ON A WITTGENSTEINIAN OBJECTION TO KRIPKE'S DUALISM ARGUMENT

Richard Double

January 23, 1981
Abstract

In 'Kripke's Argument against the Identity Theory' Michael Levin argues that the private language argument can be used to undermine Saul Kripke's Cartesian claim to be able to imagine mental states and brain states existing apart, and, thus, refute his argument for dualism. In this paper it is argued that Levin's use of the private language argument relies implicitly upon the descriptivist theory of mental language, to which Kripke has provided a plausible alternative, viz., the causal theory of reference. Thus, using the private language argument in the way Levin suggests begs the question against the Cartesian line of argument.
On a Wittgensteinian Objection to Kripke's Dualism Argument

In the Meditations Descartes reasoned that since he certainly existed although he may not have a body, he was not identical to his body. Contemporary materialists such as Place, Smart, and Armstrong replied to Descartes by introducing the notion of logically contingent identity, insisting that identities need obtain only in fact, but might be otherwise. More recently Saul Kripke argued that logically contingent identity is at base incoherent and resurrected the Cartesian argument for dualism. More recently still, Michael Levin has challenged Kripke's Cartesian reasoning with a version of the private language argument in 'Kripke's Argument against the Identity Thesis.'

Letting 'pain' designate a sensation type and 'C-fiber stimulation' some typical brain event type which materialists wish to identify with pain, we schematize Kripke's argument against the identity of sensation and brain types this way:

(1) Identity statements using two rigid designators are necessary truths, if true.
(2) 'Pain' is a rigid designator.
(3) 'C-fiber stimulation' is a rigid designator.
(4) 'Pain = C-fiber stimulation' is not a necessary truth (because one can imagine pain existing without the stimulation of C-fibers.)

---


3 According to Kripke, the distinction between rigid and nonrigid designators is one between terms that refer to or have as extensions exactly the same entities in all 'possible worlds' (where they exist) and terms whose reference or extension differs in possible worlds. 'Benjamin Franklin' is a rigid designator picking out the same individual in all worlds where he exists, while 'inventor of bifocals' is a nonrigid designator picking out Benjamin Franklin in this world, but other individuals in possible worlds where someone else invented bifocals. A correlative distinction is that between rigidly fixing the reference of a term by an essential property of the referent and...
(5) Therefore, 'Pain = C-fiber stimulation' is not true.⁴

Although Levin casts Kripke's argument in a different format, the thrust of his criticism is to deny that Kripke's Cartesian thought-experiment justifies the claim that pain is not necessarily C-fiber stimulation. For Kripke, as for Descartes, one can imagine pain, itself, existing without the presence of brain states: 'Pain . . . is not picked out by one of its accidental properties; rather it is picked out by the property of being pain itself, by its immediate phenomenological quality.'⁵ Levin believes that regardless of the phenomenal character of pain, certain constraints that speakers face in understanding ('fixing') the reference of sensation terms prevent one from focussing one's thought on pain itself in the way that Kripke believes he has. According to Levin, since speakers can fix the reference of sensation terms only nonrigidly, Kripke's thought-experiment has no more force than one where we claimed that heat is not necessarily molecular motion because we can see that it is logically possible that the cause of sensations of warmth exist apart from molecular motion. Levin argues that the reference of 'pain' can only be fixed by nonrigid topic-neutral descriptions that characterize pain inessential-ly, e.g., 'what goes on when I am stuck with a pin.' This argument is based on two premises: (a) All descriptions of pain for a speaker who is fixing the reference of 'pain' originally are equivalent in meaning to nonrigid topic-neutral descriptions, and (b) The reference of 'pain' must be fixed originally by descriptions.⁶

Levin's argument for premise (a) is this:

It is axiomatic that describing x is a matter of ascribing to x predicates that x satisfies. It is a matter, tautologically enough, of knowledge by description and not acquaintance. Thus, no matter how intimate A's awareness of x is, A cannot describe x unless A is able to say something true about x.

. . . A's immediate acquaintance with his pain in its essence does not guarantee that A knows any descriptions of nonrigidly fixing the reference by an accidental property. For instance, I can nonrigidly fix the reference (give myself to know to what I am referring) of 'gold' by gold's accidental property of being the metal used in wedding rings; I can also rigidly fix the reference of 'gold' by its essential atomic structure. See 'Naming and Necessity,' 269-284.

⁴See Kripke, 'Naming and Necessity,' 337-42.

⁵Kripke, 'Naming and Necessity,' 340.

⁶Levin, 158.
his pain other than those (others) know. Granted, A is immediately conscious of his pain; he has it; he intuits its essence. But what does A know about his pain? A knows that it happened when he put his hand in the flame. It happened at 3:00, and so on. Moreover, it seems to me, if this is the first time A has experienced such a thing, these sorts of properties are all he knows about pain, at the first encounter. So, if A fixes the reference of 'A's pain' to this new phenomenon even for himself by description, he has to do it by the same . . . description that others use.7

Levin anticipated that Kripke might believe that premise (a) is irrelevant to his argument, since Kripke may hold that the reference of 'pain' is fixed without the use of any designator. According to this view, which Levin dubs 'the direct picture,' persons notice pain in their own cases and then refer to it as 'pain' (if they are English speakers) without relying on a description.8

At this point, according to Levin, the private language argument can be used to discredit the direct picture, and thereby sustain premise (b):

. . . Wittgenstein's point (on this interpretation of PLA) is that the direct picture of how A fixes the reference of 'pain'--by noticing pain and calling it 'pain'--is incoherent because it becomes the account of a situation where A refers to pain by making essential use of the description "what I initially referred to as 'pain'." So far, the point seems well taken. For consider: on every occasion after the initial encounter with pain, when the subject asks himself what pain is in order to denote a new occurrence as 'pain,' he must, on the direct picture, think of "what I denoted by 'pain' at the first encounter." He must ask himself whether what he is presently experiencing is similar to what he referred to as "pain." . . . But now the referentiality of 'pain,' how A refers to his pain, has been presupposed, rather than explained, on this view. As Kripke puts in another context, "The question of reference is thrown back on the question of reference."9

7Levin, 159-60.
8Levin, 161.
9Levin, 164.
There is some question whether Levin intends his criticism of the direct picture to hold simpliciter or merely against a Kripkean appeal to the direct picture. I shall defend both a general and Kripkean advocacy of the direct picture, taking the latter first.

Levin believes that Kripke cannot appeal to the direct picture without violating his own principles on referring. Levin believes that a speaker attempting to re-refer to pain according to the direct picture must ask himself the semantic question "Is this what I referred to as 'pain'?" thereby running afoul of Kripke's principle:

For any successful theory of how a particular word . . . refers, the properties used, according to that theory, to determine the reference of the word . . . must not make uneliminable use of that word's reference.10

My reply is that if the direct picture per se, is defensible, then Kripke's advocacy of the direct picture would not violate his scruples on the non-circularity of reference. Pace Levin, according to the advocate of the direct picture it is not necessary to ask oneself "Is this what I referred to as 'pain'?" because it is not necessary to ask oneself anything at all. To think that one would have to ask oneself this type of question in order to re-refer to pain is already to assume that speakers require descriptions, in this case a semantic description. But if a speaker does not require any descriptions of pain in order to refer to it, then, a fortiori, he does not need semantic descriptions. If a direct picture theorist can somehow circumvent Levin's use of the private language argument, then Kripke, himself, can avoid Levin's circularity charge.

Levin disparages the idea that someone might refer and re-refer to pain according to the direct picture by saying that such a person would have to ask himself 'Is this pain, again?' For Levin, such a situation seems incredible since reference is not explained but merely presupposed.11 Levin is clearly correct, provided one's explanation of how reference is achieved requires citing intermediary steps of matching descriptions of referents with referents. This conception of explaining reference, the descriptivist theory of reference, has a venerable past. According to the descriptivist, successful reference is secured only if a speaker matches a list of properties that he associates with a name or description with the referent of the term.12

10Levin, 164.
11Levin, 164.
12Kripke, 'Naming and Necessity,' 280-1.
Thus, on the descriptivist theory, one must have descriptions of the referent at hand if one is to refer. According to the descriptivist, the direct picture is wrong because the speaker would lack the required descriptions of pain.

The question remains, however, whether reference can be explained, or at least accommodated, in some other way in the case of sensation so that the direct picture may be saved. The so-called causal theory, presented by Kripke himself, merits consideration. According to this theory, we explain reference not in terms of matching descriptions with referents, but in terms of a causal chain which goes backwards in time to an actual confrontation of a speaker and referent. For instance, the referentiality of my utterance of 'Aristotle' is explained by saying that my use of the name is an indirect causal result of other persons' utterances of 'Aristotle' which, ideally though not practically can be traced to the man himself and, perhaps also ideally, to the event of Aristotle's being dubbed 'Aristotle.' By adopting the causal theory the advocate of the direct picture can account for the referentiality of 'pain' despite the fact that the term is fixed without designators: being suitably situated with respect to one's sensations enables one to refer and re-refer to them, even without descriptions of them. Perhaps the causal theory saves the direct picture.

The question before us is not whether the causal theory is, in general, better than the descriptivist theory. Nor is the question the more specific one of whether the causal theory is a suitable model for understanding how speakers typically refer to their sensations. Kripke's argument will be vindicated against Levin's charge if it is reasonable to believe that someone, sometime, can perform the Cartesian thought-experiment of imagining pain existing apart from brain states by rigidly designating the reference of 'pain.' Kripke is not committed to any psychological claim about how speakers typically come to fix the reference of 'pain' when they first learn the word. Kripke is not even committed to the view that everyone who tries the Cartesian thought-experiment will be able to fix the reference of 'pain' without using nonrigid descriptions. Thus, Levin must advance the strong thesis that it is implausible to believe that Kripke (or anyone) can refer to pain as the direct picture theorist believes one actually does.

I believe that part of the apparent plausibility of the private language argument's disparagement of the direct picture results from misapplying a basic insight. Doubtless, if any term 'K' is a kind-term, then there must be "criteria" for the correct application of 'K' in the sense of rules that indicate what counts as a K.13 Without rules everything or nothing counts as a K, which is to say that 'K' is not a kind-term. Applying this principle to sensations, it

13J. J. Thomson, 'Private Languages,' American Philosophical
follows that if 'pain' or 'itch' is to be a kind-term in any language (public or "private"), then there must be rules for the correct application of those terms.

The exponent of the private language argument errs, however, when he infers that before speakers can refer to sensations they must know rules or criteria which state the connection of sensations to publicly observable phenomena by being formulated in topic-neutral language. The logical requirement that sensation terms have criteria of correct application can be satisfied by phenomenal rather than topic-neutral rules. Examples of such rules might be 'Whatever feels painful to me (and nothing else) counts as a pain,' and 'Whatever feels itchy to me (ditto) counts as an itch.' These rules are in an obvious sense circular, but they are not logically defective. The conviction that they are stems, I believe, from their uninformativeness. A language learner who did not already understand phenomenal language would not find these rules very useful, certainly not as useful as descriptions that relate sensations to their publicly observable manifestations. But although the need for rules is unquestioned, the need for informative recipes before one can refer to sensations is epistemologically motivated and doubtful. Uninformative rules are rules nonetheless.

We may strengthen our point by noting that although it is necessary that speakers referring to sensations have beliefs about them and follow rules governing the application of sensation terms, it is not necessary that speakers be able to articulate either their belief or the linguistic rules. The distinction between following a rule and being able to formulate that rule is a familiar one in linguistics. The distinction between having a belief and being able to articulate that belief is also clear. A speaker's use of, e.g., 'pain' may be governed by his belief that he has one of those feelings he gets in certain typical situations and the rule that pains are the feelings one gets in such situations. Or, as suggested above, both his belief and the rule may be amenable to translation into phenomenal rather than topic-neutral language--while the speaker himself is unable to articulate either the criterion of pain that he is using or his belief that what he is experiencing satisfies the criterion. Since the speaker need not be able to say things about pain in order to refer to it, the way is clear for him to fix the reference of 'pain' without the use of designators, as the direct picture holds.

This is not to claim that a speaker can pull himself up by his bootstraps with nothing to hold onto. The direct picture theorist requires a not unquestioned psychological premise if his view is to be plausible. The required premise is that each person has a pre-linguistic awareness of sensation types--that is, even before he has learned the relevant sensation words he knows there are differences between,
e.g., tickles, itches, and pains. On this view learning the sensation vocabulary is learning the terms of the public language that match-up with and refer to the speaker's own sensation types. Accordingly, although the learning of sensation terminology is in fact greatly facilitated by the use of nonrigid topic-neutral descriptions this is only a useful, but not necessary device.

In fact, there is an a priori argument (recognized by Quine, among others) that there must be such prelinguistic awareness of determinate types. (Quine calls this 'innate qualitative spacing.')¹⁴ Train a rat to respond to circles, but not to triangles. The rat's ability to learn presupposes that the rat was able to distinguish between circles and triangles before your reinforcement took effect. In a similar vein, many cognitive psychologist feel that there must be a 'language of thought' inherent in language-learners to explain how language learning could ever occur.¹⁵

What would prelinguistic awareness of sensations be like? Would it be some sort of pristine knowledge, unsullied by conceptualization, as is often ridiculed by critics of "the Given?" I do not think that prelinguistic awareness need be seen in this way. Prelinguistic awareness of sensations can simply be knowing how a sensation feels and that it feels different from other sensations without having words to describe the difference. To get an idea of what this might be like for humans, consider an area of sensation where you do not have much vocabulary, e.g., tonal qualities. Remember your elementary school hearing check-ups where you signalled the nurse when the sound in the earphone sounded "different?" This was all I could say about the different tones--that they sounded different--although I had no trouble fixing my thought on the tones.

It may be objected that even if one admits a prelinguistic ability to distinguish sensation types, this does not warrant Kripke's interpretation of the sophisticated Cartesian thought experiment. Doubtless, the thought-experiment itself is a fancy bit of abstract reasoning.

¹⁴W. Quine, 'Linguistics and Philosophy,' in S. Stich, Innate Ideas, Berkeley: University of California Press, (1975), 200: 'The very reinforcement and extinction of responses, so central to behaviorism, depends on prior inequalities in the subject's qualitative spacing ... of stimulations. If the subject is rewarded for responding in a certain way to one stimulation, and punished for thus responding to another stimulation, then his responding in the same way to a third stimulation reflects an inequality in his qualitative spacing of the three stimulations; the third must resemble the first more than the second.

But for Kripke's interpretation of the experiment to withstand Levin's criticism it is only necessary that the speaker be able to concentrate on the way sensations feel without first framing to himself a description of sensations designated nonrigidly. Is it possible that someone concentrate on Mr. A without using nonrigid designators while formulating the thought "I could imagine Mr. A existing without the presence of Mr. B?" If sensations are real, as Kripke, Levin and the nonbehaviorist majority of recent philosophers hold, and if they divide prelinguistically into types, then, for all their epistemological dissimilarities, the two thought-experiments seem equally possible.

Of course, to maintain that Kripke's Cartesian thought-experiment can be performed is not to concede that it succeeds, or even that it is very compelling. As a referee to this journal has noted, explicitly advocating the use of the causal theory of reference to save the direct picture leaves open the possibility that what ultimately causes us to use a sensation term may be, for all we know, just a brain condition. But this possibility, always lurking in the background to undermine the Cartesian intuitions, does not undermine our criticism of Levin's use of the private language argument. Levin wishes to nip Kripke's argument in the bud, but in doing so he asks us to believe that we cannot do something we have good reason to believe we can do.

Putting our discussion in perspective, it is important to distinguish between what is possible for someone attempting the Cartesian thought-experiment and what constitutes the best picture of how language typically works. A speaker left to his own devices would have to be uncommonly clever to refer and re-refer to sensations without fixing the reference of his sensation terms nonrigidly on the basis of instruction from others. We may even concede to the advocates of the private language argument that if the reference of sensation terms were fixed phenomenally, their meaning could not be transmitted from speaker to speaker. (Although there is no reason to concede that the meaning of a sensation word whose reference was fixed phenomenally by me could not be transmitted to you by my use of nonrigid, topic-neutral designators.) But what do these constraints on the communicability of language have to do with Kripke's ability to perform the Cartesian experiment? These are two very different issues.

Our conclusion, then, is that Kripke's argument is not vulnerable to the sort of Wittgensteinian rebuttal that Levin suggests. This is
not to say that the argument is acceptable, of course, since other criticisms may succeed.16

Richard Double
Department of Philosophy
East Carolina University
Greenville, North Carolina 27834

16I am grateful to Peter D. Klein, Robert Van Gulick, Robert Matthews, and an anonymous referee of Philosophy Research Archives for helpful criticisms of earlier drafts of this paper.