Why Lewis' appeal to natural properties fails to solve Kripke's rule-following paradox

(Slightly revised version of the paper published in *Philosophical Studies;*
please cite the published version)

Carla Merino-Rajme
cmerino@unc.edu

Abstract

I consider Lewis' appeal to naturalness to solve Kripke's rule-following paradox. I then present a different interpretation of this paradox and offer reasons for thinking that this is what Kripke had in mind. I argue that Lewis' proposal cannot provide a solution to this version of paradox.

Kripke (1982) famously argued that there are no facts as to what a subject means by her use of a word or sign—there are, for instance, no facts that by my use of '+' I mean addition.[1] If we call these facts *meaning facts*, Kripke's conclusion can be stated as the claim that there are no meaning facts. Instead of truth conditions, utterances of the form "I mean addition by '+'" would have, at most, assertability conditions: i.e., specifications of when it would be appropriate to make such utterances. Because Kripke's conclusion is counterintuitive and because it is arrived at via an argument whose premises are, in contrast, strongly supported by our intuitions, the argument—if sound—leaves us with a paradox.

Lewis (1983) claimed to have a solution to this paradox. He proposed that by adding an objective distinction between natural and unnatural properties to his ontology

---

[1] This is an argument that Kripke *articulates* and not one that he endorses. As Kripke puts it, in presenting this argument he is not speaking for himself. Rather, he presents it as an argument he found in Wittgenstein's *Philosophical Investigations*.

[2] In interpreting Kripke in this way, I am following Gomez-Torrente (2005).

[3] Lewis, (1983, pp. 375-376).

[4] This minimum sense of determinacy is meant to be compatible with other forms of indeterminacy. There could be facts that determine that with my use of '+' I mean addition and not quaddition even if there were no facts that determine that Joe, an unclear case of baldness, falls under the extension of 'bald' or not.

(otherwise composed of properties and classes thereof) he could offer meaning facts and thus provide a direct solution to Kripke's paradox.

In this paper, I argue that under a very reasonable interpretation of Kripke's argument, Lewis' proposal fails to solve the paradox. The problem is that according to this interpretation, meaning facts must be constituted by something capable of guiding the subject in her application of a sign or word so as to accord with what she means by this word or sign. The facts offered by Lewis, however, fail to meet this condition.

I begin Section 1 by introducing Lewis' formulation of, and proposed solution to, this paradox. In Section 2, I argue for an interpretation of the paradox that incorporates the idea that meaning facts must be guiding.[2] As I go on to explain, under this interpretation of Kripke's argument, the paradox can be seen as a more general form of skepticism about the existence of rules that a subject could grasp and follow. Not only did Lewis leave these considerations out of his formulation of the paradox but, as I explain in Section 3, what Lewis took to be meaning facts are not guiding. My conclusion is that given this interpretation of Kripke's argument, even if there are natural properties, we still haven't been provided with meaning facts.

1 The Lewisian proposal

1.1 Lewis' statement of the paradox

Here is the passage where Lewis considers the paradox:

> A well-educated person working arithmetic problems intends to perform addition when he sees the '+' sign. He does not intend to perform quaddition, which is just like addition for small numbers but which yields the answer 5 if any of the numbers to be quadded exceeds a certain bound. Wherefore does he intend to add and not to quadd? Whatever he says and whatever is written in his brain can be perversely (mis)interpreted as instructing him to quadd. And it is not enough to say that his brain state is the causal basis of a disposition to add.

---

[2] In interpreting Kripke in this way, I am following Gomez-Torrente (2005).

> Perhaps it isn't. Perhaps if a test case arose he would abandon his intention, he would neither add nor quadd but instead would put his homework aside and complain that the problems are too hard.[3]

Although we have the strong initial intuition that by using '+' the subject intends to add and not to quadd, his external behavior ('whatever he says'), as well as his internal mental states ('whatever is written in his brain'), are compatible with quaddition. Also, he may be disposed neither to add nor to quadd. However, since there seem to be no other facts to settle what the subject's intention is, we are left with nothing to support the initial intuition that the subject's intention in using '+' is to add.

To solve the paradox, we need to offer facts that determine that with his use of '+' the subject means addition and not quaddition.[4] To see why Lewis thinks he needs to bring in more ontological commitments to provide these facts, let us first briefly consider the relevant parts of his original ontology.

1.2 Lewis' original ontology is insufficient to solve the paradox

The relevant parts of Lewis' ontology for our discussion are properties and intentional states.

1.2.1 Properties

For Lewis, properties are either classes of possible particulars (actual or merely possible particulars) or classes of other classes that ultimately bottom out in these particulars, regardless of how "gerrymandered and miscellaneous and indescribable in thought and

---

[3] Lewis, (1983, pp. 375-376).
[4] This minimum sense of determinacy is meant to be compatible with other forms of indeterminacy. There could be facts that determine that with my use of '+' I mean addition and not quaddition even if there were no facts that determine that Joe, an unclear case of baldness, falls under the extension of 'bald' or not.

language, and […] superfluous in characterizing the world"[5] these classes may be. What is relevant for our purposes is that there is a very large number of them: there are infinitely many properties that any two things have in common and infinitely many more that these two things do not share.[6]

1.2.2 Intentional states

In having an intentional state, according to Lewis, a subject ascribes properties to objects. Lewis takes these properties as constituting the content of the intentional state.[7] Importantly, these properties must fit well with the subject's past and current behavior as well as with features of her other mental states. Whenever a property satisfies this constraint, I will say that it is *fitting*.

Consider, for instance, a subject who intends to mean addition by '+' in writing '68+57=125'. According to the above, the subject would be ascribing a fitting property to an object. In this case, the object would be the triple (68, 57, 125) and the property would be the class composed of the infinite triples of numbers whose members fit the subject's present and past usages of '+' as well as features of her other mental states.[8] The paradox arises, Lewis claims, because despite our confidence in that the subject means addition by her use of '+', there are infinite, wildly divergent properties—such as quaddition—that are fitting. But then this proposal does not deliver facts that determine that the content of the subject's intention is addition and not any of the other equally fitting but wildly divergent properties.

---

[5] Lewis, (1983, p. 346).

[6] According to Lewis, an object has a property when it is a member of the class that that property is. Further, he takes *relations* to be classes of n-tuples (ordered pairs, triples, etc). A relation among things would thus be "a property of 'tuples' of things". (Lewis, 1983, p. 344).

[7] Cfr. Lewis, 1984, pp. 351 and 373. More precisely, according to Lewis, to have an intentional state is to ascribe a property to an object, where the content of the intentional state *is the property thus ascribed*. To believe, desire, or intend something is thus to believe, desire, or intend that I, or some other thing, have a certain property. In referring to intentional states, he writes: "[G]iven the functional roles of the states, the problem is to assign them content. Propositional content, some would say; but I would agree only if the propositions can be taken as egocentric ones, and *I think an 'egocentric proposition' is simply a property*." (Lewis, 1984, p. 373, my emphasis). See also Lewis, 1979. For our purposes, all that matters is that for Lewis the content of an intentional state is constituted by properties.

[8] Recall that according to Lewis, relations are properties. See footnote 6.

1.3 Lewis' new ontology and the 'naive' solution

To solve this paradox, Lewis proposes a defense of what he calls *the naive solution*, which goes as follows. To add is to go on as before when the numbers to be added get bigger. To quadd, however, is to do something different for numbers larger than a certain bound. Since the subject doesn't have a further intention to act differently in different cases, the subject intends addition and not quaddition.[9] This proposal, Lewis further claims, requires the following two additions to his previous ontology.

1.3.1 Naturalness as an ontological distinction over properties

Lewis believes that one of the facts that gives rise to the paradox is that properties are abundant. A way of solving the paradox would then be to distinguish a small subset of them. He thus proposes to draw an ontological distinction that sets apart what he calls *the natural properties*. While this distinction admits of degrees, the most natural properties are sparse: there are only those needed to completely characterize the world.[10] Armed with this distinction, the degree of naturalness of a property could serve to narrow down the many, wildly divergent properties that are equally fitting to just one (or a few) and, in this way, to determine the content of the intentional state.[11]

---

[9] Cfr. Lewis (1983, p. 376).
[10] Cfr. Lewis (1983, p. 345).
[11] Lewis further claims that the natural/unnatural distinction can be drawn in different ways. For example, by adding universals, by taking the fact that some properties are perfectly natural as primitive and adding some principles to derive the less natural ones from these, or by taking as primitive an objective resemblance among certain things but not others.

1.3.2 Naturalness as a constitutive constraint on intentionality

However, merely tacking on naturalness to his ontology is not enough to solve the paradox. As Lewis points out, to be justified in appealing to it, naturalness should also constrain what the content of a subject's intention is.

What is the status of this constraint? Lewis does not take it to be a further intentional state, for this would amount to not having a real constraint. Like any other intentional state, it would be subject to divergent interpretations (maybe by my use of 'natural' I mean natural in some cases but unnatural in some others…).[12] Instead, Lewis takes it as an a priori constitutive constraint on the content of the subject's intention: the content conforms to naturalness because being the most natural (and fitting) property is what it takes to be a constituent of the content of an intentional state.[13]

By conjoining this constraint to the other constraint mentioned earlier—according to which the content of a subject's intentional state is a fitting property—we get that the content of an intentional state is the most *natural* property that best *fits* her external

---

[12] Lewis takes eligible interpretations to be those interpretations of a subject's intentional states constituted by the most natural properties (Lewis 1983, pp. 371-372). He writes (1983, pp. 376-377):

> What is the status of the principles that constrain interpretation, in particular the charitable presumption in favour of eligible content? We must shun several misunderstandings. It is not to be said (1) that as a contingent psychological fact, the contents of our states turn out to be fairly eligible, we mostly believe and desire ourselves to have not-too-unnatural properties. Still less should it be said (2) that we should daringly presuppose this in our interpreting of one another, even if we haven't a shred of evidence for it. Nor should it be said (3) that as a contingent psychological fact we turn out to have states whose content involves some properties rather than others, and that is what makes it so that the former properties are more natural. (This would be a psychologistic theory of naturalness).

[13] Lewis claims: "Only if we have an independent, objective distinction among properties, and we impose the presumption in favour of eligible content *a priori* as a constitutive constraint, does the problem of interpretation have any solution at all" (Lewis, 1983, p. 377).

behavior and other mental states. I will express this by saying that, according to Lewis, the content of a subject's intentional state is the most natural, fitting property.[14]

Given these two additions, Lewis proposes the following as a direct answer to the paradox. The objective distinction between natural and unnatural properties provides the facts that settle what the content of the subject's intention is. While both addition and quaddition are fitting properties, addition is the most natural one and this is the further fact needed to determine that addition and not quaddition constitutes the content of the subject's intention. Moreover, we are (a priori) justified in appealing to these facts because being the most natural fitting property is what it takes to be part of the content of an intentional state. We then get what we wanted: facts we are justified in appealing to and that determine that with her use of '+', the subject means addition and not quaddition.

The property that is the content of an intentional state is thus determined partly by facts having to do with the subject and partly by facts that do not have to do with her. Lewis makes it clear that naturalness only constrains the referent: "the saving constraint [naturalness] concerns the referent—not the referrer, and not any causal channels between the two. [...] Reference consists in part of what we do in language or thought when we refer, but in part it consists in eligibility of the referent. And this eligibility to be referred to is a matter of natural properties".[15] An utterance such as "I mean addition by '+'" would be thus made true by, on the one hand, facts about how well various properties fit with the speaker's behavior and her other mental states and, on the other, facts about the degree of naturalness of the most fitting properties. These facts are what Lewis takes to be meaning facts.

---

[14] As mentioned above (see footnote 12), Lewis takes eligible interpretations to be interpretations of a subject's intentional state constituted by the most natural properties (Lewis 1983, pp. 371-372).
[15] Lewis, (1983, p. 371).

2 A different way of looking at the paradox

2.1 The guidance constraint: meaning facts are guiding

Let us all agree that offering meaning facts would solve the paradox. Kripke, however,

makes it clear that to be meaning facts, the facts offered must be constituted by

something capable of guiding the subject in offering responses that accord with what she

means. Kripke writes: "[n]ormally, when we consider a mathematical rule such as

addition, we think of ourselves as *guided* in our application of it to each new instance.

Just this is the difference between someone who computes new values of a function and

someone who calls out numbers at random."[16]

　　　To illustrate this difference, contrast the following two situations. Towards the

end of your math exam, you are asked to first figure out from a previous problem what

function '@' stands for and then respond to the following questions, whose numbers, we

will further suppose, are big enough that you had not encountered their additions before:

$$243 \ @ \ 832 =$$

$$1{,}290 \ @ \ 428=$$

$$478 \ @ \ 399=$$

$$4{,}783 \ @ \ 326=$$

In the first situation, you figure out that '@' stands for addition, so you go on to perform

correctly the four computations above. In the second situation, before you manage to

---

[16] Kripke, (1982, p. 17), Kripke's emphasis.

figure out what '@' stands for, you realize that you are running out of time. Even if you

knew what '@' meant, you wouldn't have time to do the corresponding computations.

Since in this exam extra points are taken off for leaving questions unanswered, you

decide to spend the last few seconds writing down whatever numbers first come to your

mind as answers to the questions above. As it turns out, by some massive stroke of luck,

the numbers you write down in both cases are exactly the same.

The example serves to illustrate the following. In the first situation, we don't take

it that you simply wrote down whatever numbers you felt like writing. Rather, in offering

your responses, we take it that you were guided by the fact that you meant addition by

'@'. If, say, the stress of the exam had made you lose track of what you were doing,

then—assuming no computational errors were made—attending to the fact that addition

is what you meant by '@' would have been enough for you to see that 1,075 rather than 5

is the right answer to the first question. We thus take it that in the first case there was

something there, available for you to consult, that guided you in offering responses that

accorded with what you meant by '@'. This was not so in the second case. In that case,

despite offering the same numbers, there was nothing that guided you in offering your

responses: you simply wrote down the numbers that came to your mind. Thus, even if in

both cases we would say that your answers were *in accordance with* the addition function,

only in the first case we would also say that your intention to mean addition by '@'

provided you with something that *guided* you in responding in the way you did.

Note that modifying the second situation in the following ways will not suffice to

account for the difference between the cases. Suppose that, unbeknownst to you, your

answers in the second case were *caused* either by a weird disposition of yours to come up

with these numbers when under pressure or by the effects of some external factors—the room temperature, your answers to previous questions, whatever you read the night before. As Kripke explains, neither these unknown dispositions nor the fact that these external factors had effects on you are things available for you to *consult* in figuring out which answers to give. The point is that to be something that merely *caused* your responses is not enough to be a something that *guided* you in offering those responses.

Let me express this idea as follows.

(THE GUIDANCE CONSTRAINT)   Meaning facts—i.e. facts determining what a subject means by a sign or word—must be constituted by something capable of guiding the subject in applying the sign or word in accordance with what she means by this sign or word.

I will also express this idea by saying that meaning facts must be guiding.

But what facts could satisfy the guidance constraint and thus be suitable for being meaning facts? After examining several facts,[17] Kripke concluded that none could. Thus, despite the belief that by intending to use '@' in the first situation of the example above there must have been something available to guide you in responding in the way you did, if Kripke is right, there was no such thing. Your answers in the first situation were as undirected as in the second one: in both cases you were acting blindly. Your responses were just "leaps in the dark."[18]

---

[17] The following are among the putative candidates to be meaning facts that Kripke considers and rejects: a mental image, an expression of the rule, a chain of mental expressions of the rule, the subject's dispositions, a machine's dispositions, an infinite object comprising all the instances of the concept or function in question, a primitive intentional fact, an idea that would serve to grasp a platonic object, a sensation. Cfr. Gómez-Torrente (2005).

[18] Cfr. Kripke (1982, p. 10 and 15).

Kripke makes it clear that the guidance constraint is a condition that meaning facts must meet. For instance, in rejecting dispositional accounts as a solution to the paradox, he argues: "As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition […] that it should *tell* me what I ought to do in each new instance."[19]

The guidance constraint also helps us understand why the paradox is formulated as a problem about following a rule. Take again the case of '+'. My previous applications of this sign are not enough to determine a definite answer to new cases and, hence, they cannot be what guide me. Also, I cannot have the whole infinite set of triples that are instances of addition in my finite mind, so this also cannot be what guides me in responding to new cases. Moreover, as Kripke argues,[20] even if I had a table with the infinite triples that are instances of addition in my mind, it would be useless without a rule. But then, if I am capable of applying '+' to an arbitrarily large number of cases that I have not considered before it must be because, by having this intention, I access a rule that tells me how to respond when confronted with new cases.

---

[19] Kripke, (1982, p. 24), his italics, my underlining. The following formulation of the paradox also makes this point clear:

> Even now as I write, I feel confident that there is something in my mind—the meaning I attach to the 'plus' sign—that *instructs* me what I ought to do in all future cases. I do not *predict* what I *will* do [...]—but *instruct* myself what I ought to do to conform to the meaning.[...]But when I concentrate on what is now in my mind, what instruction can be found there? How can I be said to be acting on the basis of these instructions when I act in the future? The infinitely many cases of the table are not in my mind for my future self to consult. To say that there is a general rule in my mind that tells me how to add in the future is only to throw the problem back on to other rules that also seem to be given only in terms of finitely many cases. What can there be in my mind that I make use of when I act in the future? It seems that the entire idea of meaning vanishes into thin air. (Kripke, 1982, 21-22, his italics, my underlining)

Consider also the following passage:

> The idea that we lack 'direct' access to the facts whether we mean plus or quus is bizarre in any case. Do I not know, directly, and with a fair degree of certainty, that I mean plus? Recall that a fact as to what I mean now is supposed to *justify* my future actions, to make them *inevitable* if I wish to use words with the same meaning with which I used them before. This was our fundamental requirement on a fact as to what I meant..." (Kripke 1982, p. 40, his italics, my underlining).

[20] Cfr. Kripke, (1982, p. 52 fn. 34)

The challenge Kripke raises, then, takes the form of the following more general question: what could constitute a rule I could follow and, thus, something that would guide me in responding to cases I have not considered before? Kripke's paradoxical response is that, although it seems that there must be rules that I can grasp and follow, there are no such things. Kripke makes this point when he claims that the 'private language argument' "is principally to be explicated in terms of the problem of '*following a rule*'".[21] Kripke also expresses this idea when he criticizes other interpreters of the 'private language argument' for treating "the discussion of rules cursorily, virtually not at all, as if it were a minor topic."[22]

As seen in Section 1, Lewis' statement of the paradox, in contrast, is not presented as a problem about following a rule. He is not attempting to find something that a subject could grasp and follow, something that could guide her applications of a word or sign to new cases. Rather, he is concerned with finding facts that could help him narrow down the too many divergent properties that are fitting to one (or a few), which he could then take to constitute the content of the subject's intention.

Following Gómez-Torrente's (2005) interpretation of Kripke's paradox, I have laid out a way of understanding the paradox that—unlike Lewis'—makes sense of Kripke's remarks about rule following. According to this interpretation, meaning facts must be guiding. Could the facts Lewis offered as meaning facts be guiding and, thus, be a solution to the Kripkean version of the paradox that incorporates these rule following considerations?

---

[21] Kripke, (1982, p. vii), my emphasis.
[22] Kripke, (1982, p. 2, fn. 1).

From what has been said, it can already be seen that Lewis' putative meaning facts cannot be guiding. In what follows, I spell this out in more detail.

3. Lewis' proposal fails to meet the guidance constraint

First, note that the most natural fitting property itself cannot be what guides the subject's responses. For within her finite mind the subject cannot have the whole infinite number of particulars that constitute the class that is the property. She cannot, for instance, have the infinite triples that constitute the addition function, or the infinite number of objects that would constitute a color property within her mind.

Further, the subject cannot access the most natural fitting property indirectly either. For this would require her to access the property through some other mental entity. But mental entities are finite and, thus, subject also to multiple, divergent interpretations that raise the same challenge again.[23] This is partly why Kripke points out that, according to Wittgenstein, linguistic comprehension is not a mental entity: mental objects are finite, and thus they can be interpreted in many divergent ways.

In this respect, the Lewisian faces the same problem that Kripke raised against the Platonist. After considering the suggestion that there could be objects outside the subject's mind—i.e. senses—that determine the referents of the subject's words or signs, and ideas in her mind by means of which the subject grasps these senses, Kripke writes:

> [U]ltimately the skeptical problem cannot be evaded, and it arises precisely in the question <u>how the existence in my mind of any mental entity or idea can *constitute* 'grasping' any particular sense rather than another.</u> The idea in my mind is a finite object: can it not be interpreted as determining a quus function [quaddition], rather than a plus function? Of course there may be another idea in

---

[23] Kripke writes: "Such a state [the state of 'meaning addition by "plus"'] would have to be a finite object, contained in our finite minds. It does not consist in my explicitly thinking of each case of the addition table, nor even of my encoding each separate case in the brain: we lack the capacity for that. Yet (§195) "in a *queer* way" each such case already is "in some sense present" […] What can that sense be? Can we conceive of a finite state which *could* not be interpreted in a quus-like way?" (Kripke, 1982, p. 52,).

my mind, which is supposed to constitute its act of *assigning* a particular interpretation of the first idea; but then the problem obviously arises again at this new level (A rule for interpreting a rule again.) And so on. For Wittgenstein, Platonism is largely an unhelpful evasion of the problem of how our finite minds can give rules that are supposed to apply to an infinity of cases. Platonic objects may be self-interpreting, or rather, they may need no interpretation; but ultimately there must be some mental entity involved that raises the skeptical problem.[24]

Kripke does not reject senses because they cannot determine their referents. As he says, we could grant that it is in the nature of a sense to determine its referent. He rejects them instead because they could not be grasped. Because properties are not better off in this respect than senses, they should also be rejected.

Could an appeal to naturalness—the ontological distinction that Lewis brought in to solve the paradox—allow Lewis' putative meaning facts to meet the guidance constraint? To see why not, let us briefly consider an analogous proposal that Kripke rejected. Instead of naturalness, this proposal appeals to simplicity.

Suppose there was an objective measure of the simplicity of a computer program. Couldn't the subject mean by her use of a word or sign the function computed by the simplest program that approximates the subject's physical structure? Kripke argued that even if there were such a measure, and even if it could serve to predict the subject's responses to new cases, the proposal could not give us a solution to the paradox.[25] The problem is that the subject does not guide her responses to cases she hasn't considered before by appealing to this measure: "I do not justify my choice of '125' rather than '5'

---

[24] Kripke, (1982, p. 54, his italics, my underlining).

[25] Again, it is assumed here that the subject makes no computational mistakes. An acceptable solution to the paradox, however, needs to account for the possibility of making a *metalinguistic* mistake—a mistake in what function the word or sign means. A proposal according to which, for instance, an intention to mean addition by '+' is constituted by whatever caused the actual responses given by the subject could not account for this normative aspect the problem. For if this were so it would not be possible for the subject to respond in ways that do not accord with her intention.

as an answer to '68+57' by citing a hypothetical simplicity measure of the type mentioned."[26]

The analogous response in the case of naturalness would be to claim that even if there is an objective measure of the naturalness of a property, and even granting that this measure could serve to predict the subject's responses, the subject doesn't appeal to it to guide her responses. Though I think this is the right response, let me consider a way in which the Lewisian might attempt to resist it.

Against this, we might be reminded that Lewis' proposal is a defense of 'the naive solution': '125' is the right way to respond to '68+57=' because this would be to go on as before in applying the sign '+', whereas to respond '5' would be to do something different. Moreover, the subject would be justified in appealing to naturalness because addition is the most natural fitting property, and being the most natural fitting property is a constitutive constraint on the content of an intentional state. A relevant disanalogy between the appeal to simplicity above and this appeal to naturalness would therefore be this. While it is implausible to claim that the subject appeals to the simplicity of a program to guide her responses, it is plausible to claim that she appeals to the naturalness of a property, especially if we take naturalness to reflect a degree of objective similarity between particulars.

There are two ways of understanding the reply above, none of which helps Lewis' proposal meet the guidance constraint. First, it could be understood as the claim that having an intention to mean addition by '+' involves having an intention of taking '+' to mean the function that maximizes the degree of objective similarity between its instances (from those functions that accord with the subject's previous responses and other features

_____
[26] Kripke, (1982, p. 40).

of her internal states). This, however, wouldn't show that the subject has gained access to something that can guide her applications of '+' to new cases. For if having an intention to mean addition by '+' did not provide her with something that could guide her responses, how could having an intention to use '+' to mean the most natural (fitting) property (or something to this effect) succeed in providing this? This new intention will face the same problem that the original intention faced: namely, by having this intention the subject could neither directly nor indirectly gain access to a rule that can guide her in responding to new cases. This reply would thus fail because it would be an appeal to a rule to explicate a rule.

A second way of understanding the response would be to argue that, rather than a further intention to do what is most natural, what is being appealed to is *the fact* that the subject does what is most natural. According to this response, the subject does not *intend to* go on as before: she simply *goes on* as before. This response is also unhelpful. For it amounts to a concession that the subject's responses are merely *in accordance with* addition but not *guided* in any way by what she means by '+'. [27]

---

[27] Based on Kripke's discussion of an appeal to simplicity to solve the paradox, one might think that an even stronger argument could be offered against Lewis. This argument would aim at establishing that Lewis' proposal is not even a solution *to his own formulation of the paradox*. That is, it would aim to show that, even without appealing to the guidance constraint, Lewis' solution would fail. The argument would proceed in analogy to an objection Kripke raised against a different appeal to simplicity to solve the paradox. It would go as follows.

     Against the claim that one could solve the paradox by pointing out that the hypothesis that the subject means addition is simpler than the hypothesis that the subject means quaddition, Kripke argued that this would beg the question. For in order for simplicity considerations to distinguish between these hypotheses and, by so doing, provide us with meaning facts, one first needs to presuppose that the two competing hypotheses state matters of fact. But to presuppose that claims like
     (*)   The subject intends addition by '+'.
     (**)   The subject intends quaddition by '+'.
are factual is to presuppose that there are meaning facts.

     One might interpret Lewis' appeal to naturalness in a way that is prey to the same objection. For to argue that the subject means addition and not quaddition because the *facts* stated by (*) are more natural than the *facts* stated by (**) would be to beg the question. That is, to take these hypotheses as stating matters of fact would be to assume—not to show—that there are meaning facts.

Thus, an appeal to naturalness does not help the Lewisian meet the guidance constraint.

There are two things to take away from this. First, if this interpretation of the paradox is correct, Kripke's main reason for rejecting proposals that appeal to Platonic senses or to the simplicity of a computer program as solutions to the paradox is that they fail to meet the guidance constraint.[28] Second, there is nothing in Lewis' proposal that allows it to overcome this problem. In particular, his appeal to naturalness is of no help. But then, disregarding the guidance constraint and offering a solution that falls in the same position as those rejected *because they fail to meet that very constraint* would amount to not engaging with the paradox. If Lewis found a problem with Kripke's appeal to the guidance constraint or if he was concerned about a different, weaker problem and found a way to solve *it*, he should have made that clear. Presenting his solution as an answer to "the puzzle where Kripke illustrates Wittgenstein's paradox that 'no course of action could be determined by a rule, because every course of action can be made out to accord with a rule'"[29] would be a mistake.

I don't think, however, that this is the most charitable way of interpreting Lewis. Instead, we can understand him as taking the content of an intentional state as being constituted by those properties that, from those that are fitting, are the most natural. In this way, while naturalness would indeed serve to distinguish between different *properties,* it would not be brought in to distinguish between different meaning *facts.* In this way, his appeal to naturalness would not beg the question—quotes like those included in footnotes 12 and 13 above suggest just this. (For this reason, I think Lewis is best understood as offering a response analogous to the appeal to simplicity that Kripke considers in the footnote to his discussion of simplicity, and not as analogous to the appeal to simplicity that Kripke discusses in the main text). Cfr. Kripke, (1982, pp. 39-40).

[28] Note that adding an ontological measure of the naturalness of a sense analogous to the one Lewis proposed for properties would not help the Platonist meet the guidance constraint more than it helped the Lewisian. Under this interpretation, the guidance constraint would also be Kripke's main reason for rejecting dispositionalist accounts.

[29] Lewis, (1983, p. 375), my emphasis.

3.1 Where does this leave us? A final analogy

Let us understand *zombies* in the familiar way: as creatures very much like us but who are completely deprived of phenomenal states. Though they have beliefs, desires, and intentions and they act in every way like regular people, for them, there is nothing that it is like to see, hear, touch, taste, or smell anything. It is natural to think that zombies are thereby defective, or at least deficient, in that respect.

If Kripke is right, we are perhaps an even more disturbing kind of creature. Each of us would be a meaning-zombie: a creature who, while perhaps having all sorts of phenomenal states, would have no understanding.

It would not console a regular zombie to be told—if she, perhaps per impossible, could understand what is being said to her—that there exist, completely inaccessible to her, phenomenal qualities that correspond to some of her mental states. So long as these phenomenal qualities remain out of her reach, this should make no difference to her: for this zombie, there would still be nothing that is what it's like to see red or smell roses.

Similarly, it would hardly console a meaning-zombie to be told—if she, perhaps per impossible, could understand this one single thing—that there are, completely inaccessible to her, meanings that correspond to some of her mental states. So long as these meanings remain out of her reach, this should make no difference to her: being in one mental state or another will not provide her understanding. While she could know what it's like to see red and smell roses, she would understand nothing. She would be meaning-blind.

4. Conclusion

Given the interpretation of Kripke's paradox that has been defended, offering meaning facts that satisfy the guidance constraint is crucial to solving the paradox. I have argued against Lewis that his proposal does not offer facts that satisfy this constraint. Therefore, if the interpretation favored here is correct, Lewis has failed to provide a way of avoiding Kripke's paradoxical conclusion.

It should be noted, though, that if there were sufficient reasons for rejecting the guidance constraint, Lewis' proposal would indeed be a solution to the paradox. The Lewisian should thus either accept that she has not provided a solution to this paradox—and that therefore one of her reasons for embracing naturalness as an ontological distinction over properties has been undermined—or else she must provide an argument for rejecting the guidance constraint. In this latter case, she must explain why the meaning-zombie should, after all, be consoled when told about the existence of meanings she has no access to.[30]

References

Gómez-Torrente, Mario (2005). "El Wittgenstein de Kripke y la analogía entre reglas y

fundamentos", Diánoia, 50 (55), 55-94.

Kripke, Saul (1982). *Wittgenstein on Rules and Private Language*, Harvard University

Press.

Lewis, David (1979). "Attitudes *De Dicto* and *De Se*", *Philosophical Review,* 88, 513-

543.

Lewis, David (1983). "New work for a theory of universals", *The Australaisan*

Journal of Philosophy, 61(4), 343-377.