speaker's meaning for both.

That is a mistake. If I lie to you, I don't believe the semantic meaning of my utterance. I do intend you to take me to believe that semantic meaning, but unlike a case of sincere speech, I do not actually do so. Indeed, I have no belief or thought whose content corresponds to the semantic meaning of the utterance. So it cannot be that what I, as a

speaker mean in uttering my lie is the same as it would be if my utterance were truthful. Since what I mean is different from the sincere case, there must be a corresponding difference in speaker's meaning.

Seeing speaker's meaning as a matter of the content of the thought a speech act expresses gets all this right. If I speak insincerely, I am not in any intentional state whose content matches the semantic meaning of the utterance. So I do not mean anything at all in producing the utterance, despite whatever communicative intentions I may have. And since I literally do not mean anything by my insincere utterance, that utterance has no speaker's meaning. We must distinguish what one means by a speech act from what one wants to achieve by performing that speech act.

The expressing account usefully differentiates insincere from sincere speech. And the expressing account also points to an informative explanation of what does occur in insincere speech. In speaking insincerely, one does not perform a speech act with illocutionary force at all; one simply acts as though one is doing so (Rosenthal 2005: 87–93 and 100–102; Frege 1956: 294; cp. Austin, 1970/1979: 248–251). One pretends to express a thought with the relevant content; one pretends, that is, to be somebody exactly like oneself except for actually having that thought. Insincere speech is in that way like play acting. It is a case of what Sellars called parroting speech (1964). The expressing account does justice to our commonsense understanding of insincere speech, whereas a communicative-intention view cannot say anything at all about it.

In these ways, an expressing account is arguably superior to one based on communicative intentions. But the two accounts also differ in a way that has implications for how we come to have the conceptual resources needed to describe thoughts.

On Sellars' expressing account, we understand speaker's meanings of independently of the thoughts the speech acts express, so that we can then infer from those speech acts to the nature of the relevant thoughts. The expressing account thereby explains how we come to have the conceptual resources needed to characterize thoughts.

On a communicative-intention account, by contrast, we understand speaker's meaning only by appeal to communicative intentions. And that leaves it unclear how we could come to have those conceptual resources. Since we can't derive those resources by inferring from speaker's meanings, advocates of a communicative-intention account might tacitly assume that consciousness by itself delivers those resources. If so, advocates of that account would face the difficulties raised in sections II and IV for that picture of consciousness.

Sellars' account avoids those difficulties by holding that we understand speaker's meaning independently of any appeal to thoughts. But that explanation of speaker's meaning might prompt another objection. Sellars argues that we understand speaker's meaning in roughly functionalist terms, appealing to moves in language games (Sellars 1954). So when we extrapolate from speech acts to the thoughts those speech acts express, we will also understand intentional content in roughly functionalist terms.

Functionalism describes intentional content in part by the relations each thought has with every other thought in an individual's repertoire of thinking. We can understand these relations in terms of dispositions for one thought to cause various others; the relevant relations among thoughts are inferential dispositions. So the resulting account of intentional content is a type of semantic holism.

Such functionalist holism also appeals to dispositional ties intentional states have with perceptual stimuli and with nonverbal behavior, ties that reflect the language-entry and language-departure moves in Sellars' account of language games (Sellars 1969). A striking advantage of such a holist account of human thought is that it may help explain some of the shortcomings of large-language models, whose dispositional connections are highly limited and local.

As noted in section III, it is natural to see Sellars' extrapolation from speaker's meanings to thoughts as some type of language-of-thought hypothesis. But because that extrapolation relies on sentence-sized content properties, it does not by itself underwrite imputing subsentential syntactic structure to the contents of thoughts. By contrast, Fodor's (1975) classic language-of-thought hypothesis holds that thoughts do have such syntactic structures. But as argued in section IV, there is little reason to hold that thoughts themselves exhibit such structures, as against the speech acts that express them.

The holism of Sellars' account of the content of thoughts constitutes a second, related way in which it differs from Fodor's view. Fodor assigns subsentential syntactic structures to thoughts, and claims that the sentence-sized

contents of thoughts are constructed syntactically from term-sized concepts. Sellars, by contrast, holds that each sentence-sized intentional content is determined by functionalist ties to other thoughts and to perceptual inputs and behavior.

It is sometimes argued that to explain the inferential ties that hold among thoughts we must appeal to their syntactic components and structures. But content holism lets us explain those inferential ties by appeal to dispositions that take one from one thought to another. And we can also explain many inferential connections by appeal to the syntactic structures of the speech acts that we are disposed to use in expressing our thoughts. These explanations do not require imputing syntactic structures to the thoughts themselves.

Content holism explains intentional content and inferential dispositions together, as a package. One explains sentence-sized contents by robust, regular dispositions to pass from one thought to another when those dispositions are independent of idiosyncratic, passing associations. These inferential dispositions will not carve off semantic meaning from well-entrenched nonsemantic information. But since a serious line between analytic and synthetic is unlikely (Quine 1951), it is natural to expect that the contents of thoughts will often reflect such nonsemantic information.

It is also sometimes argued that we can explain the seemingly unlimited range of thoughts we can have only if the thoughts themselves are built up syntactically from subsentential components. This is also wrong. Language is compositional. So when one uses novel linguistic constructions to express one's thoughts, one will often come to have thoughts that are novel in respect of their intentional content. And the process of thinking can result in changes in the inferential dispositions that hold among one's thoughts, thereby generating thoughts with novel contents. Again, we need not impute syntactic structures to the thoughts themselves.

All that to one side, Fodor and Ernest Lepore (1992) have argued that content holism faces an objection that is dispositive. No two individuals have exactly the same repertoire of thoughts. So Fodor and Lepore note that if content is global inferential potential, none of the thoughts of one individual will have exactly the same content as any thoughts of somebody else. And since everybody's repertoire of thoughts changes over time, the content of any individual's thoughts will change accordingly.

This objection rests, however, on a wildly unrealistic absolutism about intentional content. If global inferential potential determines content, each token thought will indeed change its content whenever anything changes among one's other thoughts. But that change in content will be indiscernible, both subjectively and in respect of any effects on behavior, communicative success, or anything else that matters. And though the thoughts of different people that are expressible by the same forms of words will also always differ, the differences will again be indiscernible in those ways. These differences contrast sharply with differences that do affect communication, and often do get detected.

One might indeed question why we would expect two individuals ever to have thoughts with precisely the same content, or why one would expect one individual's thoughts to be precisely constant in their content over time. We rarely if ever expect things to have precisely the same properties, whether of color or size or shape, though we still speak of sameness when there is no relevant discernible difference. Why think otherwise for the intentional content of thoughts?

The answer likely relies on the picture that we grasp and individuate contents by the way they are present to consciousness. Consciousness would not reveal these indiscernibly minute differences in content. So that picture would encourage the absolutist individuation of content presupposed by Fodor and Lepore's objection. If consciousness reveals everything about intentional contents, the indiscernible differences that result from by content holism would seem objectionable.

But we do not individuate content by how it is present to consciousness, and we could not do so, since we must in any case calibrate our own contents with those expressed by the speech acts of others. Just as the Gricean account of speaker's meaning forces us to see consciousness as the source of our conceptual resources for characterizing thoughts, so Fodor and Lepore's objection against content holism likely rests on the picture of consciousness rejected in sections III and IV. That picture figures in some way in the most prominent reasons to reject a theory like Sellars'.

7 | WHY ARE THOUGHTS EVER CONSCIOUS?

Sellars urges that Jones's contemporaries would come to report their thoughts using the conceptual resources of Jones's theory but without relying on self-observation. This would constitute first-person access to our thoughts. Thoughts reported in that way would be conscious. And because such reports would express thoughts with content that matches the speaker's meanings of those reports, this points to a HOT theory of what it is for thoughts to be conscious.

But as noted in section IV, it is not obvious why Jones's contemporaries would actually come to report their thoughts in this way. Sellars writes: "[I]t turns out—need it have?—that [Jones's contemporaries] can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe [their own] overt behavior" (1956: 320). Is training really required? Is that outcome inevitable, as Sellars' elusive parenthetical question, "need it have?", might suggest? If so, why?

One possibility is that extensive habituation in making observation-based self-ascriptions would somehow lead to self-ascriptions that are independent of observation. But it's not obvious that habituation by itself would have that result, as against simply facilitating observation-based self-ascriptions. For Sellars' account to accommodate first-person access, we must explain why HOTs that are independent of observation would arise.

And such an explanation is available, relying once again on Sellars' account of speech acts as expressing thoughts with the same content. Consider language users with the ability to describe thoughts as folk-theoretical posits along the lines of Jones's theory. They can self-ascribe thoughts relying on third-person considerations, saying, for example, "I have a thought that it's raining." And they can also simply say, "It's raining."

Those two speech acts have distinct truth conditions, the first pertaining to the speaker's thought and the second to the weather. But the conditions for appropriately performing those two speech acts are roughly the same. Whenever it's appropriate to say that one thinks it's raining it's equally appropriate simply to say that it's raining, and conversely. So given the ability to describe one's thoughts at all, the two speech acts are virtually interchangeable in respect of appropriateness conditions.

The near equivalence of conditions for appropriately performing these two types of speech act is enhanced by a similar near equivalence that is altogether independent of speaking about one's thoughts. As Sellars notes (e.g., 1968: 71–73,155–156), being disposed to say something goes hand in hand with being disposed to say one is inclined to say that thing. (I am grateful to an anonymous referee for stressing this point.) This near equivalence of dispositions occurs even without one's having any concept of a thought. Once one does gain that concept, the near equivalence of saying something and saying that one thinks that thing will come to be automatic and second nature.

Habituation plays a significant role in that process. On Jones's theory, if one says something, one has a thought with the same content. As Jones's contemporaries habituated to apply the theory, the connection between saying something and having a corresponding thought will become second nature for them. So it will become second nature that whenever one says that it's raining, one could as readily have said instead that one thinks that it's raining. And if one could as readily have said that one thinks it's raining, one must have had the thought that one thinks it's raining.

That would be a HOT that one thinks it's raining. And it arises only as a result of its being second nature that saying that it's raining and saying that one thinks it's raining have equivalent appropriateness conditions. So that HOT arises independently of observation. So verbally expressing one's thought will give rise to a HOT that one has that first-order thought. These considerations explain why it is that verbally expressed thoughts are always conscious (Rosenthal 2005: ch. 10).

And the same mechanisms will be operative even when one's thoughts are not verbally expressed. A HOT arises whenever one verbally expresses some first-order thought. But many thoughts are disposed to be verbally expressed, even though that verbal expression does not occur. And when a thought is disposed to be verbally expressed, that will induce a disposition to say that one thinks that thing. So HOTs will also arise for many unexpressed thoughts. The more strongly disposed one is to express a thought verbally, the more likely that thought is to be conscious.

The idea that consciousness has some strong connection to speech has long history, though the nature of such a connection has long remained unexplained. The foregoing explanation of why verbally expressed thoughts are always conscious is therefore notable for providing an informative explanation of that connection.

Sellars' parenthetical, "need it have?', rhetorically raises the question of whether first-person access to our thoughts is inevitable on his account, but leaving it unanswered. The foregoing shows that it is inevitable if the appropriateness conditions of expressing thoughts that self-ascribe thoughts come to be second nature for speakers. And habituation in self-ascribing thoughts using Jones's theory would almost certainly have that result.

But apart from those developments, speakers will have thoughts but lack first-person access to them. HOTs about one's own thoughts would be unlikely to arise in any other way. A HOT that one has a particular first-order thought requires conceptual resources needed to characterize that first-order thought. And as argued earlier, it is unclear how one could come to have those resources without something like Jones's theory. So the thoughts of creatures such as infants and nonhuman animals, which lack the linguistic resources needed to describe thoughts, will arguably not be conscious.

This result might seem unexpected. Still, it conforms reasonably well to our pretheoretic ideas about infants and nonhuman animals. We do have a strong pretheoretic sense that their mental states are often conscious. But that sense arguably pertains only to qualitative mental states, such as bodily and perceptual sensations, and not in any clear way to thoughts. I'll come back to qualitative states in a moment.

Sellars was reluctant to ascribe full-fledged thoughts to nonlinguistic animals and prelinguistic infants. But his account of thoughts arguably applies independently of language. We understand highly purposive nonlinguistic behavior as involving states whose content is determined by functional role, and that applies to some behavior of infants (e.g., Bohus et al. 2023) and nonlinguistic animals (e.g., Gruber et al. 2019). Though the contents of states that give rise to nonlinguistic behavior will be vastly less fine-grained and sophisticated than that of speech acts, we can still in the spirit of Jones's theory extrapolate from such behavior to corresponding thoughts.

In an important discussion of this issue, Willem deVries shows that Sellars did in later work become more accommodating about nonlinguistic creatures having thoughts (deVries online). But that aside, the foregoing explanation of how HOTs about thoughts arise suggests an explanation of why Sellars might have had some hesitation. On that explanation, whatever thoughts nonlinguistic creatures may have would never be conscious. So perhaps Sellars' concerns about whether nonlinguistic creatures have full-fledged thoughts was due to a sense that the states of such creatures that function much as thoughts do in adult humans would nonetheless never occur consciously. And as noted, this likely reflects our commonsense view about whatever thoughts nonlinguistic creatures may have.

Human thoughts often occur unconsciously, and we still regard them as genuine thoughts. But should we regard states that never occur consciously as full-fledged thoughts? Indeed, should we regard a creature as having thoughts at all if that creature has no concept or conception of a thought? If a type of state has the functionalist ties to behavior, perceptual input, and other such states that adult human thoughts have, it is compelling to see them as thoughts. Issues about having a concept of a thought and first-person access are arguably irrelevant to what thoughts are.

Though we have no pretheoretic convictions about whether the thoughts of nonlinguistic creatures are conscious, our pretheoretic sense is that the qualitative states of nonlinguistic creatures, such as perceptual and bodily sensations, are often conscious. And the argument from appropriateness conditions for speech acts, which explained how thoughts come to conscious, is plainly irrelevant to qualitative states.

But we can independently explain how prelinguistic infants and nonlinguistic animals come to have HOTs about many qualitative states. To interact successfully with its environment a creature must calibrate sensory inputs that occur in one modality with those that occur in others. It is crucial to know which visual inputs, for example, result from the same object or event as various audiology inputs. The inputs on their own offer no help with this. A creature must come to register which inputs from each modality go with which inputs from others. That includes registering correlations of bodily sensations such as pains with perceptual inputs that pertain to their causes.

Such registering requires having thoughts about which inputs co-occur. Without such thoughts, a creature's perceptual abilities would remain very primitive indeed. But such thoughts are about the various sensory impacts that

the environment makes on the creature's sensory systems. So these thoughts are HOTs that one is in the relevant qualitative states. The cross-modal calibration required for perceptual interaction with the environment by itself gives rise to HOTs about qualitative states, making those states conscious at a very early stage (Rosenthal 2024: §6).

Sellars' myth of Jones, combined with a HOT theory of consciousness, provides a highly informative theoretical account of what it is for a state to be intentional, how we come to have the conceptual resources needed to characterize our intentional states, and how those states come to be conscious. And these theoretical accounts do all this independently of unsupported and questionable assumptions about consciousness.

ORCID

David Rosenthal https://orcid.org/0000-0002-9323-0424

REFERENCES

Austin, J. L. (1970/1979), "Performative Utterances," in J. L. Austin, Philosophical Papers, J. O. Urmson and G. J. Warnock (eds.). Oxford: Oxford University Press, 2nd and 3rd edns., pp. 233–253 (1st edition 1961, pp. 220–239).

BonJour, Laurence (1985), *The Structure of Empirical Knowledge*. Cambridge, Massachusetts: Harvard University Press.

Bohus, Kinga Anna, Nicolò Cesana-Arlotti, Ana Martín-Salguero, and Luca Lorenzo Bonatti (2023), "The Scope and Role of Deduction in Infant Cognition, *Current Biology*, *33*, 18 (September 25, 2023): 4014–4020.

Carey, Susan (2023), "Do Nonlinguistic Creatures Deploy Mental Symbols for Logical Connectives in Reasoning?", *Behavioral and Brain Sciences*, *46*, e261: 28–31.

Chisholm, Roderick M., and Wilfrid Sellars (1958), "Intentionality and the Mental," in Herbert Feigl, Micnael Scriven, and Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, vol. 2. Minneapolis: University of Minnesota Press, 507–539.

Dennett, Daniel C. (1971), "Intentional Systems," *The Journal of Philosophy*, *68*, 4 (February 25): 87–106.

Dennett, Daniel C. (1987), *The Intentional Stance*. Cambridge, Massachusetts: MIT Press/Bradford Books.

deVries, Willem (online), "Sellars, Animals, and Thought," on the "Problems from Wilfrid Sellars" website. URL: http://www.ditext.com/devries/sellanim.html

Ferguson, Kyle, and Jeremy Randel Koons (2023) (eds.), *The Metaphysics of Practice: Writings on Action, Community, and Obligation*. Oxford: Oxford University Press.

Fodor, Jerry A. (1975), *The Language of Thought*. New York: Thomas Y. Crowell.

Fodor, Jerry, and Ernest Lepore (1992), *Holism: A Shopper's Guide*. Oxford: Blackwell Publishers.

Fodor, Jerry A., and Zenon W, Pylyshyn (1988), "Connectionism and Cognitive Architecture: A Critical Analysis," *Cognition*, *28*, 1-2 (March): 3–71.

Frege, Gottlob (1956), "The Thought: A Logical Inquiry," tr. P. T. Geach, *Mind*, *65*, 259 (July): 289-311.

Grice Paul (1957), "Meaning," *The Philosophical Review*, *66*, 3 (July): 377–88.

Grice, Paul (1968), 'Utterer's Meaning, Sentence Meaning, and Word-Meaning', *Foundations of Language*, *4*, 3 (August): 225–242.

Gruber, Romana, Martina Schiestl, Markus Boeckle, Anna Frohnwieser, Rachael Miller, Russell D. Gray, Nicola S. Clayton, and Alex H. Taylor (2019), "New Caledonian Crows Use Mental Representations to Solve Metatool Problems," *Current Biology*, *19*, 4 (September 19): 686–692.

Nagel, Thomas (1974), "What Is It Like to Be a Bat?", *The Philosophical Review*, *83*, 4 (October): 435–50.

O'Shea, James R. (2021), "What Is the Myth of the Given?", *Synthese*, *199*, 3-4 (December): 10543–10567.

Pavese, Carlotta, and Alexandru Radulescu (2023), "Meaning without Gricean Intentions," *Analysis*, *83*, 4 (October): 708–716.

Rosenthal, David (2005), *Consciousness and Mind*. Oxford: Clarendon Press.

Rosenthal, David (2010), "How to Think about Mental Qualities," *Philosophical Issues*, *20* (October): 368–393.

Rosenthal, David (2016), "Quality Spaces, Relocation, and Grain," in James R. O'Shea (ed.), *Sellars and His Legacy*. Oxford: Oxford University Press, pp. pp. 149–185.

Rosenthal, David (2022), "Mental Appearance and Mental Reality," in Josh Weisberg (ed.), *Qualitative Consciousness: Themes from the Philosophy of David Rosenthal*. Cambridge: Cambridge University Press, pp. 243–271.

Rosenthal, David (2024), "Methodological Considerations for the Study of Mental Qualities," in Juraj Hvorecký, Tomáš Marvan, and Michal Polák (eds.), *Conscious and Unconscious Mentality: Examining Their Nature, Similarities and Differences*. Oxford and New York: Routledge, 2024, pp. 91–129.

Quine, W. V. (1951), 'Two Dogmas of Empiricism," *The Philosophical Review*, *60*, 1 (January): 20–43.

Searle, John R. (1983), *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Sellars, Wilfrid (1954), "Some Reflections on Language Games," *Philosophy of Science*, *21*, 3 (July): 204–228; reprinted in Sellars (1963), pp. 321–358.

Sellars, Wilfrid (1956), "Empiricism and the Philosophy of Mind," in Herbert Feigl and Michael Scriven (eds.), *Minnesota Studies in the Philosophy of Science*, vol. I. Minneapolis: University of Minnesota Press, pp. 253–329; reprinted in Sellars (1963), pp. 127–196.

Sellars, Wilfrid (1962), "Truth and 'Correspondence'," *The Journal of Philosophy*, *59*, 2 (January 18): 29–56; reprinted in Sellars (1963), pp. 197–224.

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

Sellars, Wilfrid (1964), "Notes on Intentionality," *The Journal of Philosophy*, *61*, 21 (November 12): 655–665; reprinted in Sellars (2011), pp. 117–128.

Sellars, Wilfrid (1965), "Scientific Realism or Irenic Instrumentalism: A Critique of Nagel and Feyerabend on Theoretical Explanation," in Robert Cohen and Marx Wartofsky (eds.), *Boston Studies in the Philosophy of Science*, vol. II, New York: Humanities Press, pp. 171–183; reprinted in Sellars (2011), pp. 141–166.

Sellars, Wilfrid (1968), *Science and Metaphysics*. London: Routledge & Kegan Pau.

Sellars, Wilfrid (1969), "Language as Thought and as Communication," *Philosophy and Phenomenological Research*, *29*, 4 (June): 506–527.

Sellars, Wilfrid (1981), "The Lever of Archimedes," *The Monist*, *64*, 1 (January): 3–36.

Sellars, Wilfrid (2011), *Philosophical Perspectives: Metaphysics and Epistemology*. Atascadero, CA: Ridgeview.

Sellars, Wilfrid, and Hector Castañeda (1961–62), Unpublished Correspondence, March 6, 1961-August 8, 1962, online, Andrew Chrucky (ed.). https://www.ditext.com/sellars/corr.html

How to cite this article: Rosenthal, D. (2025). Thought, consciousness, and the given. *European Journal of Philosophy*, e13039. https://doi.org/10.1111/ejop.13039