The Sceptical Paradox and the Nature of the Self

Tony Cheng, University of California

Abstract

In interpretation, understanding does not become something different. It becomes itself. Such interpretation is grounded existentially in understanding; the latter does not arise from the former. – Martin Heidegger, *Being and Time*, §32

In the present article, I attempt to relate Saul Kripke’s “sceptical paradox” to some issues about the self; specifically, the relation between the self and its mental states and episodes. I start with a brief reconstruction of the paradox, and venture to argue that it relies crucially on a Cartesian model of the self: the sceptic regards the Wittgensteinian “infinite regress of interpretation” as the foundation of his challenge, and this is where he commits the crucial mistake. After the diagnosis, I attempt to sketch my own model of the self and its mental states and episodes. This tentative picture binds meaning and the self together, stressing the subjective aspect of meaning without committing the same fallacy. The solution ventured here is relatively independent of the secondary literatures on the sceptical paradox of following rules, for it aims to provide a new angle to understand and meet the challenge presented by the sceptic.

I.

Both the rule-following considerations in Wittgenstein and the rule-following “paradox” in Kripke are extremely important for contemporary discussions of meaning in mind and language. The latter is also called “the sceptical paradox” concerning meaning and understanding. The literatures on them and related topics are dauntingly enormous. The present article attempts to diagnose the Kripkenstein paradox by connecting it to another important issue in philosophy, that is, the nature of the self. It aims to provide a new angle to understand the whole debate, no matter the positive account of the self sketched in the final section is promising or not. My way of connecting the two issues may appear to be unnatural for those who are familiar with Kripke’s dialectic, but I shall begin by introducing his way of constructing the paradox briefly and make intelligible my approach later.
We normally think that our linguistic behaviours are rule-governed: a concept is a rule; when we use it, we need to follow the rule given by its content. This is where normativity comes in. The rule determines correct and incorrect uses of the concept in question. The Kripkenstein paradox challenges this conception of language. Kripke summarises the paradox as follows:

The Sceptic doubts whether any instructions I gave myself in the past compel (or justify) the answer “125” rather than “5.” He puts the challenge in terms of a sceptical hypothesis about a change in my usage. Perhaps when I used the term “plus” in the past, I always mean quus: by hypothesis I never gave myself any explicit directions that were incompatible with such a hypothesis. (Kripke 1982: 13)

The problem is this: one’s past performances are finite, and we can always fit them into more than one rule, however deviant they might be. There is no principled reason to decide which one is uniquely correct. But if that is so, then the normativity vanishes, for it requires there being a principled reason. Given a set of past behaviours or intentions, including linguistic ones, the sceptic claims that we can always interpret them as confirming infinitely different rules. There is no principle to prevent him from doing so. Kripke points out that his opponent might appeal to some other resources, for example we may just “count” things (in his example, marbles) to settle the issue (Kripke 1982: 15–16). This does not work, according to Kripke’s Wittgensteinian reply, for the very trouble for “plus” is there for any other notion: if we want to settle the issue by “counting” things, how can we prevent the sceptic from applying the same manoeuvre, saying that it is equally possible that what we are doing is “quounting” (Kripke 1982: 16)?

Here Kripke is applying again the infinite regress of interpreting rules. This time, the regress is not within a single symbol, say “+”, but between different symbols. The basic insight, if any, is the same: a rule is never self-interpreting; we can always assign infinite interpretations to it, and if we attempt to interpret it with other rules, the series of interpretation is also infinite. Hence, the paradox.¹

Let me now briefly introduce Kripke’s “sceptical solution”. He says this kind of solution “begins on the contrary by conceding that the sceptic’s negative assertions are unanswerable. Nevertheless our ordinary

1. Actually, “infinity” is not always a problem, especially when we notice examples from mathematics. I argue that the infinity in the present context is indeed a problem because it violates the normativity of meaning; it prevents us from deciding non-arbitrarily, which interpretation is correct.
practice or belief is justified because . . . it need not require the justification the sceptic has shown to be untenable” (Kripke 1982: 66, his italics). He goes on to say that “Wittgenstein proposes a picture of language based, not on truth conditions, but on assertability conditions or justification conditions: under what circumstances are we allowed to make a given assertion” (Kripke 1982: 74, his italics)? He then concludes by saying that “[t]he success of the practices . . . depends on the brute empirical fact that we agree with each other in our responses” (Kripke 1982: 109, italics added). This is the general guise of Kripke’s sceptical solution.2

In the next section, I will argue that we should not treat understanding as a kind of interpretation at the very beginning. In particular, I will show the implausibility of this assumption by placing it into the context of philosophy of mind. In section four I will further relate the discussion of rule-following to the issue of the self.

III.

As we have seen, the first principle of the Kripkenstein paradox is the infinite regress of interpreting rules: we can equally interpret the “+” sign to mean “plus” or “quus” or infinitely other functions, and if we attempt to determine its meaning by invoking another notion, say “count”, the sceptic can still interpret it as “quount”, which again is a deviant interpretation. This infinite regress is the foundation of the paradox. In this section, I attempt to prevent the sceptic from getting off the ground by undermining this foundation.

I would like to begin by a quotation from Wittgenstein:

It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us at least for a moment, until we thought of yet another standing behind it. What this shews is that there is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call “obeying the rule” and “going against it” in actual cases. (Wittgenstein 2001: §201)

2. Readers might have noticed that I do not have any surprising interpretation of Kripke. One thing to be noted, however, is that I do not accept George Wilson’s reductio interpretation of the paradox (1994), which is supposed to cast doubt on the standard interpretation I adopted here for the sake of argument. I side with Alexander Miller in thinking that Wilson’s interpretation violates Kripke’s distinction between the straight solution and the sceptical solution, which seems to me perfectly legitimate. See Rule-Following and Meaning (2002). Obviously, this should not be regarded as a sweeping objection to Wilson, whose discussions are abundant and delicate. I do not discuss his interpretation here because that will lead me too far away from my central concern, i.e., putting meaning and the self back together to resolve the paradox. This remark applies to all other different interpretations to Kripke: I do not pretend that the summary above is anything new or special.

© 2015 John Wiley & Sons Ltd
Many authors have taken this to be a *reductio ad absurdum* of Kripke’s assumption, and to be evidence that Kripke fails to appreciate the second paragraph of §201: “rule-following behaviours” do exist, so if the “understanding as interpretation” model fails to preserve it, the model must be given up. Here, I tend to agree with this standard reaction, but I am not simply accepting this and using it to argue against Kripke. Other authors (e.g., Winch 1983; McGinn 1984) have pointed out this possibility, and there are also philosophers siding with Kripke at this point (e.g., Weir 2007). It is not my purpose to take a side on this interpretative issue, and nothing below hinges on this.

Now bracketing the exegetical issues of Kripke’s reading of §201, consider the way Wittgenstein introduces the infinite regress with this following figure: imagine an ambiguous picture containing a human-like item and a slope-like line, and then we can interpret it as a person climbing up or sliding down, or even no human being there. There does not seem to be a principled reason to decide which hypothesis is uniquely correct.

To fill in the details or invoke other pictures would not help, for we can still interpret those pictures and make them about anything. So far, so good. But why think that we have this very situation when the topic is our thoughts? When one entertains the plus function in my thought, one seldom interprets what one is thinking. Moreover, if one tries to interpret it by *looking inward* and trying hard to do some interpretation, one would presumably find that s/he cannot do anything like the above scenario suggests. The above scenario is *perceptual* in character: we *perceive* a picture in front of us, and go on to *interpret* it in many different ways. The precondition of doing so – of interpreting the picture wildly – is that the relation between the picture and us is *perceptual*. If, on the contrary, the relation is not perceptual, but some more intimate sort, there is no reason to think that we can do wild interpretations, or indeed any interpretation, in the scenario.

Now a natural doubt against me is that the sceptical paradox is not about perception, so even if what I just said is true, it must be a serious digression. This doubt is perfectly legitimate, but I think what it shows is just that the notion of “perception” is not essential in my diagnosis. What is crucial is that the case I just described involves a *distance* between a subject and an *object*: a perceiving self and a picture in front of her. What

---

3. This should be clear when we notice that many philosophers who espouse the perceptual model of self-knowledge do not invoke the literal meaning of “perception”. The literal meaning of perception is fixed by our relations to the external world, so if those philosophers do use the notion of perception literally in the context of self-knowledge, their position would be outright self-contradictory.
generates the possibility of wild interpretations is the distance in question, and that the fact that the above scenario involves perception is just a contingent one. I chose that example mainly because it vividly illustrates what I want to stress. Now, we shall go back to Kripke’s own text and apply the lesson we got above to it.

In many places, the sceptical paradox looks essentially temporal in character, for Kripke often contrasts the past and the present. What I want to take issue here is about his remarks on past intentions and mental history, but the point to be argued is neutral about whether the temporal dimension is essential to the paradox. Kripke explicitly states the challenge in this way: “[h]ow do I know that ‘68 + 57,’ as I meant ‘plus’ in the past, should denote 125” (Kripke 1982: 12)? He goes on to ask “why I now believe that by ‘plus’ in the past, I meant addition rather than quaddition” (Kripke 1982: 12)? A perfectly reasonable reaction to this is “why not?” In order to have some genuine doubt here, as the sceptic does, we need to distance ourselves to our past intentions, treat them as objects, so as to make room for deviant interpretations. The same problem occurs when Kripke says that “[b]ut I can doubt that my past usage of ‘plus’ denoted plus” (Kripke 1982: 13). Yes, but only on the assumption that there is some space for us to conduct alternative interpretations. We have no reason to think that there is such a space in the case of understanding, however. Generally speaking, when I entertain “plus” in my thought, there is no distance between it and myself, so to speak. The relation here is not between a subject and an object at all.

Let us apply this thought to another passage. Kripke writes: “my past mental history is equally compatible with the hypothesis that I meant quus, and therefore should have said ‘5’.” (Kripke 1982: 15; and he says something similar in 21). But the talk about compatibility presupposes we have different interpretations to begin with: there is some distance for us to assign different interpretations on it, and all of them are compatible with my mental history. My reply here is pretty much similar to Kripke’s own reply to the simplicity suggestion, that simplicity presupposes we have a set of genuine hypotheses to be evaluated (Kripke 1982: 38). Here I want to say, pace Kripke, that to invoke compatibility presupposes we have a set of interpretations at hand, and that in turn presupposes there is some distance between my thoughts and myself, i.e., treating my thoughts as objects for a homunculus – me. To repeat, the relation between my thoughts/past intentions and me is a more intimate sort, which precludes the possibility of (deviant) interpretations. In order for my thoughts to be meaningful, interpretations are not required. To think otherwise is to

---

4. I will explain more about the homunculus presently.

© 2015 John Wiley & Sons Ltd
violate the crucial difference between the intentionality of mind and of language.5

Maybe the sceptic will simply go back to his original query: “just cite any fact to refute other interpretations!” This way of posing the challenge, however, presupposes certain reductionism about meaning: we can grant the supervenience thesis about meaning, but to ask us “cite the fact” is asking something more than that. “Supervenience” is an ontological thesis; it does not further require that we are able to cite the very fact. I cannot go into the reductionism debate here, but it seems plausible to reply to the sceptic that the requirement of “citing the fact” commits reductionism, and we have no reason to accept it without further arguments from the sceptic.

Let me put this in another way. The picture Kripke offers suggests that when we entertain a thought, we need to consult our own past intentions in using that symbol. The falsity of this way of thinking is that this makes ownership of thoughts a myth. If we need to consult our mental history whenever we entertain relevant thoughts, the difference between spontaneously entertaining our own thoughts and attributing thoughts to our fellow speakers vanishes.

Some might think that my move here is Cartesian in the bad sense6: I claim that there is certain intimate relation between my thought and myself, and this is supposed to answer the sceptic. But is this not exactly what the sceptic attempts to refute? He shows that we thought we are sure about our own meanings, but it turns out that we are not, because interpretations are required even in our own cases.

But who is Cartesian? The sceptic assumes that when we entertain the plus function in our thoughts, we always have to embark on interpretations. The precondition of this, as I argued above, is that we conceive our knowledge of our own thoughts, i.e. self-knowledge, in a quasi-perceptual way.7 According to this model, we are just like homunculi, sitting in front of our mental screens and wondering what the relations

5. It is helpful to think this in John Searle’s (1983) terms. Without buying into his philosophy of language and mind as a whole, still we should not treat our thoughts’ intentionality as derivative ones, just like linguistic intentionality; the former is not observer relative. But to treat them as objects is to think they are observer relative.

6. Descartes himself, in the Fifth Meditation, points out the problem of this conception, so to label it “Cartesian” is not really fair to him historically. Here, I am using the term only figuratively.

7. Kripke emphasises that his point is ontological rather than epistemic (Kripke 1982: 14), and this might suggest that my talk about self-knowledge is a red herring. In fact, it is not. What I am trying to argue is that in order for the whole argumentation to get off the ground, a certain model of self-knowledge has to be presupposed. Now I find something false in that model, so the sceptical case is in fact a non-starter. Nothing here involves a conflation between ontology and epistemology.

© 2015 John Wiley & Sons Ltd
between things on the screens and external state of affairs are. This is the famous “homunculus fallacy” in the philosophy of mind: if we attempt to explain our mental faculties by postulating homunculi inside of our minds, then we explain nothing, for we can then ask how the homunculi achieve those mental functions.8 Here we get another regress: the infinite regress of homunculi. So actually, we can have two challenges to Kripke at this point. First, the “distance” problem I introduced above; second, the “lack of explanatory power” problem, that postulating homunculi just postpones the original issue. Now, notice that almost no one in the philosophy of mind accepts the homunculus explanation, for it conflicts with the fact that we do have those mental faculties. Then why cannot we say something parallel in the case of the sceptical paradox? Why cannot we just say that both the infinite regress of interpretation and the infinite regress of homunculus are incompatible with the fact that we do mean things, so the “understanding as interpretation” model should thereby be rejected?

One reason might be this: many philosophers do not see that actually Kripke assumes a Cartesian model when he elaborates the sceptical paradox; it is not obvious because the paradox is primarily about language, not about self-knowledge.9 They mistakenly think, by contrast, that the line I take is Cartesian, for I claim that there is certainly more intimate relation between my thoughts and myself. But what I insist is just the asymmetry between the first-person and the third-person knowledge. I do not need to assume that the former is infallible or incorrigible.10

To repeat, from many of Kripke’s remarks, we can see that he tacitly assumes some dubious distance between our thoughts and ourselves, and that generates the problematic infinite regress of interpretation. And through this reflection, we discover that the way Kripke sets up his sceptical case is essentially a homunculus model, and this generates a further problem for him, that is, the problematic infinite regress of homunculus explanation. Both of them suggest that the sceptical paradox should be blocked at the very beginning.

8. For the connections between the homunculus model, Cartesian Theater, and Cartesian Materialism, see Daniel Dennett (1992). For an exposition of it, see Susan Blackmore (2003).

9. Similar considerations apply to Kripke’s (1972) argument against the mind–body identity theory. There, he relies on the possibility of imagining C-fibre’s firing without pain, and vice versa. This possibility of imagination embodies a Cartesian line of thought.

10. Here, I side with Colin McGinn (1984). Some might think that the picture I share with McGinn might neglect the social element of meaning. Admittedly, I do not say anything about the social settings for meaning, but I think what I say here leaves room for that: the asymmetry I stressed does not imply anything like a private language. The importance of fitting the social element is recognised, to be sure.
To identify Kripke’s assumption with the homunculus fallacy is a way to make the problem vivid. Here is another way to state the same problem. If we assume the homunculus model, then the screens would be mental intermediaries between the world and us. The images on the screen are not intrinsically related to the world, according to the sceptic. Those images are free-standing objects; to use a Wittgensteinian locution, they just “stand there as signposts”. We need to interpret them in order to form their relations to the world. Put this way, the model becomes more and more like the sense-datum theory: what we directly perceive are sense-data, mental intermediaries. They are by themselves non-intentional; that is, devoid of intentionality. They need to be interpreted to be about external state of affairs. That does not seem to be a good picture to have. I am not saying that Kripke’s picture is exactly a version of the sense-datum theory; if I were, I need to insist on the notion of perception, but that is beside the point, as we have seen. Nor am I claiming that the sense-datum theory is hopeless. What I do want to stress is that when we uncover the nature of the sceptical paradox, we find a homunculus model inherited in it, and we can further burden it by identifying its close connection to the sense-datum theory. As a result, the “foundation” of Kripke’s whole argumentations becomes shakier than it originally appears.

Some might reply for Kripke in this way. Let us grant that in some places Kripke does implicitly commit the Cartesian thinking in the way I described. Still, he can simply insist on his demand on justification: what fact is supposed to be cited to justify the answer in a particular case? Kripke says that “[w]hen I respond in one way rather than another to such a problem as ‘68 + 57,’ I can have no justification for one response rather than another” (Kripke 1982: 21, similar remarks come again in 23 and 37). My answer to this challenge, not surprisingly, is that the question should not arise in the first place. Or put in another way, the challenge presupposes that one’s past intentions are objects to be consulted whenever one entertains relevant thoughts. The competing interpretations arise from regarding thoughts as free-standing meaningless objects, and I have shown that this is a wrong model for the ownership between the self and its thoughts.

11. See Howard Robinson (1994) for a recent defence of the theory. For an even more drastic position, see John Foster (2000).
12. John McDowell also connects his criticisms to Kripke to a Cartesian way of thinking, but in a more complex way. See his “Meaning and Intentionality in Wittgenstein’s Later Philosophy” (McDowell 1993). He further links this to his view on singular thought; see his “Singular Thought and the Extent of Inner Space” (McDowell 1986) and “Putnam on Mind and Meaning” (McDowell 1992). Donald Davidson (1997) elaborates a similar line of thought in his “The Emergence of Thought”.
Wittgenstein writes in *The Blue and Brown Books*:

> We are tempted to think that the action of language consists of two parts; an inorganic part, the handling of signs, and an organic part, which we may call understanding these signs, meaning them, interpreting them, thinking. (Wittgenstein 1958: 3, italics added)

I take this to be on the same line as my argumentations in this section. It is tempting to start by treating words or thoughts as *object*, as *inorganic* things, and add *interpretations* on them in order to make them organic. This cannot be a correct model for the following reason:

> [W]hy should the written sign plus the painted image be alive if the written sign alone was dead? – In fact, as soon as you think of replacing the mental image by, say, a painted one, and as soon as the image thereby loses its occult character, it ceases to seem to impart any life to the sentence at all. (Wittgenstein 1958: 5)

The point seems to be that if we start with some inorganic object, adding anything else would not solve the problem, for they are just other pieces of dead furniture. This picture is bad enough even without Kripke’s elaborations of the sceptical paradox.

IV.

If I am right so far, at least in the main line, then the root of the sceptical paradox is the Cartesian model of the self and its mental states and episodes. I am not saying that Kripke’s detailed discussions of the paradox are not valuable, or without other problems; all I want to argue is that the paradox seems to in some way presuppose a problematic model of the self, so if we want to dissolve the paradox completely, we need to develop a plausible alternative to that model. That will be the main task of the present section.

If the model at fault is Cartesian in some sense, a natural response is to trace the history of philosophy and consider a Kantian model. Like my previous use of the term “Cartesian”, and the term “Platonic” used in Kripke’s book and many other literatures in analytic philosophy, the label “Kantian” is not meant to denote Kant’s philosophy in particular. Instead, it denotes a way of conceiving certain phenomena we want to understand (in the present case, the self and its mental states and episodes). Still, it is helpful to look at some of Kant’s remarks on relevant issues. In the first Paralogism, Kant argues against Cartesian substantial self, i.e., self as substance. In the third, he further argues that “[t]he identity of the consciousness of myself at different times is therefore only a *formal* condition of my thoughts and their coherence…” (Kant 2003: A363,
emphasis added). The self must be able “to accompany all my representations” (Kant 2003: B131).

Now, if what we are doing is interpreting Kant’s philosophy, we need to make an effort to relate the passages I quoted and fill in many details. This is not my purpose, however. Therefore, I suggest a pictorial way of understanding it. The Cartesian model, at least as I explained in the previous section, has it that the self sits behind a mental screen that plays some passing shows; there is some distance between the self and its mental states and episodes, which means we can conceive the self independent of the mental goings-on. On the contrary, the Kantian model, at least as I explained in the present section, has it that it must be able to accompany all of its mental episodes, which at least hints that the self and its mental goings-on should not be conceived independent of each other. The mental episodes are not something outside the self, standing there like signposts, but are parts of the self.

This is relevant to the sceptical paradox for the following reason. Recall that the crux of the paradox is the infinite regress of interpretations, and I have argued that the precondition of it is the distance between the self and its mental states and episodes. Now the Kantian formal self seems to be a way out of it, for it does not treat mental states and episodes as something outside the self. In this way, it blocks the possibility of interpretations, deviant or not, in the first place.

Or consider a simple schema, “I think that P”. In the Kantian framework, “I think” is apperception/self-consciousness; “I” denotes the self, and “P” is any given representation accompanied by one’s self-consciousness. Propositional attitudes are not objects before the self; rather, the self is part of those attitudes; they are constitutively inseparable, as the simple schema shows. Once we remove the false Cartesian assumption about the self, scepticism about meaning can be dissolved.

Unfortunately, this Kantian model is not without its problem. The main problem for the Cartesian is that the self is too substantial in the model. Now the opposite problem for the Kantian is that the self seems to be too formal. As John McDowell puts it, “[the Kantian self] would not provide for it to conceive itself . . . as a bodily element in objective reality – as a bodily presence in the world” (McDowell, 1996: 103, my emphasis). The complaint seems to be that it goes too far to say that the self is only a formal condition, for after all, the self has its bodily presence in the world; it must be substantial in some sense, although not the Cartesian one. I will not go into McDowell’s discussions here, for that would lead us too far. All I want to stress are, first, the issue about meaning is in a significant way connected to the issue about the self, and second, more boldly, a middle course between the Cartesian and the Kantian is
necessitated by the ungrounded way of thinking in the sceptical paradox and the inability to account for our bodily presence in the world.13

References


13. This paper originated from extensive discussions with Barry Stroud, John Campbell and Alva Noë back in 2006. An early version of the paper was presented at Soochow Early Analytic Philosophy conference in 2007, where I gained various comments from Anthony Anderson, Richard Fumerton, Danielle Macbeth, David McCarty, Edwin Mares, Christopher Pincock, Scott Soames and Kenneth Williford, among others.


