KRIPKE'S WITTGENSTEIN AND KRIPKE'S CAUSAL-HISTORICAL PICTURE OF REFERENCE[*]

Alexander Miller, University of Otago

Proper names are a confounded business. For example, suppose I wanted to call this chair Jacob. What did I really give the name to? The shape or the chair? … The possibility of giving names to things presupposes very complicated experiences.

(Wittgenstein to Friedrich Waismann 25 December 1929, as recorded in *Ludwig Wittgenstein and the Vienna Circle*)[1]

1. Introduction

Saul Kripke's *Naming and Necessity* (1981, hereafter "NN") and *Wittgenstein on Rules and Private Language* (1982, hereafter "WRPL") have both played a pivotal role in shaping late 20th Century and contemporary philosophy of language and mind. In Lecture II of NN, Kripke outlines a "picture" of how the references of certain sorts of natural language expressions are determined: a picture, in other words, of the facts that determine reference. In Chapter 2 of WRPL, on the other hand, Kripke presents a number of arguments, suggested to him by a reading of the later Wittgenstein's "rule-following considerations", the conclusion of which is that there are no facts capable of determining reference. There thus appears to be a straightforward conflict between Kripke's two famous texts. In this chapter I will be concerned with the nature of the relationship between them and with the question whether this appearance of conflict is genuine or not.

I'll proceed as follows. In Section 2, I'll give a brief reminder of the sceptical argument developed by Kripke in Chapter 2 of WRPL, followed in section 3 by a very brief reminder of the causal-historical picture of reference outlined by Kripke in NN. In Sections 4 and 5, I'll consider arguments by (respectively) Colin McGinn and Penelope Maddy, the intended upshot of which is that (something like) the causal-historical picture of reference in NN is capable of disabling the sceptical argument of WRPL. I'll argue that McGinn and Maddy are both wrong, and that the causal-historical picture of reference fails to neutralise Kripke's Wittgenstein-inspired scepticism about the existence of facts capable of determining reference. I'll go on in Section 6 to argue that the underlying source of the conflict between Kripke's two books is actually very direct: the causal picture of reference defended in NN and the dispositional theory of meaning attached in WRPL are arguably instances of the same kind of naturalistic approach to the determination of meaning and reference. It seems that NN, in its espousal of a causal-historical picture of reference, and WRPL, in its scepticism about dispositional theories of meaning, respectively advocate and reject instances of the same general form of naturalism about meaning and reference. In Section 7, I'll develop the idea that the appearance of contradiction

---

[1] Wittgenstein (1979: 51).

here might be obviated by drawing on an analogy with a pair of positions in metaethics and normative ethics respectively. It is possible, without inconsistency, to hold an expressivist position in metaethics at the same time as a broadly utilitarian position in normative ethics (in which we reject at the metaethical level the idea that moral judgements express beliefs about utility while holding at the level of normative ethics that utility is the standard of right action). In similar fashion, we could perhaps view WRPL as advocating a kind of expressivism about semantic judgement at the meta-level (on which ascriptions of meaning do not express beliefs about speakers' dispositions), while advocating a causal picture of reference at the level of first-order semantics (on which speakers' dispositions provide standards for selecting the referents of linguistic expressions), so that there is in fact no contradiction between the positions defended in the two books. In section 8, however, I'll argue that promising as it sounds, although this maneuver may work in the ethical case, there are specific considerations which suggest that it is bound to fail in the case of meaning and reference. The conclusion of the chapter, then, is that ultimately there is indeed a direct conflict between the position defended in NN and the view suggested to Kripke by his reading of Wittgenstein in WRPL.

2. Kripke's Wittgenstein's Sceptical Argument

In Chapter 2 of WRPL Kripke's Wittgenstein's sceptic (hereafter "KW's sceptic") argues for a "sceptical paradox": there are no facts in virtue of which ascriptions of meaning, such as "Jones means *addition* by '+'", are true . Since the argument generalises, there are no facts in virtue of which any speaker attaches a determinate meaning to any of the expressions of his language.

It is worth noting that although Kripke doesn't highlight it himself, the sceptic's argument exploits the idea that the meaning of an expression determines its reference: any fact which constitutes the meaning of the "+"-sign must determine that it refers to the *addition* function. Given this principle, the sceptic can argue that since there are no suitable facts capable of determining that "+" refers to the *addition* (as opposed – see below - to the *quaddition*) function, it follows that there are no suitable facts capable of conferring truth on an ascription of meaning to the "+"-sign.[2]

Suppose that Jones is asked to answer the query "68 + 57 = ?", a calculation that he has never before been asked to perform and a calculation in which both of the arguments (here 68, 57) are larger than any of the numbers in calculations he has performed previously. (We know that such an example and threshold exist given that Jones, a finite creature, has performed only finitely many computations in the past).

Jones confidently answers "125". This answer seems to be correct in two ways: first, it is *arithmetically* correct, given that the *number* 125 is indeed the *sum* of the *numbers* 68 and 57; it is also *metalinguistically* correct, given that the "+" sign denotes the *addition* function (the function that gives the sum of two numbers presented to the function as arguments). KW's sceptic argues that Jones's confidence that "125" is the correct answer is not well-placed:

---

[2] Kripke puts it in terms of the *denotation* of the "+"-sign at the start of Chapter 2 (1982, pp.7-8), but later puts it in terms of reference (1982, p.54).

This sceptic questions Jones's certainty about his answer, in ... the "metalinguistic" sense. Perhaps, he suggests, as Jones used the term "plus" in the past, the answer he intended for "68 + 57" should have been "5"! Of course the sceptic's suggestion is obviously insane. Jones's initial response to such a suggestion might be that the challenger should go back to school and learn to add. Let the challenger, however, continue. After all, he says, if Jones is now so confident that, as he used the symbol "+", his intention was that "68 + 57" should turn out to denote 125, this cannot be because he explicitly gave himself instructions that 125 is the result of performing the addition in this particular instance. By hypothesis, he did no such thing. But of course the idea is that, in this new instance, he should apply the very same function or rule that he applied so many times in the past. But who is to say what function this was? In the past he gave himself only a finite number of examples instantiating this function. All, we have supposed, involved numbers smaller than 57. So perhaps in the past he used "plus" and "+" to denote a function which we will call "quus" and symbolize by "⊕". It is defined by

$$x \oplus y \quad = x + y, \text{ if } x, y < 57$$

$$= 5, \text{ otherwise.}$$

Who is to say that this is not the function Jones previously meant by "+"? (adapted from 1982, pp.8-9).

The challenge posed by KW's sceptic is: find a fact about Jones which constitutes his meaning *addition* rather than *quaddition* by the "+" sign, which makes it the case that "+" as he uses it refers to the *addition* (and not the *quaddition*) function. Kripke (1982, p.11, p.26) imposes two constraints on candidate responses to this challenge. First, any response must provide an account of the type of fact that makes it the case that Jones denotes *addition* (and not a *quaddition*-like function) by his use of "+". Second, it must be possible to "read off" from this fact what constitutes *correct* and *incorrect* use of the "+" sign.[3] In other words, it must show why Jones is *justified* in giving the answer "125" to the query "68+57=?" and why "125" is the answer Jones *ought* to give.[4]

   In challenging us to find a suitable meaning-constituting – or reference-determining - fact KW's sceptic allows us unlimited and omniscient access to facts of two types: (a) facts about the previous linguistic behaviour and behavioural dispositions of Jones and fellow members of his speech community; and (b) facts about Jones's mental history and "inner life". The sceptic considers and rejects a variety of possible meaning-constituting (reference-determining) facts. These include facts about: Jones's *previous behaviour* (1982, pp.7-15); *general thoughts or instructions* that Jones might have given himself (1982, 15-17); how Jones is *disposed to use* the "+"-sign (1982, pp.22-38); the relative *simplicity* of hypotheses about what Jones means or refers to by the "+"sign (1982, pp.38-40); Jones's *qualitative, introspectible mental states (including mental images*) (1982, pp.41-51); *sui generis and irreducible mental states* of Jones's that are "not to be assimilated to sensations or headaches or any 'qualitative' states"(1982, pp.51-53); and Jones's relation to *objective, Fregean*

---

[3] For our purposes here, giving a response to a query that includes the "+" sign counts as a use of that sign.

[4] The nature and plausibility of this second, "normativity", constraint is a controversial matter: for an overview of some of the relevant literature see Miller (2022) and section 4 of Miller and Sultanescu (2022).

senses (1982, pp.53-54). KW's sceptic insists that in each case the proposed meaning-constituting (or reference-determining) fact violates one or both of the two constraints imposed on candidate responses. It seems, then, that facts about meaning – and determinate reference - have as Kripke puts it, "vanished into thin air"(1982, p.22).

For the purposes of our discussion in this chapter, the most important of the responses to KW's sceptic to consider is the dispositionalist response. According to a simple form of the dispositional theory, Jones's meaning addition by "+" is constituted by the fact that he is disposed to respond to queries of the form "x + y = ?" by giving the sum of the numbers denoted by "x" and "y".

KW's sceptic argues that this response fails to determine the addition function as the referent of "+" (1982, pp.26-7). Jones's dispositions to respond to arithmetical queries are finite: some numbers are simply so large that Jones's brain lacks the computational wherewithal to process calculations involving them and indeed so large that Jones will be dead long before he is even able to grasp them. Define the *skaddition* function as follows:

> x ψ y   = x +y, if x, y are small enough that Jones can grasp them
>   and perform calculations involving them
>
>   = 5, otherwise

The dispositions that Jones actually possesses are consistent with "+" as he uses it referring to the skaddition function (and indeed an open-ended and potentially infinite set of functions with similar singularities), so they fail to determine the addition function as the referent of "+".

As Kripke observes (1982, pp.27-32) the dispositionalist about meaning may try to respond to this objection by invoking ceteris paribus or ideal conditions for the manifestation of the relevant dispositions: for astronomically large numbers I'll die before I finish responding to arithmetical queries involving them, but if we include in the ideal conditions a proviso to the effect that I live long enough, we can say that I'm disposed to respond in ideal conditions with the *sum* (and not the *skum*) of the relevant numbers. So according to this more sophisticated form of dispositionalism, Jones's meaning addition by "+" is constituted by the fact that he is disposed in ideal conditions to respond to queries of the form "x + y = ?" with the sum of the relevant numbers.

Kripke argues that this move fails as a defence of dispositionalism. The dispositionalist aspires to give a *reductive* account of meaning: the meaning-constituting dispositions – and the attendant ideal conditions – are to be specified in wholly non-semantic and non-intentional terms. How plausible is it that this can be done? The obtaining of the ideal conditions has to guarantee that Jones does not respond with something other than the sum of the relevant numbers. But note that if Jones means *subtraction* by "+", he'll respond with "11" and not "125". And if he means *multiplication* by "+" he'll respond with "3876". Indeed, the set of functions *f* such that if Jones meant *f* by "+" he would respond with something other than the sum has an infinite number of members. So the obtaining of the ideal conditions has to guarantee that none of these eventualities obtains. How could a set of conditions specified in entirely non-semantic and non-intentional terms guarantee that an open ended and

potentially infinite set of alternative meaning hypotheses (in which Jones means one of these other functions) fails to obtain? Even if we include provisos guaranteeing that Jones lives long enough and has sufficient cognitive capacity to carry out the calculation, and so on, without an explicitly semantic proviso to the effect that Jones means *addition* by "+", we'll fail to guarantee that Jones will in the ideal conditions respond with the sum. And we can't include such an explicitly semantic proviso without rendering the account circular in a way that stymies its reductionist aspirations.[5] The reductive dispositionalist view thus fails to satisfy the first of the two constraints on answers to KW's sceptic: it fails to determine *addition* (as opposed to some *quaddition*-like alternative) as the function denoted by "+".[6]

If reductive dispositionalism and the other sorts of response all fail to provide a plausible answer to the sceptic, it seems that we can find no fact capable of making in true that a speaker means something by the expressions of his language.[7] This threatens to spiral into the "insane and intolerable" (1982, p.60) conclusion that "all language is meaningless" (1982, p.71). Kripke describes this conclusion as "incredible and self-defeating" (ibid.), and KW tries

---

[5] I take this to be the sort of worry Kripke is pointing to in his discussion of the more sophisticated form of dispositionalism (1982, pp.27-32). For further elaboration, see Boghossian (1989, pp.164-177). This sort of consideration still poses a problem for views on which facts about reductively characterised dispositions are held to metaphysically necessitate meaning facts. For example, Scott Soames (1997) claims that the Kripkean point (familiar from NN) "that many necessary consequences of propositions are not a priori consequences of them" (1997, p.231) opens up space for the idea that some non-intentional fact about a speaker's dispositions metaphysically necessitates the fact that he means addition by "+". Soames concedes that Kripke gives reasons (concerning, broadly, the "normativity" of meaning (see note 5 below)) for thinking that the meaning fact cannot be an a priori consequence of the reductively characterised dispositional fact; but he argues that if we switch to thinking of the putative relationship between the two facts as one of non-a priori consquence Kripke has simply given us no reasons to doubt the plausibility of the view. He writes:

> [KW's sceptic] has insisted that if I meant anything in the past, then what I meant must be determined by nonintentional facts; and I have agreed, provided that the relation is one of necessary consquence. I grant that if I meant anything in the past, then what I meant must be a necessary consequence of nonintentional facts about me, my environment, my community, and so on. But it is not evident that there is a problem here, since none of the sceptic's arguments show that such a relation fails to hold. Indeed, they scarcely even attempt to show this (1997, p.230).

Contra this, note that in order for the fact that Jones is disposed to give the sum to metaphysically necessitate the fact that he means addition by "+", the conditions in which the disposition would be manifested have to be specified in a way such that their holding metaphysically guarantees that Jones responds with the sum. So they have to metaphysically guarantee that he does not mean any of the functions in the open-ended and potentially infinite set of functions that yield something other than the sum of the relevant input numbers. How could a non-semantically and non-intentionally characterised set of ideal conditions achieve *that*? Soames fails to show that there is any basis for confidence that this question can be answered in a way that favours the reductive dispositionalist. (An interesting question, which I can't go into here, is how the considerations broached in this note relate to Wittgenstein's remarks in PI §183).

[6] KW's sceptic also argues that dispositionalism fails to satisfy the second constraint and the idea that meaning is normative (1982, p.23, p.37). In addition to the works cited in footnote 3 above, see Sultanescu (2022) for a helpful discussion.

[7] A response which attempts to answer the sceptic by providing a meaning-constituting fact is what Kripke calls a "straight solution" (1982, p.69).

to avoid it by developing a "sceptical solution" in which Jones can with perfect propriety be described as e.g. meaning addition by "+" even given the conclusion that there is no fact capable of making this so (1982, p.69). We'll return to the sceptical solution in section 7 below. Our immediate concern, in the next section, is with the causal-historical picture of reference outlined by Kripke in NN.

3. The Causal-Historical Picture

In Lectures I and II of NN, Kripke launches a full-frontal attack on what he calls the Frege-Russell view of names. According to the Frege-Russell view a name has the reference it does because it is synonymous with, or an abbreviation of, a definite description. Kripke's famous modal argument suggests that the synonymy relation postulated by the Frege-Russell view implies that many clearly contingently true sentences have to be regarded – implausibly – as necessary truths. On Kripke's alternative picture, names are what he calls rigid designators: unlike definite descriptions they refer to the same individual no matter what counterfactual situation is under discussion. Moreover, in general we cannot even view definite descriptions as fixing the reference of these rigid designators: it is in general neither necessary nor sufficient for a use of a name to refer to a given individual that the user of the name associate it with a definite description uniquely satisfied by that individual.

As an alternative to the Frege-Russell view, Kripke proposes what has come to be known as the "causal-historical"[8] picture:

> Someone, let's say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say, Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman is a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely ... a chain of communication going back to Feynman himself has been established, by virtue of his membership in a community which passed the name on from link to link (1981, p.91)

At the origin of the chain, the relevant object or individual will be christened with the name in a baptismal ceremony. This may be via ostension, but it may also be via the use of a description. This does not constitute a concession to the Frege-Russell view, since, as John Burgess nicely explains "this description need not remain permanently associated with the name" (Burgess 2006, p.172). Having been baptised, the object or individual "continues to be denoted by that name even if the description used and every other circumstance of the baptism is forgotten or misremembered"(ibid.). In order to avoid fairly obvious counterexamples, moreover, for speakers at various points in the chain "the new

---

[8] Kripke himself dislikes the "causal-historical" label (comments in a CUNY seminar attended by the author in February 2022) and would prefer simply "historical chain" (see also Burgess 2006, pp.172-73) but the "causal" label has become entrenched in the literature (it is used e.g. by Hattiangadi, Kusch, McGinn, and Maddy) so I'm hoping I'll be forgiven for continuing to use it in this chapter.

user should intend to use the name for the same object the old user was using it for"(ibid.). As Kripke himself puts it:

> An initial "baptism" takes place. Here the object may be named by ostension, or the reference of the name may be fixed by a description. When the name is "passed from link to link" the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it (1981, p.96)

In Lecture III of NN, Kripke suggests that a similar picture be adopted for natural kind terms. As Burgess summarises the view:

> Using a description, perhaps involving demonstratives and requiring supplementation by ostension, that is true of them or at least that the baptist thinks is true of them, a natural kind of individual may be picked out and given a "common name". This common name or natural kind term thereafter passes from speaker to speaker, with the original description being perhaps very soon completely forgotten (2006, 181).

For example, if the baptist uses the description "the shiny, yellow, malleable metal in front of me" to fix the reference of "gold", the kind term will refer to the stuff with atomic number 79, and this reference will be passed along the causal-historical chain to future users of the term.

There have been some attempts in the literature to use the causal-historical picture of reference to respond to KW's sceptic's argument. I'll consider two such attempts, those developed by Colin McGinn (1984) and Penelope Maddy (1984). In the next two sections I'll argue that neither of these is successful.

4. McGinn

Colin McGinn (1984, pp.164-6) considers whether we might reply to KW by picking up on the suggestion in NN that a name *n* correctly applies to an object o if and only if o lies at the origin of the causal chain leading up to applications of *n* to o, and the parallel suggestion for natural kind terms, that e.g. "gold" applies to an object if and only if that object is of the same kind as the original sample which initiated the causal chain leading up to uses of "gold".[9]

---

[9] In this context, McGinn construes the causal theory as providing a set of necessary and sufficient conditions for the correct application of a name or natural kind term (1984, p.165), possibly going beyond Kripke's reticence about suggesting necessary and sufficient conditions for reference (1981, p.93). I'll let this point pass as I'm going to argue that even if we construe the causal theory as offering necessary and sufficient conditions it fails as a straight response to KW's sceptic (and if it is not offering necessary and sufficient conditions it's hard to see how it could even be a candidate for offering a straight solution). That said, there are some very interesting residual questions in this vicinity: could it be that in rejecting the search for necessary and sufficient conditions for reference here, Kripke is actually signalling that the causal-historical view in NN is *not* a form of reductive dispositionalist "straight solution", as in WRPL? This matter deserves a fuller discussion than I can attempt here. I will say, though, that I'm sceptical as to whether the reticence about necessary and sufficient conditions in NN signals a principled contrast like the contrast between straight and sceptical solutions in WRPL: whereas in WRPL there's a relatively clear contrast between truth-conditional accounts of meaning and assertibility-conditional accounts of meaning, there seems to be nothing comparable in NN to yield a point of contrast, in general terms, with accounts of a concept that eschew the search for necessary and sufficient conditions for its application. That said, I think that treating the attack on reductive dispositionalism in WRPL and the causal-historical view in NN as operating on

How might these suggestions secure determinate reference for e.g. "Kripke" and "tiger" in the face of KW's sceptic's challenge? Before answering this question it's worth noting McGinn's comment that Kripke fails to consider the causal theory of reference as a potential solution to the sceptical argument in WRPL because in the latter "he tends to formulate his sceptical problem in terms of the notion of meaning and not that of reference"(1984, p.166). In fact, McGinn is mistaken about this, since as we saw in section 2 above, there is no such separation in the sceptical argument. The sceptic assumes that meaning determines reference, so that whatever constitutes the fact that an expression has the meaning that it has must determine its reference.

Why, then, might we think that the relevant causal facts are incapable of securing the fact that the name "Kripke" refers to Kripke and the fact that "tiger" refers to tigers? Consider an application of the name "Kripke" to Hilary Putnam at some future time t*, and consider the sceptical suggestion that this application is correct, because by "Kripke" we actually mean *Kripnam*, where

x = Kripnam iff (a) it is time t prior to t* and x = Kripke or (b) it is time t later than or equal to t* and x = Putnam

All of the causal facts about the use of "Kripke" prior to t* are consistent with the name referring to Kripke and with the name referring to Kripnam, so they fail to determine the reference of the name. Likewise, consider an application of "tiger" to an aardvark at time t* and consider the sceptical suggestion that this application is correct because by "tiger" we actually mean *tigvark*, where

x is a tigvark iff (a) it is time t prior to t* and x is a tiger or (b) it is time t later than or equal to t* and x is an aardvark

Again, KW's sceptic will argue that all of the causal facts up to t* are consistent with the alternative hypothesis, so that ultimately they fail to determine the class of tigers as the things to which "tiger" is correctly applicable.

McGinn thinks that the causal theorist has a ready answer to these sceptical suggestions:

> [T]he non-standard extension Kripnam for "Kripke" … will not qualify for the simple reason that Putnam is causally isolated from my present use of "Kripke" (or we can suppose as much): the sceptical hypothesis was that "Kripke" correctly applies to Putnam after some future time t*, but the causal theory can exclude this possibility by observing that it is Kripke who lies at the origin of the causal chain leading up to my present use of "Kripke" – I need have had no causal contact with Putnam at all, still less the kind of causal contact that determines reference (1984, p.165).

Likewise:

> [M]y current use of "tiger" has a sample of tigers at its causal origin and not any aardvarks so that the sceptic is defeated if he claims that "tiger" might correctly apply to aardvarks after some future time t* (1984, p.165-6).

different levels is the best chance for a rational reconstruction of Kripke's views that would allow us to view the two texts as consistent: see §§7-8 below for an attempt along these lines and an argument that it is bound to fail. The issue of Kripke's "reticence" in NN would certainly bear further discussion (I'm very grateful to Olivia Sultanescu for raising this issue).

Neither of McGinn's suggestions convincingly overturns the sceptical argument, however: the sceptic can reply that it is irrelevant that Putnam isn't the causal origin of uses of "Kripke" since it is nonetheless true that *Kripnam* is at their causal origin. After all, given the sceptic's definition, the object at the causal origin of uses of the name is indeed Kripnam. Likewise, the objects at the causal origin of uses of "tiger" are, given the sceptic's definition, tigvarks.

How, then, can the causal theorist respond to the sceptical challenge? It seems that the causal theorist would need to cite some fact along the following lines: were I presented with Putnam at some time at t* or beyond I would not apply the name "Kripke" to him and were I presented with an aardvark at some time at or beyond t* I would not apply "tiger" to it. In other words, I'm not disposed to apply the name "Kripke" to Putnam at time t* or beyond, and I'm not disposed to apply "tiger" to aardvarks at time t* or beyond. Clearly, we are now back again replying on something like a dispositional theory, hence back again facing the problem afflicting dispositional theories rehearsed in section 2 above.[10]

Thus, McGinn's attempt to deploy the causal theory of reference against KW's sceptical argument fails to overturn it.[11]

5. Maddy

Penelope Maddy (1984) concentrates on the case of a natural kind term like "gold". She notes that the causal theorist tells a story according to which my current use of "gold" is linked by a historical chain of reference-preserving links that stretches back to an initial event in which the word is introduced in a baptism. As Maddy notes (1984, p.475-6 n.7) it could be argued that indeterminacy afflicts the historical chain, but like Maddy, we'll put this worry to one side and focus on the initial baptismal act:

> The baptist picks out samples of the metal in question. He points at these, pronounces "gold!", and from that moment, the word refers to whatever is like this, that is, to all members of the natural kind containing the samples. (1984, p.464)

Maddy correctly notes that the sceptic can question whether the baptismal act secures referential determinacy:

> If the extension of the term "gold" is to contain everything "like this", the referent of "this" must be determinate. Wittgenstein[12] argues that it is not. The baptist points

---

[10] As we'll see later, this is not at all surprising given that causal theories are in fact forms of dispositionalism.

[11] It appears that McGinn requires something along the lines of David Lewis's account on which certain sceptical possibilities for the referents and extensions of expressions can be ruled out on grounds of "unnaturalness" (Lewis 1983). This would be a major supplementation, requiring much work (and amongst other things a response to Kripke's remarks about the candidate response to the sceptic which invokes the relative simplicity of meaning hypotheses (1982, pp.38-40)). For some discussion and pointers to the relevant literature, see the postscript to Section 4 in Miller and Sultanescu (2022).

[12] Maddy puts it in terms of "Wittgenstein", but as she notes "My account of the rule-following problem derives from Kripke's extremely helpful book" (1984, p.475 n.2).

towards the sample, but who is to say whether he is pointing at the metal, or at its shape, or at its colour?" (ibid.)[13]

She argues, however, that the indeterminacy which threatens here can be closed down via an appeal to "neurological speculation" (p.465) and a focus on "a more structured perceptual connection" (p.464) between the baptist's perceptual states and features of his environment.

I'll outline a slightly simplified version of Maddy's argument. The reason that the baptist's ostensive act picks out the kind of metal (gold) exemplified by the sample rather than its shape (say, square) or colour (say, yellow) is that the baptist is *perceiving* the kind of metal exemplified by the sample rather than its shape or colour. How so? Maddy's idea turns on the thought that the perceptual state is underpinned by a kind of neural structure (Maddy calls these "cell assemblies") and that the content of the perceptual state is determined by facts about how the neural structure in question is correlated causally with features of the environment. The perceptual state caused by the sample of gold is underpinned by neural structure ψ (say). Square cardboard cut outs don't produce ψ, and neither do yellow flowers. However, gold triangles do cause ψ, as do gold bracelets. This set of complex causal facts makes it the case that in the initial baptismal event the baptist was perceptually responding to the nature of the sample qua metal, rather than to its shape or colour. Since the nature of the sample qua metal consists in its having atomic number 79, it is this that the baptist picks out when he pronounces "gold" in the initial baptismal ceremony.

We can see, then, that for Maddy the determinate reference of "gold", as it is used in the baptismal ceremony, is derivative on the content of the perceptual state that the baptist occupies when performing the baptism, and the content of the perceptual state is determined by facts about how it is causally correlated with features of the environment. This means, however, that Maddy's strategy faces problems very similar to those (described in section 2 above) that beset the reductive dispositionalist account of meaning. Contentful perceptual states can represent things as they are but they can also misrepresent them: I can have a perceptual experience with the content that such and such is gold when in fact the item causing me to have the experience is not in fact gold, and in the presence of a piece of gold I can fail to have an experience with that content. Thus, we cannot simply identify the content of a perceptual experience with the features of objects that cause me to have it, on pain of ascribing a disjunctive content that it doesn't possess and that would render it impossible for the perceptual experience to be a misrepresentation. It would seem that the content of the experience would need to be regarded as determined, if at all, by a certain select *subset* of the causal correlations it is capable of entering into. Characterizing this select subset in non-semantic and non-intentional terms is – for all that Maddy has said – of a piece with the problem faced by the reductive dispositionalist in characterizing a suitable set of ideal conditions in terms that do not take the notion of content for granted.

So, the causal theorist faces some severe problems in determining the reference of "this" in the "like this" of the baptismal ceremony, in other words, in

---

[13] This is what is known as the "qua problem" (Devitt and Sterelny 1993, chapter 4): is he pointing at the sample qua metal, qua shape, or qua colour?

determining the relevant feature of the sample that is the focus of the baptist's act.

There are further shortcomings in Maddy's story. Put to one side the problem outlined in the previous paragraph, that of determining the reference of "this" in the baptist's "like this". As Maddy notes:

> The next problem naturally concerns the "like". The extension of the term is to consist of whatever is like those samples. What determines this? (1984, p.469)

Recall from above the idea that the perceptual state caused by samples of gold is underpinned by neural structure ψ. Given that whatever else stimulates the neural structure ψ will be in some respect similar to the sample

> The temptation is to say that this similarity is the "like" in "like this", that the extension of the term "gold" consists of whatever stimulates [ψ] (Ibid.)

Tempting as it is, Maddy notes that this suggestion faces some serious problems. Since some pieces of gold (those too small or too far away, for instance) will fail to stimulate ψ while some pieces of Iron Pyrites will, we again face the problem of narrowing down the extension of "gold" by regarding it as containing a subset of those things that stimulate ψ and also some of the things that don't stimulate ψ. And Maddy makes it clear that she thinks no progress is to be made by invoking ceteris paribus conditions, given the success of Kripke's argument (against reductive dispositionalism) that it will not be possible to characterize these conditions in non-semantic and non-intentional terms (1984, pp.476-7, n.20).[14]

Maddy argues, however, that there is no need for the causal theorist to go down the route of seeking to identify suitable ceteris paribus conditions:

> The interpretation of the word "like" in "like this" was by no means left up for grabs; the whole idea is that by isolating a sample, the baptist fixes the reference of the term to members of its natural kind. It isn't up to the baptist to determine what belongs in the same kind as the sample; the world determines that. (1984, p.470)

Maddy argues that it is an objective, mind- and language-independent fact that the piece of gold in front of me and the piece of gold in a far distant part of the universe have the atomic number that they do, and that their doing so accounts for the phenomenal properties that regulate our use of "gold". And she notes that Wittgenstein can object that "the world is not pre-packaged into natural kinds independently of our linguistic activity"(1984, p.471), but this is to open a substantial can of worms so that "the debate [between the causal theorist and Wittgenstein] is once more at a standoff"(ibid.).

It seems to me, however, that the appeal to objective mind- and language-independent natural kinds will not help with the problem about determinacy. In response to the question "What makes it the case that the item in front of me and

---

[14] As noted above, Maddy fails to see that a problem of this sort is faced by her account of how the content of the baptist's perceptual states is determined. In addition to the problem concerning how the extension of "gold" is determined, Maddy notes (1984, p.477, n.22) that there is the further problem of seeing how the normativity of meaning can be accommodated by the causal theorist. She is optimistic about the causal theorist's prospects for dealing with this, but doesn't go into detail in the paper under consideration.

the item on Neptune belong to the extension of 'gold'?", the observation that they belong to the same kind will help only if it has been determined that the word "gold" refers to that kind. And this, as yet, has not been established. That is to say, Wittgenstein can concede that the world itself determines that the two samples belong to the same kind, but what he needn't concede is that the world itself determines that the word "gold" refers to that kind. Thus, Maddy's play with the notion of objective scientific kinds presupposes that the determinate reference of "gold" has been fixed – and we are still to be given an account of how that is so. It seems, then, that Maddy confuses two distinct claims:

> (i) For any natural kind that the baptist's sample belongs to, the world itself determines whether any other given object belongs to it

and

> (ii) The world itself determines what natural kind the expression "gold" refers to.

And as yet, we have no plausible story from Maddy as to how the causal theory of reference can deliver (ii).

Maddy's attempt to use the causal theory of reference to neutralise KW's sceptic's argument, like McGinn's, thus falls prey to the arguments developed by Kripke in Chapter 2 of WRPL.[15]

6. Dispositional Theories and the Causal-Historical Picture

If the causal-historical picture of reference adumbrated in NN provided resources to undermine the claim in WRPL that there is no "straight" solution to KW's sceptic's challenge, that would constitute a clear source of tension between the two texts. The arguments of the previous two sections suggest that this is not the case. But this does not mean that there is no such tension. Indeed, although more work would need to be done to drive the point home, prima facie at least it appears that KW's sceptic's argument against reductive dispositionalism undermines the "causal-historical" picture of reference sketched in NN.

I will now suggest further that NN's picture of reference and WRPL's stance on reductive dispositionalism are straightforwardly incompatible. To see the incompatbility, note Paul Boghossian's observation that "[I]n all essential respects, a causal theory of meaning is simply a species of a dispositional theory of meaning" (1989, p.164). Boghossian continues:

> The root form of a causal/informational theory is given by the following basic formula:

---

[15] In order to stop the gap in her account, Maddy could try appealing to the ideas of "reference magnetism" and natural properties suggested by the work of David Lewis (see note 10 above). Although some of the things Maddy says gesture in this direction (see esp. (1984, p.472)), the suggestion requires a development and defence that she fails to provide. And as we noted above in response to the thought that McGinn might deploy a Lewisian story at this point, this is far from being a straightforward matter. For congenial critical discussions that also conclude that the causal theory of reference is unable to neutralise KW's sceptic's argument, see Hattiangadi (2007, pp.141-144) and Kusch (2006, pp.133-136).

O means (property) P by predicate S iff (it is a counterfactual supporting generalization that) O is disposed to apply S to P (ibid.).

Although it is perhaps not immediately obvious, the causal-historical picture outlined in section 3 above fits this rubric. In the first instance, the view would be along the following lines in the case of a natural-kind term like "tiger":

(a) Speaker O refers to kind K by "tiger" iff K is the kind exemplified by the sample which lies at the origin of the causal chain leading up to O's current use of "tiger".

We saw in section 4 that in order to secure determinacy in the face of the sceptic's *tigvark* suggestion, (a) would require supplementation by something along the lines of:

(b) "Tiger", as used by speaker O, correctly applies to x iff O is disposed to apply "tiger" to x.

And the same thing goes in the case of names of individuals. Initially we have:

(c) Speaker O refers to individual o by "Kripke" iff the individual o lies at the origin of a suitable causal chain leading up to O's current use of "Kripke".

Additionally, in order to secure determinacy of reference in the face of the sceptic's *Kripnam* suggestion, a supplement along the following lines would be required:

(d) "Kripke", as used by speaker O, correctly applies to x iff (O is disposed to apply "Kripke" to x).

Given this, the incompatibility of NN and WRPL is clear: in advocating a causal-historical picture of reference, NN is effectively advocating a form of dispositional theory of meaning, whereas the rejection of dispositional theories of meaning is the centrepiece of chapter 2 of WRPL. To put it in the broadest terms, we can say that if (reductive) dispositional theories of meaning are a paradigm form of semantic naturalism, NN advocates that paradigm form of semantic naturalism while WRPL rejects it.[16]

---

[16] It's worth noting that as Boghossian points out (1989, p.164) causal-informational theories are only one form a dispositional view can take. Conceptual role theories would be another form of the same general dispositional view. This shows that McGinn is wrong in his conjecture (1984, p.166) that Kripke doesn't consider causal theories of reference in WRPL because of his extensive focus on the mathematical example of the "+" sign (mathematical objects, being abstract, are generally regarded as causally inert). Contra McGinn, Kripke *does* consider causal theories of reference albeit indirectly via considering dispositional theories in general, and this doesn't commit him to attributing causal powers to to abstract objects such as numbers and functions, since expressions referring to these get dealt with by conceptual role theories, another type of dispositional view.

7. A Possible Escape Route?

In Chapter 2 of WRPL, the causal-dispositionalist view of meaning is repudiated, while in NN it is defended. One way in which we might try to obviate the appearance of contradiction here would be to return to the distinction we mentioned briefly at the end of section 2, between straight and sceptical solutions to the argument of KW's sceptic. The sceptical solution attempts to undercut the sceptical argument by denying that meaning-constituting facts are necessary for the propriety of our meaning-ascribing practices. One way of developing the sceptical solution involves viewing it as an expressivist/quasi-realist account of semantic judgement.[17] On this way of looking at things, the claim that no straight solution is plausible would amount to the rejection of cognitivist views of semantic judgement on which they express beliefs about semantic states of affairs. To see how this might help dissolve the appearance of contradiction mentioned above, let's think about a superficially similar scenario in the ethical domain, and in particular how it might be possible to be a metaethical expressivist about moral judgement while holding e.g. a utilitarian view in normative ethics. I'll do this by taking Simon Blackburn's quasi-realism about moral judgement as our stalking horse.[18]

In the first instance, Blackburn sees himself as giving an *explanatory* story about the nature of moral judgement. A cognitivist explanation, for Blackburn, is an explanation that proceeds by attempting to identify distinctively moral states of affairs and then construing moral judgements as expressing beliefs that these states of affairs obtain. Cognitivism faces a crippling dilemma. Either we identify moral states of affairs with natural states of affairs or we construe them as non-natural and sui generis. If we take the former route we face the challenge posed by Moore's open-question argument: moral judgements appear to have a normative aspect, and/or an internal link to motivation, not possessed by beliefs about naturalistic states of affairs. On the other hand, if we take the latter route and attempt to construe moral judgements as expressing beliefs about non-natural and sui generis states of affairs, the account succumbs to the sorts of metaphysical and epistemological challenges faced by Moore and the intuitionists.[19]

Blackburn's way of avoiding this dilemma involves giving an explanation of moral judgement that does not help itself to the idea of a distinctively moral state of affairs. In mounting his alternative explanation, Blackburn adopts what might be called *methodological non-cognitivism*: in our explanation we can avail ourselves of natural states of affairs, beliefs about natural states of affairs and non-cognitive sentiments or attitudes directed at natural states of affairs – but not beliefs about distinctively moral states of affairs. We start out with the assumption e.g. that the judgement that X is wrong expresses the sentiment B!(X) and then attempt to *construct* a notion of moral truth: roughly, and as a first approximation, we could say that a moral judgement is true if it belongs to

---

[17] See Miller (2020) for a fuller elaboration (and also Miller (2010) and (2015) for a critique of other, non-expressivist construals of the sceptical solution).

[18] Blackburn (1984a), chapters 5 and 6.

[19] This is just an oversimplified thumbnail sketch to give the shape of Blackburn's position. The issues are in fact much more complicated (e.g. "Cornell realists" will argue that moral properties are both natural and sui generis: see Chapter 7 of Miller (2013)).

M*, the set of attitudes that would remain after all opportunities for improvement in attitudes, and dispositions to form them, have been taken (Blackburn 1984, p.198). This would allow us to view some moral judgements as genuinely true or false: and since we would have earned the right to do so on the basis of methodologically non-cognitivist materials, we would not have "sold out" to moral cognitivism in having done so. Blackburn's quasi-realist doesn't deny the existence of moral states of affairs (hence his realism), he simply refuses to take them for granted in the materials he deploys in his explanatory account (hence his *quasi*-realism).[20]

The multiple challenges faced by accounts of moral judgements along these lines are well known (see chapter 4 of Miller (2013) for an overview), but what is important for our current concern is that it seems possible to adopt it while holding a position in normative ethics that initially appears to align with naturalistic cognitivism. A utilitarian view, on which the standard of right action is maximising utility (say), might appear to be committed to a metaethical cognitivist view on which the judgement that X is right expresses the belief that X maximises utility. But there is no necessary connection between utilitarianism and cognitivism: it is possible to embrace utlitarianism from an expressivist perspective. To see how, think of the disposition to express the attitude of approval (H!) to actions insofar as they have the characteristic of maximising utility. Call this disposition, D. We could embrace utilitarianism by taking up an attitude of approval towards this disposition itself: H!(D). And having done so, we can think about whether this attitude would belong to the set of attitudes M* that would remain after all opportunities for improvement in attitude, and dispositions to form them, have been taken. This latter exercise is essentially what we do when we engage in normative ethics, and if it turns out that H!(D) in fact is a member of M*, we can conclude that it is *true* that actions are right if and only if they maximise utility. But we will not have left ourselves open to the open-question argument since our account involved no attempt to *explain* moral judgement by identifying moral rightness with maximising utility and using such an identity to characterise the content of moral beliefs. We can reject metaethical naturalistic cognitivism and the idea that moral judgements are to be explained in terms of naturalistic states of affairs while arguing for a utilitarian normative ethic.[21]

---

[20] As Blackburn puts it, we "distinguish where we start, as we attempt to give a theory of ethics, from where we end up" (1996, p.91). This is why I called Blackburn a *methodological* non-cognitivist rather than a non-cognitivist *simpliciter*: his non-cognitivism, such as it is, concerns the materials he allows himself to start out with, not where he ends up (where he'll happily embrace moral knowledge and moral truth). I'm grateful to Cameron Ogle for suggesting the "methodological non-cognitivism" label. For a paper that does nicely highlight this facet of Blackburn's view, see McDowell (1987). (Of course, McDowell goes on to criticise Blackburn on other grounds, but these are not our concern here).

[21] Note that the methodological non-cognitivist starting point in effect imposes a "no circularity" constraint on what the quasi-realist can use in his explanatory story, but it does not follow from this that in the attempt to construct of moral truth we have to start out from a point where all our moral judgements are collectively suspended. In that (justificatory) part of the project, we are "working from within", and not "playing the fake game of trying to certify values without values" (Blackburn 1996, p.89). (This is a point missed by McDowell in his otherwise exemplary (1987) exposition of Blackburn).

If we view the sceptical solution in Chapter 3 of WRPL as proposing a metasemantic expressivist account of semantic judgement, might it be possible to square this with acceptance of a causal-dispositional story at the level of first-order semantic theory?[22] Just as facts about the utility of actions might form the basis for the selection of right action even though moral judgements are not to be construed as expressing beliefs about the utility or otherwise of actions, could causal/dispositional facts form the basis for selecting the referents of linguistic expressions even though judgements about meaning are not to be construed as expressing beliefs about causal/dispositional states of affairs? If so, NN could perhaps be regarded as engaging in the first of these – arguing that causal/dispositional facts are an appropriate basis on which to assign referents to linguistic expressions – while WRPL could be regarded as taking a stance on the second – arguing that meaning-ascribing judgements should not be construed as expressing beliefs about causal/dipositional states of affairs.[23] Just as in normative ethics we don't simply aim to classify actions as right or wrong but attempt to delineate the general principles which govern this classificatory process, in philosophical semantics we won't just aim to assign referents to expressions but attempt to identify the general principles which govern this classificatory process:[24] so the idea would be that just as a metaethical expressivist can be a utilitarian in normative ethics, a purveyor of the sceptical solution might embrace a causal-historical/dispositional view in philosophical semantics.[25]

8. No Way Out

The suggestion mooted in section 7 strikes me as the best prospect for squaring the views of NN with the arguments developed in WRPL. In the end, however, the analogy between the moral case and the semantic case breaks down at a crucial

---

[22] Note that "metasemantic" here is to be distinguished from "metalinguistic" as used in the exposition of the sceptical scenario in §2 above. The expressivist position in semantics considered in the present section is a position in metasemantics in the same sense in which quasi-realism in the moral case is a position in metaethics.

[23] Blackburn is perhaps advocating something along these lines in his paper on KW (1984b,p. 37).

[24] So on the way of looking at things proposed here, practical ethics would correspond to what in the philosophy of language is called by Robert Stalnaker (2017, p.903) "descriptive semantics", normative ethics would correspond to Stalnaker's "foundational semantics", and metaethics would correspond to what I term "metasemantics". (I'm unsure whether Stalnaker would follow me in distinguishing between foundational semantics and metasemantics in a way that parallels the distinction between normative ethics and metaethics).

[25] The fact that the causal-historical picture explicitly involves intentions to preserve reference without raising any question about what constitutes the fact that *these* intentions have the content that they do suggests that Kripke's concerns in NN are of a different order from those in WRPL. (This is also suggested by his inclusion, among the facts that determine reference, of facts about what we think. See e.g. Kripke (1981, p.95)). The maneuver considered here would provide one way of capturing this difference: the causal-historical picture would be viewed as best systematizing our intuitive judgements about reference rather than providing a constitutive account of putatively reference-determining meaning facts. (Only the latter is governed by a no-circularity constraint, so the circularity objection to reductive dispositionalism outlined in section 3 doesn't rule out the proposed maneuver: the real circularity problem at this point in the story is the one outlined in the next section).

point, undermining the idea that the two texts can be viewed as compatible. The problem concerns the resources available for use by the expressivist explanation in the two cases. Recall that in the moral case, the expressivist is allowed: non-cognitive sentiments and attitudes, natural states of affairs, and beliefs about the obtaining of natural states of affairs. Without helping himself to a notion of moral truth the ethical quasi-realist seeks to construct a notion of moral truth out of the materials at his disposal by focussing on the notion of improvement in attitude and identifying the true moral judgements as those which express attitudes belonging to M*. As Blackburn summarises the proposal:

> [T]he root idea is that the virtue of truth is constructed from the virtues of method (Blackburn 1984a, p.237).

The crucial reflection here is that, in any given case, a *method* is something that can be deployed in ways that are *better* or *worse*. In other words, the notion of a method presupposes the notion of a *standard* for assessing attempts to apply it. Or to put it another way, the notion of a method presupposes the notion of a *rule* which sorts items into different evaluative categories. Now the notion of a rule is of a piece with the notion of meaning: just as a rule sorts behavioural episodes into those which comply with it and those which don't, the meaning of an expression does likewise with respect to uses of that expression. This is why Wittgenstein focusses on, precisely, *rules and language* in the passages which inspired WRPL, and indeed why the notions of *rule* and *language* feature in the book's title.

A consequence of this is that the expressivist about meaning, unlike the moral expressivist, is barred *by his own methodology* from taking for granted the notion of a standard and its attendant notions of compliance and non-compliance. The "root idea", then, that allowed the moral expressivist to argue his way to a utilitarian position in normative ethics is simply not available to the quasi-realist about meaning. While utilitarianism might conceivably emerge in good standing from the metaethical expressivist's construction of moral truth, the causal-historical/dispositional picture cannot emerge in similar fashion from KW's sceptical solution.

9. Conclusion

It seems, then, that he causal-historical picture adumbrated in NN presupposes a straight solution to KW's sceptical challenge, while WRPL denies that a straight solution is possible. In conclusion, though, we should note that pointing out this apparent inconsistency does not necessarily amount to a criticism of Kripke: after all, he himself doesn't explicitly *endorse* the arguments developed in WRPL. He writes that WRPL "should be thought of as expounding neither 'Wittgenstein's' argument nor 'Kripke's': rather Wittgenstein's argument as it struck Kripke, as it presented a problem for him" (1982, p.5). Even so, he clearly does find the arguments developed in WRPL to be challenging and important. The question of how they relate to the views outlined in NN is likewise an important and challenging matter for devotees of Kripke's two classic texts.

References

Blackburn, Simon (1984a) *Spreading the Word: Groundings in the Philosophy of Language.* Oxford: Clarendon Press.

Blackburn, Simon (1984b) The Individual Strikes Back, *Synthese*, 58 (3), 281–301. Reprinted in Miller & Wright 2002.

Blackburn, Simon (1996) Securing the Nots: Moral Epistemology for the Quasi-Realist in W. Sinnott-Armstrong and M. Timmons (eds.) *Moral Knowledge: New Readings in Moral Epistemology*. Oxford: Oxford University Press.

Boghossian, Paul (1989) The Rule-Following Considerations *Mind*, 98 (392), 507–49. Reprinted in Miller & Wright 2002.

Burgess, John (2006) Saul Kripke: *Naming and Necessity*, in John Shand (ed.) *Central Works of Philosophy volume 5 (The Twentieth Century: Quine and After)*. Chesham: Acumen Publishing.

Devitt, Michael and Kim Sterelny (1993) *Language and Reality: An Introduction to the Philosophy of Language*. Cambridge, Mass.: MIT Press.

Hattiangadi, Anandi (2007) *Oughts and Thoughts: Rule-Following and the Normativity of Content*. Oxford University Press.

Kripke, Saul (1981) *Naming and Necessity*. Oxford, Blackwell.

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

Kusch, Martin (2006) *A Sceptical Guide to Meaning and Rules*. McGill-Queen's University Press.

Lewis, David (1983) New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61, 343-377.

Maddy, Penelope (1984) How the Causal Theorist Follows a Rule, *Midwest Studies in Philosophy* IX, 457-477.

McDowell, John (1987) Projectivism and Truth in Ethics, reprinted in his *Mind, Value and Reality*. Cambridge, Mass: Harvard University Press 1998.

McGinn, Colin (1984) *Wittgenstein on Meaning*. Oxford: Blackwell.

Miller, Alexander (2010) Kripke's Wittgenstein, Factualism and Meaning, in D. Whiting (ed.)*The Later Wittgenstein on Language*. Basingstoke: Palgrave MacMillan.

Miller, Alexander (2013) *Contemporary Metaethics: An Introduction* Revised and Expanded 2nd edition. Cambridge: Polity Press.

Miller, Alexander (2015) Rule Following, Error Theory and Eliminativism, *International Journal of Philosophical Studies*, 23 (3), 323–36.

Miller, Alexander (2020) What Is the Sceptical Solution?, *Journal for the History of Analytical Philosophy*, 8 (2).

Miller, Alexander (2022) The Normativity of Meaning and Content", in P. Stalmaszczyk (ed.) *The Cambridge Handbook of the Philosophy of Language*. Cambridge: Cambridge University Press.

Miller, Alexander and Olivia Sultanescu (2022) Rule-Following and Intentionality, *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), Edward N. Zalta (ed.), URL =

https://plato.stanford.edu/archives/sum2022/entries/rule-following/.

Miller, Alexander, and Crispin Wright, eds (2002) *Rule-Following and Meaning*. Montreal & Kingston: McGill-Queen's University Press.

Soames, Scott (1997) Skepticism about Meaning: Indeterminacy, Normativity, and the Rule-Following Paradox, *Canadian Journal of Philosophy*, 27 (supplement), 211–49.

Stalnaker, Robert (2017) Reference and Necessity, in Bob Hale, Alexander Miller and Crispin Wright (eds.) *A Companion to the Philosophy of Language* Revised and Expanded 2nd Edition. Oxford: Wiley-Blackwell, 902-919.

Sultanescu, Olivia (2022) Meaning, Rationality, and Guidance, *Philosophical Quarterly* (advance online publication), (https://doi.org/10.1093/pq/pqac004)

Wittgenstein, Ludwig (1979) *Ludwig Wittgenstein and the Vienna Circle: Conversations recorded by Friedrich Waismann*. Edited by Brian McGuinness, translated by Joachim Schulte and Brian McGuinness. Oxford: Basil Blackwell.