Yet Another Victim of Kripkenstein's Monster: Dispositions, Meaning, and Privilege[*]
Andrea Guardo

Semantic dispositionalism is a very attractive option in metasemantics; it stems from quite an intuitive idea – namely that the meaning of a linguistic expression is a matter of how speakers are disposed to use it – and fits nicely the naturalistic inclinations shared by many of us. Several philosophers, however, find it deeply problematic, which has led to a rather large literature about its very viability. As large as it is, though, this literature usually focuses on quite a narrow set of challenges to the dispositional view, challenges first raised by Saul Kripke in his *Wittgenstein on Rules and Private Language* (1981) and therefore customarily attributed to "Kripkenstein".[1] The main goal of this paper is to broaden the playing field by calling attention to a strangely neglected argument against semantic dispositionalism, which I will refer to as "the Privilege Argument" (or just "Privilege", for short).

The Privilege Argument builds on some, rather sketchy, remarks by Paul Boghossian (1989, esp. § 21) and Anandi Hattiangadi (2007, pp. 106-108 and 117). The strength of such a line of argument has not been appreciated enough, maybe not at all, and people writing on dispositionalism routinely ignore it. My hope is to remedy this situation by (1) offering what, as far as I know, is the first systematic development of the intuitions behind Boghossian's and Hattiangadi's remarks and (2) trying to distinguish them as clearly as possible from the ideas lying at the root of other anti-dispositionalist arguments.

Clearly distinguishing Privilege from those other arguments is, I think, especially important, since I suspect that the main reason why people tend to ignore

[1] The portmanteau "Kripkenstein" originated as a way of talking about ideas that, in his essay, Kripke contentiously ascribes to Wittgenstein but stops short of fully endorsing. People do not want to say, of the ideas in question, that they are Wittgenstein's (because that would be controversial), nor do they want to say that they are Kripke's (since he does not fully endorse them); so they say that they are Kripkenstein's. This means that, strictly speaking, the standard arguments against semantic dispositionalism should not be said to be "Kripkenstein's", for Kripke clearly takes them to be sound and he nowhere attributes them to Wittgenstein. That being said, people routinely refer to these arguments as "Kripkenstein's", and I think there is no point in going against common usage, at least in this case.

the considerations the argument gives voice to is that they conflate them with other, more familiar, lines of reasoning. This is why, after having sketched, in section 1, the basics of the dispositional approach and briefly summarized the state of the debate about its viability, I will introduce Privilege against the background of two other, well-known, anti-dispositionalist arguments and then immediately take care to distinguish it from them – and from a third, much-debated, argument against dispositionalism. Section 3 then discusses three strategies dispositionalists might be tempted to embrace in an attempt to answer Privilege – and to avoid any vulgar suspense, let me anticipate that the conclusion of that discussion will not be favorable to the dispositionalist program.

For the sake of concreteness – but without, I hope, any loss of generality – I will often focus on Jared Warren's (2020) recent version of the dispositional strategy. Just as in Borges' short story – where, by looking at the Aleph, one can see the whole universe – the current state of the debate on dispositionalism can practically be seen through the lens of Warren's discussion. Warren displays admirable ingenuity in his attempts to deal with the standard anti-dispositionalist arguments, and yet his theory ends up failing – becoming, to use his own turn of phrase, another victim of Kripkenstein's monster – simply because it does not really address the problems raised by Privilege.[2]

1. Semantic Dispositionalism and the Standard Arguments against It

In the sense in which I use the term, there is more to dispositionalism than just the idea that the meaning of a linguistic expression is in some way a matter of how speakers are disposed to use it. Such a use of "semantic dispositionalism" is not especially idiosyncratic, but it is nonetheless worth explaining. "Semantic dispositionalism", in this sense, refers to a family of views Kripke (1981, pp. 22-37 and 110-112) criticizes in the context of his discussion of (what he takes to be) Wittgenstein's rule-following paradox. And such views do not share just the commitment to the very general idea that meaning is in some way a matter of dispositions; they also share two other important features: they argue for *an a priori reduction* of meaning to speakers' dispositions and they assume a certain view about the relation between the notion of meaning and that of *reference*. Let us start with the second point.

Semantic dispositionalism, in the Kripke-inspired sense in which I am using the term, relies on a certain picture of the concept of meaning (for which see, e.g., Chalmers 1996 and Kaplan 1989), what is sometimes called the "referential picture" of meaning. According to this picture, the meaning of a linguistic expression

---

[2] The Privilege Argument is not an argument of Kripkenstein's, but I take it to be Kripkensteinian in spirit (see below, note 19). So there is a sense in which, by failing to address it, Warren's account does end up being a victim of the monster.

is something like a rule explaining how the referent, the denotation, of that expression depends on certain parameters – in more formal terms, a function which takes possible referents as values.[3] The meaning of "I", for example, would tell us that its referent depends (roughly) on the identity of the speaker, and the meaning of "the president of the United States" would tell us how its referent depends on features of the conversation's topic (are we talking about the latest news or about the Cuban missile crisis? About this world or about some merely possible one?). Now, against the background of the referential picture, asking what makes it the case that a certain expression has the meaning it has is the same as asking what makes it the case that, given the value of the relevant parameters, a certain expression refers to what it refers. What determines the reference of, say, "red"? What makes it the case that "+" refers to addition, and not to some other function?[4] It is to such questions that the dispositionalist answers that it is all a matter of the speakers' dispositions.[5]

---

[3] Wittgenstein's (1921) view, on which Kripke focuses in his book, can be seen as a degenerate version of this picture (degenerate in a sense akin to that in which a triangle formed by three collinear points is said to be degenerate), an über-simplified variant in which the meaning function is always constant.

[4] Examples involving mathematical symbols are extremely popular in the literature. The main reason is, I think, that the referents of such symbols do not depend either on the identity of the speaker (unlike the case of "I"), or on the speaker's environment (unlike "water"), or on the conversation's topic (unlike "the president of the United States"), so that we can at least pretend that their meaning is just their referent – we can, for example, identify the meaning of "+" with a set of ordered triples. Note, however, that analogous moves can be made in the case of other expressions, too – if you take its reference to be the class of all the red things, "red" is topic-sensitive (its referent depends at least on whether we are talking about the actual world or not), but if you take it to be the class of all the determinate (but repeatable) shades of red, then there is no topic-sensitivity whatsoever.

[5] Brian Skyrms (1996, 2004, and 2010) has maintained that meaning is a matter of the speakers' embracing certain strategies (in the game-theoretical sense), which in turn seems to be a matter of dispositions. However, Skyrms explicitly denies that intentionality (i.e. reference) is required to give an adequate account of the informational content (i.e. of the meaning) of signals (2010, pp. 43-44). Therefore, his is not the kind of view I have in mind when I speak of "semantic dispositionalism". That being said, there is clearly a sense in which Skyrms is a dispositionalist – after all, in his view meaning is a matter of dispositions. So here is a question: how does such a non-referential dispositionalism fare against the Privilege Argument? The answer is that it depends on the details of the view! If a dispositionalist goes non-referential while trying to retain the notion that there is an objectively correct way to use words and symbols (or even just that there are objective facts about what means what), then I think it is pretty clear that Privilege can be run against their view, too. However, if the theory is developed in a more anti-realistic direc-

However, talk of dispositions *determining* reference, or *making it the case* that a certain sign denotes a certain entity, is somewhat vague (in the everyday, not the technical, sense of the word). Is the idea that facts about reference *supervene* on facts about the speakers' dispositions? Or are such claims better interpreted, say, in terms of grounding? Or maybe in terms of identity? And is dispositionalism a reductionist view or not? And if it is, what kind of reduction are we talking about? An a posteriori reduction, such as that of water to H_2O? Or is the dispositionalist making a conceptual analysis claim, thereby championing an a priori reduction of meaning and reference to speakers' dispositions?

Kripke is not very explicit about such issues, but it seems clear to me that the dispositionalists he has in mind must be arguing for some form of a priori reduction. In reading Kripke this way I agree with Boghossian (2015, p. 336); however, unlike Boghossian, I also believe that Kripke had excellent reasons to focus on conceptual analysis varieties of the dispositional view. Kripke begins his exposition of (what he takes to be) Wittgenstein's argument by casting doubt on *the very intelligibility* of *the notion* of meaning, referentially conceived, and so the challenge his book revolves around is primarily to make sense of that concept. And in order for it to be addressing this problem, dispositionalism must, of course, be put forward as an exercise in conceptual analysis.

Here, then, is the kind of view I have in mind when I speak of "semantic dispositionalism". You start with the background assumption that the concept of meaning must be made sense of in terms of reference, roughly along the lines explained above. Then you are confronted with the usual battery of Kripkensteinian arguments aimed at showing that the referential picture of meaning is of dubious coherence. And this leads you to try to make sense of the results of your referential analysis of meaning in terms of dispositions. Semantic dispositionalism, in the sense in which I am using the term, is therefore the general view that meaning, *referentially conceived*, can be *a priori reduced* to speakers' dispositions.

That the dispositionalists' goal is to provide a reductive account of the concept of meaning is especially worth stressing, for it entails that – on pain of their analysis turning out to be viciously circular – dispositionalists cannot make, in their account, any use of that concept. This is a key requirement that any attempt to answer Privilege – and, more in general, any attempt to defend dispositionalism – must meet, and which Kripke himself (1981, pp. 27-32) explicitly calls attention to. We will encounter it again in section 3, during our discussion of some of the ways in which a dispositionalist may try to meet the challenge raised by Privilege.

---

tion the argument is not a problem anymore – of course, the theory will have to deal with all the objections which are commonly raised against its ilk, but that is another story.

Having clarified what the dispositionalists' thesis amounts to, it is useful to say something about its prima facie plausibility. This comes, I believe, mainly from the fact that the dispositional strategy is suggested quite naturally by the standard story about past usage's inability to determine reference.[6]

Let us imagine that we have been using "karatakai" to talk about the color a certain fruit is when perfectly ripe – where the description should be read de re. We now bump into a hue ("H", for short) we had never seen before. Should H things be labeled as "karatakai", or not? Since H is a hue *we had never seen before*, it is clear that *past* usage, by itself, is not enough to answer such a question.

One could, however, try to tell a more complex story, of which past usage is but one component. One could, for example, try to argue that every time we *used* "karatakai" we also *grasped* a certain universal, and that that is enough to determine whether H things should be labeled as "karatakai" or not – if H things instantiate the universal in question, it is correct to call them "karatakai"; otherwise, not.[7]

But the most natural way to muscle up the intuition that the way a word should be used must have something to do with the way it has been used in the past is, of course, to go dispositional. Every time we used "karatakai", our utterance was the manifestation of just one of our countless dispositions concerning that word. Most of those dispositions have never been manifested, usually because we have never bumped into the relevant stimuli, but they were there nonetheless; and among them there were also dispositions concerning H, which determine whether H things should be labeled as "karatakai" or not – if these dispositions were dispositions to categorize H things as karatakai, then it is correct to call H things "karatakai"; otherwise, not.

As natural as it is, the idea that the reference of words and sentences is determined by the speakers' dispositions has, as I have already noted, many enemies. Following Kripke (1981), skepticism about the dispositional strategy is usually motivated by an appeal to some subset, proper or not, of a group of three arguments.

First, since we are all disposed to make (what we would usually regard as) mistakes, e.g. categorizing white objects as red under certain illumination conditions, a dispositional analysis assigns the wrong meanings to our words and can therefore be discarded out of hand as empirically inadequate (the Mistake Argument, or just Mistake).

---

[6] I myself believe that the standard story is false and that this can be used to build yet another, quite general and extremely powerful, anti-dispositionalist argument, but here we can bracket such issues.

[7] I discuss this strategy at greater length in Guardo 2018.

Second, we do not have enough dispositions, and so – once again – a dispositional analysis does not assign the correct meanings to our words (the Cardinality Argument, or just Cardinality). Take, for example, the case of "+". The dispositional analysis tells us that "+" means *plus* if and only if our dispositions concerning that sign track addition, but our dispositions cover only a finite segment of addition – since we lack any dispositions about additions involving huge numbers – and therefore the dispositional analysis tells us that "+" does not mean *plus*. But, of course, "+" does mean *plus* and, therefore, the dispositional analysis is empirically inadequate.[8, 9]

Third, meanings entail *categorical* oughts. If, for example, "+" means *plus*, then you ought to answer "125" if asked for "68 + 57" – and this ought does not depend on your desire to give the correct answer, nor on any other particular desire of yours.[10] But having certain dispositions does not entail any such ought. Therefore meaning cannot be analyzed in terms of dispositions (the Ought Argument, or just Ought).[11]

These are, of course, also the arguments on which the friends of dispositionalism focus when they have to defend the strategy. There is, therefore, something like a consensus that the destiny of dispositional analyses of meaning hinges on the status of these arguments (see, e.g., Båve 2020, Blackburn 1984, Cheng 2009 and 2010, Ginsborg 2011, Heil and Martin 1998, Jones and Podlaskowski 2012, Kowalenko 2009, Kusch 2005, Podlaskowski 2012, and

---

[8] Note that the problem is not that our dispositions are finite: the notion that our dispositions cover only a finite segment of addition is consistent with the idea that we have infinitely many dispositions concerning "+". This is why I prefer "Cardinality Argument" to the somewhat misleading "Finitude Argument".

[9] As stressed by several critics of semantic dispositionalism, the Cardinality Argument is perfectly general, applying also to non-mathematical cases; mathematical symbols, however, provide an especially vivid illustration of the problem.

[10] This is the standard way the first premise of this third argument is formulated (see, e.g., Boghossian 2003 and 2005, Gibbard 2012, Glüer and Wikforss 2009, Hattiangadi 2006 and 2007, Miller 2010, Whiting 2007 and 2009, and Wikforss 2001). Such formulations, however, may be a bit misleading, in that they may suggest that the relevant notion is that of a no-matter-what ought (you ought to answer "125" even if, say, you know that, if you do that, the Society for Postmodern Mathematics will break your legs), which, of course, is not the idea – the point is just that the oughts in question, *whatever their strength*, do not depend for their existence on the speakers' desires. It would be less misleading to talk about *reasons* which, though pro tanto, exist independently of the speakers' desires.

[11] Ought is usually called "Normativity Argument", since it is thought to capture Kripke's remarks about the normativity of meaning. This is controversial (see, e.g., Guardo 2014, Sultanescu and Verheggen 2019, and Zalabardo 1997), therefore I settled for a more neutral label.

Schlosser 2011). This is unfortunate, since there are other arguments which provide, I think, stronger reasons to be wary of this approach. The Privilege Argument is one of them.

2. Cardinality, Mistake, and Privilege

The best way to introduce the Privilege Argument is against the background of Cardinality and Mistake. This is because the Privilege Argument can be viewed as arising from a weakness in the standard strategy for dealing with Mistake, which is in turn related to Cardinality.

Let us start, then, by asking what a successful reply to Cardinality must look like. The answer is, I think, obvious. In order to successfully deal with it, a dispositionalist must show that even though it may seem that I have no disposition to give the arithmetically correct response "Z" if asked for the sum of two huge numbers X and Y, in fact I do have such dispositions.

Building on previous work by Blackburn (1984), Shogenji (1993), and others, Warren (2020, pp. 262-268) has recently offered a nice implementation of this strategy, and having taken a look at it here will come in handy later on. The basic idea behind Warren's approach is simple. For some pairs of (smallish) numbers, there is no doubt that I have the disposition to give the arithmetically correct response if queried about their sum. For some others, even though I do not have any *simple* disposition to give the correct answer, I have suitable *complex* dispositions: if asked for "1,091 + 2,906" I would answer "3,997" as a result of the execution of some intermediate steps (i.e. adding 1 to 6, 9 to 0, etc)[12] – even though I have no disposition to give that answer *directly*, without going through the various steps of the algorithm. When the addenda are really huge I have neither a simple nor a complex disposition to give the arithmetically correct response – even if I were to start going through the relevant intermediate steps, I would die of old age before getting halfway through;[13] however, in such cases I still have an appropriate *composite* disposition, where a subject S has a composite disposition to perform action A if and only if A would be the output of the manifestation of certain *linked* dispositions of S's.

The key claim is, of course, the last one, so let us take a closer look at it. Let $d_1 \ldots d_n$ and $g_1 \ldots g_m$ be, respectively, an n-digit and an m-digit number – where n and m are assumed to be huge and, for simplicity's sake, it is also assumed that n > m. One way of having a complex disposition to give the correct response to

---

[12] For simplicity's sake, I am ignoring the possibility of mimics or finkishly lacking dispositions (for mimicking see Johnston 1992, for the finkish lack of dispositions see Martin 1994).

[13] This time I am ignoring the possibility of masks (for which see Johnston 1992) or finks (for which see Martin 1994).

7

"$d_1...d_n + g_1...g_m$" would be to have the disposition to go through the following steps: first, I add d_n to g_m, I get the result t_1t_2 (where t_1 is either 1 or 0), and I write down "t_2"; second, I add d_{n-1} to g_{m-1}, I get the result t_3t_4, to which I add t_1, thereby getting t_3t_5 (of course, sometimes $t_4 = t_5$), and I write down "t_5" to the left of "t_2"; and so on – modifying the procedure in the obvious way once I get to d_{n-m}. Now, let us say that this is a disposition I do not have – because by the time I get to d_{n-p} (where p is assumed to be smaller than n but still huge) I will be long dead. My lacking this disposition is perfectly consistent with the notion that, for each digit d_i of $d_1...d_n$, I do have the disposition to implement the $n-i+1^{th}$ step of the algorithm above. And having these dispositions is just having a composite disposition to give the correct response to "$d_1...d_n + g_1...g_m$". Let us imagine, for example, that I am *really* slow at adding and that, because of this, I lack the complex disposition to correctly compute "714,231,723,055,439 + 231,135,230,349,825" – by the time I get to the fourteenth step of the algorithm above I will be six feet under. This does not mean that I do not have, say, the disposition to compute "1 + 3" and write down the result!

Warren's strategy to deal with Cardinality looks, I think, promising enough. Let us now turn to Mistake and ask what a successful reply to *this* problem would look like. I think it is clear that a solution to Mistake will have to be somewhat more complex than one to Cardinality. More precisely, I believe that a little reflection is enough to see that this time the advocates of the dispositional strategy must both (1) convince us that there is a kind of disposition k such that, for any pair of numbers, the answer I am disposed$_k$ to give if asked for their sum is the arithmetically correct one and (2) modify the original, unqualified, dispositional analysis so that the meaning of a word depends only on our dispositions$_k$.

Warren (2020, pp. 270-273) offers, once again, quite a neat concrete case to focus on. First, let us define the notion of a *general* disposition by saying that S has a general disposition to perform A in a class of situations γ if and only if in the overwhelming majority of the situations in γ S performs A (this is what Warren calls an "M-general" disposition). Second, let an answer to a given question be *stable* if and only if checking and rechecking tends to produce the very same answer over and over again. Third, let us define the *normal* conditions as those in which neither external nor internal factors are interfering with the functioning of our cognitive apparatus.[14] Finally, let us put these three notions together and define a disposition of kind gsn (a "disposition$_{gsn}$", for short) as follows: S has the disposition$_{gsn}$ to give a certain response to a certain addition problem if and only if S has the *general* disposition to *stably* give that response when queried about that

---

[14] Warren assumes that we can make sense of the notion of *the* function of something in etiological terms, which I agree with (see, e.g., Wright 1973).

8

problem in situations in which the conditions are *normal*. Now, Warren maintains that, for any pair of numbers, the answer I am disposed$_{gsn}$ to give if asked for their sum is the arithmetically correct one (see point (1) above) and, accordingly, he suggests that we should take "+" to mean *plus* if and only if our dispositions$_{gsn}$ concerning that sign track addition (see (2) above).

Keeping all this in mind, we can now introduce the Privilege Argument. As we have just seen, dispositionalists are committed to the idea that there is some kind of disposition k such that, for any pair of numbers X and Y, the answer we are disposed$_k$ to give if queried about "X + Y" is "Z" – which, here and in what follows, will always denote the sum of X and Y; however, dispositionalists do not deny that for some other kind of disposition j and some pair of numbers X and Y, the answer we are disposed$_j$ to give if queried about "X + Y" is *not* "Z" – we are disposed$_j$ to make a mistake.[15] In other words, even though they believe that, for some k, our dispositions$_k$ track addition, dispositionalists are willing to grant that there are js such that our dispositions$_j$ do *not* track that function – if one is drunk, or overly tired, they will make some mistakes. Dispositionalists, however, are keen to assure us that this does not really matter, since the dispositions on which the meaning of "+" depends are only our dispositions$_k$; dispositions$_j$ are entirely irrelevant. But why so? Why are our dispositions$_k$ privileged over all our other dispositions? Why does a word's referring to a certain entity reduce to our being disposed$_k$, rather than disposed$_j$, to use it in a certain way?[16] Dispositionalists do

---

[15] In principle, one could try to deny this by arguing that mimics and/or finkishly lacking dispositions are ubiquitous. This, however, looks like an extremely ad hoc, and in fact also pretty desperate, kind of move.

[16] Note that the challenge is not to explain why certain, as a matter of fact addition-tracking, dispositions we actually possess are privileged over certain other, as a matter of fact non-addition-tracking, dispositions we could but do not actually possess. *That* challenge would not be that hard to meet – indeed, I take it to be obvious that if meaning facts are to be identified with dispositional facts, the relevant dispositional facts must be facts about the dispositions we have in the *actual* world (see Guardo 2012, pp. 203-204). But that is not Privilege's challenge to the dispositionalist! The disposition to answer "5" if I were asked for "68 + 57" *while under the influence of LSD* – which, as far as I know, I may actually possess – is not to be confused with the disposition to answer "5" if I were asked for "68 + 57" *which I would acquire were I to take a generous dose of LSD* – a disposition which in the actual world, trust me, I do not possess. In the former case, the mention of LSD is part of the specification of the relevant stimulus; in the latter case, it is not. And it is the first kind of disposition we are interested in here. Privilege's challenge to the dispositionalist is to explain why certain, as a matter of fact addition-tracking, dispositions are privileged over certain other, as a matter of fact non-addition-tracking, dispositions – which we all possess, the latter no less than the former, in the actual world (just as

not tell us – and, moreover, it is not clear what they could say. And if dispositionalists cannot answer such questions, the dispositional strategy does not really explain what it was supposed to explain, and it should therefore be abandoned. This is, in a nutshell, the Privilege Argument.[17]

That Privilege raises quite a serious problem for any dispositional view of reference becomes especially clear if we pay due attention to the fact that facts about reference are normative in nature: if "rosso" *refers* to the class of all the red things, then it is *correct* to call "rosse" the red things but not the non-red ones. If we keep this well in mind, the questions dispositionalists find themselves facing take a different form: why should the answer we are disposed$_k$ to give also be the answer we *ought* to give? And where does the alleged *normative oomph* of these dispositions come from? And as soon as we turn our attention to such questions, the dispositionalist's inability to provide satisfactory answers becomes, I believe, glaringly apparent.

It is important, though, not to misunderstand *the nature* of the relevant normativity. Unlike the oughts around which the Ought Argument revolves, the oughts (and the normative oomph) I have in mind here are unremarkably *hypothetical*. The relevant notion of the answer I ought to give is just the notion of the answer I ought to give *if I care about mathematical truth* – and, more in general, the relevant notion of the way one ought to use a word is just the notion of the way one ought to use that word *if they want to tell the truth*, or *if they want to communicate*. This means that what is usually referred to as "the debate on the normativity of meaning" (see, e.g., Boghossian 2003 and 2005, Glüer and Wikforss 2009, Hattiangadi 2006 and 2007, Whiting 2007 and 2009, and Wikforss 2001) is utterly orthogonal to my argument. The notion of normativity at issue in that debate is a

_____

in the actual world I am both disposed to say "Yes" if you ask me for a favor politely and disposed to say "No" if you ask me for that very same favor, but rudely). *This* challenge cannot be met just by insisting that if meaning facts are to be identified with dispositional facts, the relevant dispositional facts must be facts about the dispositions we have in the *actual* world – for the simple reason that this time all the dispositions involved are dispositions we *actually* possess.

[17] This is not the Privileging Problem of Bird and Handfield 2008. Bird and Handfield focus on pairs composed of (1) a disposition to give the arithmetically correct answer if queried about a given addition problem *involving huge numbers* and (2) a disposition whose manifestation prevents that of (1), e.g. a disposition to wither and decay within a certain number of years. On the other hand, the pairs I focus on are composed of (1*) a disposition to give the arithmetically correct answer if queried about a given addition problem (not necessarily involving huge numbers) and (2*) a disposition to give an arithmetically incorrect answer if queried about that same problem. (2) is a fink, (2*) is not, so the structure of the two arguments is quite different – even though they admittedly rely on somewhat analogous intuitions.

very strong one: the oughts which people like Daniel Whiting have argued follow from semantic facts are *categorical* (they do *not* depend on *any* desire of ours). None of the participants to the debate on the normativity of meaning regard the idea that there are merely hypothetical semantic oughts as particularly problematic (see, e.g., Hattiangadi 2006, esp. § 1 and 2007, esp. what she says about normativity vs norm-relativity). Even more importantly, that there are such oughts is something that semantic dispositionalists themselves accept. Indeed, that there are such oughts is a straightforward consequence of one of their key theses, namely of the thesis that meaning can be made sense of in terms of reference (see again § 1 of Hattiangadi 2006, but also Boghossian 1989, pp. 513 and 530). If "rosso" *refers* to the class of all the red things, then it is *correct* to call "rosse" the red things but not the non-red ones, which in turn straightforwardly entails that *if* one wants to use that word correctly they *ought* to call something "rosso" only if it is red.

It is also important not to misunderstand *the role* that considerations about normativity play in the challenge raised by Privilege. The challenge is to explain why the fact that a certain word has the meaning it has reduces to facts about *a certain kind* of disposition – what is so special about *those* dispositions? Hence, the challenge can be formulated without any mention of oughts and the like, and in principle it could be issued to any view which identifies meaning facts with facts about *a certain kind* of disposition – even to views which reject the notion that meaning is in any sense normative. Now, since, as a matter of fact, dispositionalism (in the sense of the term defined in section 1) assumes meaning facts to entail certain (hypothetical) oughts, the challenge can also be phrased in normative terms. And phrasing the challenge this way no doubt makes it especially vivid, which is why I think not only that it is useful to give this formulation, too, but also that sometimes it is more effective to focus on this dimension of the argument. But that should not obscure the fact that the challenge is primarily that of explaining why meaning has to do with that particular kind of disposition – no mention of oughts needed. So there is definitely a sense in which the Privilege Argument can be described as a normativity argument (provided, of course, that one is clear about the kind of normativity involved),[18] but by so doing one would be focusing on a somewhat superficial feature of the problem.

The challenge raised by Privilege – this, too, is definitely worth stressing – is also quite different from the one raised by Cardinality and Mistake.[19] As a matter

---

[18] As noted in the introduction, there are some remarks of Boghossian's that clearly revolve around the very same idea Privilege is a development of, and there is no doubt that Boghossian thinks about this issue in terms of normativity.

[19] That being said, there is no doubt that Privilege fits the general argumentative strategy of Kripke's book quite well, especially its emphasis on a-rule-for-interpreting-a-rule ar-

of fact, dispositionalists do not want their analysis to entail any revisionist corollaries about the referents of the words of our language; they do not want to find themselves committed to the notion that, say, "+" refers to something other than the addition function. The point of both Cardinality and Mistake is that such corollaries do follow from the dispositional view, either because we do not have enough dispositions or because we are all disposed to make (what we would usually regard as) mistakes. This is not the point of Privilege. The point of Privilege is not that dispositionalism sometimes assigns the intuitively wrong referents to the words of our language but, rather, that it cannot justify the assumption – which it *has to* make – that certain dispositions have a normative oomph that all the others lack and, more in general, that only those dispositions are relevant to the determination of a word's meaning and reference.

One way to appreciate the difference between, on the one hand, Privilege and, on the other, Cardinality and Mistake is to focus on the fact that answering Cardinality and Mistake does not, in and by itself, give us an answer to Privilege. Let us assume, for example, that Warren's answers to Cardinality and Mistake are successful. That would mean that if, first, we take care to analyze the meaning of "+" in terms of the relevant dispositions$_{gsn}$ (i.e. our general dispositions to stably give certain responses when queried about "+"-problems in situations in which the conditions are normal) and, second, we keep in mind the need to quantify over composite dispositions too, our analysis will deliver the desired result about the reference of "+" (i.e. that "+" refers to the addition function). That, however, is not yet having shown that the correct way to use a word is the way we are disposed$_{gsn}$ to use it.

The point is that Warren has proved, at most, that *if* the correct way to use a word is the way we are disposed$_{gsn}$ to use it, then it is not the case that we are hopelessly wrong about the correct way to use most (maybe all) of the words of our language – let me stress it: a *conditional* statement. If we agree that any analysis that entailed that we *are* hopelessly wrong about the correct way to use most words is a non-starter, that is a fine result; Warren has shown that the hypothesis that the correct way to use a word is the way we are disposed$_{gsn}$ to use it is not, unlike other dispositional analyses of meaning and reference, clearly absurd. But Warren has not proved the *antecedent* of the conditional, i.e. that the correct way to use a word is indeed the way we are disposed$_{gsn}$ to use it.[20]

---

guments – which, however, Kripke does not explicitly employ in his attack on semantic dispositionalism.

[20] In the fourth section of his paper (starting on p. 276) Warren discusses a passage in which Boghossian gives voice to some of the intuitions behind Privilege (Boghossian 1989, p. 513). His initial reading of the passage does not, I think, do justice to Boghossian's remarks and his first answer (pp. 277-278) looks like a version of the one

Here is another way to see that the point I am calling attention to has little to do with the general line of argument instantiated by Cardinality and Mistake. As I have already noticed, both Cardinality and Mistake rely on the assumption that any analysis that entailed that we are hopelessly wrong about the correct way to use most words (i.e. any analysis that assigns the intuitively wrong referents to most words) is a non-starter. Let us now try to drop this assumption and see what happens. Of course, neither Cardinality nor Mistake has any force anymore: if revisionism is not a flaw, entailing that "+" does not refer to addition cannot be a mortal sin. Not so with Privilege, though. Whatever dispositions the dispositionalist ends up identifying as the meaning-determining ones, one can still ask why those and not others.[21] Why is the answer we are disposed to give in these conditions the right one? Why not the answer we are disposed to give in those other conditions? And so on.

One can think of Privilege as having three steps. The first step consists in realizing that dispositionalists like Warren are making the above assumption, i.e. that the only meaning-determining dispositions are our dispositions$_k$. Second, one has to recognize that that assumption is in need of justification. And, finally, we end up realizing that no such justification is forthcoming. Since dispositionalists usually make the assumption in question tacitly, the first step is not trivial, but by now it should be clear that, indeed, dispositionalists do make this assumption. In the remainder of this section I want to say something more on the argument's second step, before finally turning to the third one in the next section.

The reason why I believe it is useful to say something more on Privilege's second step is that one may feel that when we ask the dispositionalist to justify the assumption that the only meaning-determining dispositions are our dispositions$_k$ we are making a demand that we would never issue to the proponents of other reductive theories – which, in absence of a convincing explanation of the asymmetry, would of course give us very strong reasons to suspect that here there is really no genuine explanatory burden for the dispositionalist.

---

discussed below, in section 3.1. He then goes on to give a reading of the passage seemingly closer to what I and (I believe) Boghossian have in mind (pp. 278-279), but some of the things he says suggest that he still does not fully appreciate the strength of Boghossian's remarks – for one, he seems to think that, assuming his answers to Cardinality and Mistake to be successful, he can take for granted that dispositions to use "+" explain the meaning of "+" (first paragraph of p. 280), which is precisely what Privilege denies (see again section 3.1 below).

[21] Actually, one can even ask why *dispositions*, and not something else – which shows that (as noticed in note 19 above) Privilege is, in a certain sense, really just another a-rule-for-interpreting-a-rule argument.

My answer to this, very natural, worry is that, contrary appearances notwithstanding, there is no asymmetry here. The explanation we are asking of the dispositionalist is no different in kind from the ones that we routinely expect from a reductive theory. Hence, the point of Privilege is just that dispositionalists do not seem to be able to meet the explanatory standard that successful reductive theories of other phenomena effortlessly meet.[22]

Take, for example, the case of memory. The common wisdom in neural science is that facts concerning the storage of general knowledge about the world, say that elephants have trunks, should be identified with certain neocortical facts (see, e.g., Schacter and Wagner 2013, pp. 1449-1451).[23] Now, when one makes such a claim, they must be able to justify it: what is so special about (the relevant areas of) the neocortex? And that is indeed what any good neural science textbook tries to do. Of course, ideally, what we would want is some kind of causal (see Lewis 1986) and/or mechanistic (see Machamer, Darden, and Craver 2000) story showing exactly how the manifestation of the knowledge in question is linked to the relevant neocortical facts – which is not really what we find in a neural science textbook. However, we are at least given, first, a general idea of how such a story might go and, second, some reasons (having to do with the observation of patients with brain lesions and/or with the use of neuroimaging techniques) to believe that here there indeed is a story to be told.

In a parallel fashion, if dispositionalists want to identify, say, "+"'s referring to addition with certain dispositional facts, they have to show us how our having the dispositions in question explains the properties of the phenomenon they are trying to reduce/the facts that constitute "+"'s referring to addition. Of course, there are also important differences between the two cases. For one, facts about what refers to what are normative in nature, in the sense explained above (if "rosso" *refers* to the class of all the red things, then it is *correct* to call "rosse" the red things but

---

[22] In this paper I am trying to avoid committing to any positive view about the nature of meaning; my thesis is purely negative: dispositionalism, in the sense defined in section 1, does not work. However, I feel that at this particular juncture it may be of some help to point out that I believe that the reason why dispositionalists cannot meet the explanatory demands their view commits them to is that the phenomenon they are trying to reductively explain (meaning referentially conceived and, more in general, meaning realistically conceived) is not a real phenomenon. That being said, I hope it is clear that the Privilege Argument can be accepted also by, say, a primitivist – the only difference between me and the primitivist will concern the ultimate cause of the dispositionalist's explanatory failures.

[23] I focus on the storage – rather than, say, the encoding or the retrieval – of information to avoid getting bogged down in issues concerning the (ir)reducibility of phenomenal consciousness.

not the non-red ones), while mnestic facts are not[24] – which suggests that the required explanation would have to be somewhat sui generis. Still, the general dialectic is the same.

3. Three Strategies to Try to Handle Privilege

One may agree that Privilege calls attention to a perfectly legitimate explanatory demand and still be skeptical that the argument raises any serious problem for dispositionalism – simply because they are confident that the dispositional view has the resources to meet the challenge. In this section I will argue that there are strong reasons to believe that such trust in the dispositionalist program is misplaced.

I shall discuss three strategies a dispositionalist can try to develop to meet the explanatory demands introduced in section 2, starting with the least and concluding with what I take to be the most promising: the answer I am disposed$_k$ to give

---

[24] Commenting on a previous (and much less explicit) version of this section, a reviewer for this journal suggested that Warren may want to start by explaining semantic facts in terms of dispositions and only then explain the relevant normative facts in terms of the semantic facts. Just as we start by explaining thermal facts in terms of molecular motion and then we explain things like the fact that you ought not to put your hand in the boiling water in terms of thermal facts (it's scalding hot!) together with various background assumptions concerning the effects of heat on our bodies and our attitude towards these effects, it seems that Warren can start by explaining the fact that "+" refers to addition in terms of dispositions and only then go on to explain the fact that one ought to answer "12" to "7 + 5" in terms of that semantic fact together with various background assumptions concerning addition (namely that it takes the value 12 for the arguments 7 and 5) and our respect for mathematical truth. Now, to a certain extent, this is, of course, perfectly fine: you can definitely explain the fact that one *ought* to answer "12" to "7 + 5" in terms of the fact that "+" refers to addition together with the relevant background assumptions. The problem is that, unlike thermal (or mnestic) facts, facts about what refers to what *are themselves normative in nature*, in the sense that a word's referring to, say, a certain class of entities *consists* in the fact that it is *correct* to apply that word to those, and only those, entities – it is not just that if you told me that even though "rosso" refers to the red things it is not correct to call "rosse" the red things (or that even though it is correct to apply it to the red things "rosso" does not refer to the red things) I would have no idea of what you are trying to say; the point is that "rosso" seems to refer to the red things *because* it is correct to call "rosse" the red things but not the non-red ones (for the methodological background see, e.g., Wasserman 2018, pp. 15-18). Therefore, there are normative facts concerning the use of words for which the strategy above cannot be used. One cannot start by explaining facts about reference in terms of dispositions and only then go on to explain the facts concerning correct use in terms of facts about reference – for facts about reference *are* facts concerning which applications of a word are, and which are not, correct.

is privileged because it is the *right* answer, the answer I am disposed$_k$ to give is privileged because it is the answer I would give in *ideal* conditions, and the answer I am disposed$_k$ to give is privileged because it is the answer I would give in *standard* conditions.

The lessons I will draw from my discussion are quite easy to generalize to other cases. That being said, the primary goal of this section (and, indeed, of the paper as a whole) is just to show that the challenge raised by Privilege should be taken seriously and that, by ignoring it, dispositionalists are dodging an issue that, on the contrary, they should regard as quite pressing. Hence, here I am not committing myself to the thesis that my arguments are generalizable to *any* attempt to answer Privilege dispositionalists may come up with – even though I suspect they are.[25]

3.1. Right
A strategy that, I suspect, some may find rather natural is to say that dispositions$_k$ are privileged because the answer I am disposed$_k$ to give is the *right* one. A little reflection, however, is enough to see that this will not do. The answer I am disposed$_k$ to give if queried about "X + Y" is "Z", i.e. the sum of X and Y. But answering "X + Y" with their sum is giving the correct answer only if "+" means *plus*. And "+" means *plus* only if dispositions$_k$ are privileged. Therefore, saying that dispositions$_k$ are privileged because the answer I am disposed$_k$ to give is the right one is not that much better than saying that dispositions$_k$ are privileged because they are privileged. Quite Münchhausen, but not that good, as a reply to Privilege.

This is a key point, and it is worth dissecting. First, let me stress again that dispositionalism, at least as it is commonly conceived, relies on a particular picture of the concept of meaning, namely what I have called "the referential picture". Against such a background, saying that the meaning of a word is a matter of how we are disposed$_k$ to use it is saying that what makes it the case that, given the value of the relevant parameters, that word *refers* to what it *refers* is that we are disposed$_k$ to use it in a certain way.

---

[25] It is also worth emphasizing, once again, that in the sense in which I use the term there is more to dispositionalism than just the idea that meaning is a matter of dispositions. Therefore, there are views which, albeit *in a broad sense* dispositionalist, simply fall outside of the scope of the Privilege Argument. In note 5 I mentioned Skyrms' theory, which likely falls outside of the scope of the argument because (arguably) anti-realistic. But another possible example may be Ginsborg's (2011) view, which would fall outside of the scope of the argument because non-reductionist.

Now note that what the dispositionalist is trying to do is *reduce* reference, and hence meaning, to dispositions$_k$, and to be regarded as successful a reduction must make sense of the various properties of the reduced phenomenon – the reduction of water to H_2O is an instance of a successful reduction only because it explains water's various properties, i.e. that (unlike most liquids) it becomes less dense as it freezes, that (at a pressure of 1 *atm*) it freezes at 0 °C, etc.[26]

Finally, do not forget that a peculiar feature of the phenomenon of reference is its normativity. If "red" refers to these shades, but not to those, then it is *correct* to apply it to these shades, but not to those. And if "+" refers to addition, then one *should* answer "125" when asked about "68 + 57" (at least as long as one cares about mathematical truth).

Now, if instead of focusing on the general question that lies at the heart of Privilege (why does a word's referring to a certain entity reduce to our being disposed$_k$, rather than disposed$_j$, to use it in a certain way?) we focus for a moment on the normative side of the challenge (why is the answer we are disposed$_k$ to give also the answer we ought to give?), we can say that Privilege holds that if we try to identify semantic facts with dispositional facts, thereby reducing the former to the latter, we cannot make sense of the relevant normative oomph – and so the attempted reduction ends up failing. The fact that we are *disposed$_k$* to answer "125" when asked about "68 + 57" does not entail that that is what we *ought* to do. And this becomes especially clear if we ask ourselves why the correct answer to "68 + 57" should be the one we are disposed$_k$ to give rather than the one we are *disposed$_j$* to give. What is so special about dispositions$_k$ and why should they be privileged over all our other dispositions?

It is at this juncture that we encounter the idea introduced at the start of this section: dispositions$_k$ are privileged because the answer we are disposed$_k$ to give is the right one. Now, as I have already noted, one may find such a suggestion rather natural, but it is question-begging. When we ask why dispositions$_k$ should be privileged over dispositions$_j$ part of what we are asking is why *the right answer* to

---

[26] The analogy with the case of water is a useful one, but (just like the analogy with the case of memory) it must be handled with care. For one thing, that there is such a thing as water and that water has certain properties nobody doubts, whereas that there is such a thing as reference, that "+" refers to the addition function and therefore the right answer to "68 + 57" is "125", even that *there is* an objectively correct answer to a "+"-problem and – more in general – an objectively correct way to use the words of our language, these are all things we *cannot* take for granted: Kripke's Wittgensteinian arguments show that the very notions of reference and of the objectively correct way to use a word are at least problematic. Identifying facts about reference with facts about the speakers' dispositions would be a way to show that there is indeed such a thing as reference. Nothing like that is needed in the case of water.

"68 + 57" should be *the one we are disposed$_k$ to give* rather than the one we are disposed$_j$ to give. Answering this question by saying that *the answer we are disposed$_k$ to give* is *the right answer* is clearly, indeed blatantly, question-begging. It is, in a certain sense, no answer at all.[27]

But even leaving aside the fact that part of what we are asking when we ask why dispositions$_k$ should be privileged is why the *right* answer should be the one we are disposed$_k$ to give, it should be reasonably clear that the strategy we are considering is question-begging. The answer we are disposed$_k$ to give when asked about "X + Y" is an utterance of the numeral denoting the sum of X and Y. But answering "X + Y" with the numeral denoting the sum of X and Y is giving the right answer only if "+" means plus. And "+" means plus only if dispositions$_k$ should be privileged. Therefore, justifying the privileging of dispositions$_k$ by saying that the answer we are disposed$_k$ to give is the right answer already assumes that the privileging of dispositions$_k$ is justified, which means that the justification is circular and therefore useless.

3.2. Ideal

A less naïve suggestion is that dispositions$_k$ – i.e. that particular kind of disposition such that, for any X and Y, the answer I am disposed$_k$ to give if queried about "X + Y" is "Z" (which, as always, denotes the sum of X and Y) – are, as a matter of fact, dispositions in *ideal* conditions[28] and this can provide the required normative oomph. Is this a viable option? Three different sets of considerations suggest that the answer is no.

First, a little reflection is enough to see that it is extremely hard to find a set of dispositions in ideal conditions which meets the following two conditions: (1) the

---

[27] There is, of course, a sense in which the dispositionalist can just *assume* that the right answer to "68 + 57" is "125": they can take that to be one of the facts that their reductive account of reference should explain – just as we took water's becoming less dense as it freezes to be one of the facts that its identification with H_2O had to explain. But the dispositionalist's reductive account of reference must indeed *explain* this fact. Hence, that the right answer to "68 + 57" is "125" cannot be one of the account's *assumptions*. It must be, rather, one of its *observational consequences*.

[28] Locutions such as "the answer I am disposed to give in ideal conditions" suffer from a scope ambiguity, so let me make clear that when I speak of dispositions in ideal conditions what I mean are things like my actual disposition to answer "125" if asked for "68 + 57" in ideal conditions, not things like the disposition to answer "125" if asked for "68 + 57" which I would have if I were in ideal conditions.

elements of the set track addition and (2) they are what I will call "actuality-grounded dispositions".[29]

The rationale behind (1) is obvious. However, (2) needs some justifying – and, of course, some explaining, too. Let us start with my use of "actuality-grounded dispositions".

Let β be an ordinary Bohemian glass. β is, of course, somewhat fragile; it has, in other words, the disposition to break if struck. And saying that β has this disposition is saying that it has the property of having the kind of microphysical structure which would be causally responsible for its breaking if it were struck.[30] Now take δ. δ is a diamond cup and, as such, it is far from fragile. However, δ has the "disposition" to break if struck *after its microphysical structure has been changed into that of a Bohemian glass*. But saying that δ has this "disposition" is not saying anything about the properties it has (in the actual world), since the property of having the kind of microphysical structure which would be causally responsible for its breaking if it were struck is not one δ currently has. Now, β's disposition to break if struck is an actuality-grounded disposition, whereas δ's "disposition" to break if struck after its microphysical structure has been changed into that of a Bohemian glass is not.[31]

Now, when dispositionalists say that what determines the meaning of a word are (past usage and) the speakers' dispositions, the idea is that a word has the meaning it has because of what goes on inside the speakers' skulls; at least in principle, the meanings of the words of our language could be read off of our brains (plus information about past usage).[32] But this is clearly incompatible with the meaning-determining dispositions being non-actuality-grounded "dispositions", for such "dispositions" have little to do with what *actually* goes on – as opposed to what *could* go on – inside our skulls. This is the rationale behind condition (2).

Here is another way to see why (2) is a perfectly reasonable requirement. Dispositionalism is a form of what, for lack of a better term, we can call "mentalism". According to mentalism, a word has the meaning it has because of

---

[29] What follows improves on the argument of Guardo 2012. For somewhat analogous remarks see also Boghossian 2015.

[30] In case you are wondering, yes, the argument I am about to sketch is consistent both with the view that dispositions are identical with their causal bases and with the view that they are not (see Lewis 1997, pp. 151-152 for some background).

[31] As I see it, non-actuality-grounded dispositions are not real dispositions. That being said, nothing in my argument requires you to agree with me on this.

[32] I am assuming that the information about past usage comprises also information about the environment in which we made our utterances – so no, I am not committing myself to any kind of naïve internalism.

what people usually mean by it – a mentalist thinks that the phenomenon of *word* (and sentence) meaning can be explained in terms of a particular *mental state*, namely the mental state of meaning, or intending, something by a sign. To this, dispositionalism only adds the thesis that this mental state should be analyzed in terms of dispositions. But the mental state of meaning something by a sign, if there is such a thing, is definitely actuality-grounded; saying that I mean addition by "+" is saying that I *currently have*, in the *actual* world, a certain property – not just that I *could* acquire it. Therefore, if the mental state of meaning something by a sign can really be analyzed in terms of dispositions, the dispositions in question must be actuality-grounded dispositions.

Having explained both the meaning of (2) and the rationale behind it, I can now make clear why I am skeptical one can find a set of dispositions in ideal conditions which meets both (1) and (2), i.e. a set of dispositions in ideal conditions whose members are both addition-tracking and actuality-grounded. To this end, I will discuss two, prima facie rather promising, proposals – each one corresponding to a very natural way to flesh out the notion of ideal conditions.

The first proposal is to define the ideal conditions in terms of the following two clauses: (a) I live long enough to perform the relevant calculations and (b) my brain was stuffed with all the extra matter needed to grasp the numbers in question and perform additions involving them. For brevity's sake, I will refer to the relevant dispositions in ideal conditions as "dispositions in ideal conditions$_1$".

Are dispositions in ideal conditions$_1$ addition-tracking? I am willing to grant that they are and, therefore, that they meet condition (1).[33] However, a little reflection is enough to see that dispositions in ideal conditions$_1$ are not actuality-grounded and, therefore, fail to meet condition (2). After all, the property which would be causally responsible for my giving the arithmetically correct answer if asked about the sum of two authentically huge numbers – i.e. the property of having a brain stuffed with all the extra matter needed to grasp such numbers and perform additions involving them – is quite clearly not one I currently have; at most, it is one I could acquire.

Faced with the failure of this first proposal, one may suggest substituting (b) with something along the lines of (b*): I have a lifelong supply of paper and pencils. For the sake of brevity, I will refer to this kind of disposition in ideal conditions as "dispositions in ideal conditions$_2$".

Since, unlike (b), (b*) does not make any mention of enhanced brains, there is no reason to suspect that dispositions in ideal conditions$_2$ are not actuality-

---

[33] Even though it is worth noting that this is not uncontroversial. Kripke (1981, pp. 27-28) argued that we should be agnostic about the truth value of the relevant conditional and if one has the strong intuition that it is true, it is because they read it in a way that makes the dispositional analysis circular.

grounded; therefore, we can assume that condition (2) is met. Unfortunately, however, it seems clear that there will be speakers who – for at least some pairs X, Y – are disposed to give some arithmetically *incorrect* answer if asked about "X + Y" in ideal conditions$_2$. The reference to indefinite lifespans and lifelong supplies of paper and pencils can take care of cardinality-related issues, but it is just not designed to deal with incorrigible tendencies to make *mistakes*. Therefore, this time it is condition (1) which is not met.

These are, of course, only two examples, but they should give the reader an idea of the difficulties one encounters while trying to make the notion of a disposition in ideal conditions precise enough. I now turn to the second reason for being wary of ideal-condition answers to the Privilege Argument.

Let us say, for argument's sake, that the above difficulties can be overcome, and let us call the dispositions in ideal conditions that we are assuming satisfy requirements (1)-(2) "dispositions in ideal conditions$_3$"; the question now is: why are dispositions in ideal conditions$_3$ privileged over dispositions in ideal conditions$_1$ and dispositions in ideal conditions$_2$? The only answer to this question I can think of is that dispositions in ideal conditions$_3$ are privileged over dispositions in ideal conditions$_1$ because they are actuality-grounded and over dispositions in ideal conditions$_2$ because they are addition-tracking. Is this a satisfactory answer? Well, the first part is, I think, fine. As noted above, the mental state of meaning something by a sign, if there is such a thing, is actuality-grounded; therefore, privileging actuality-grounded dispositions over non-actuality-grounded ones makes sense. However, saying that our dispositions in ideal conditions$_3$ are the meaning-constituting ones because they are addition-tracking is blatantly question-begging. What is so special about the addition function? The only possible answer seems to be that it is to that function that "+" refers. But this will not do. "+" refers to addition only if dispositions in ideal conditions$_3$ are the meaning-constituting ones. Therefore, saying that dispositions in ideal conditions$_3$ are the meaning-constituting ones because they are addition-tracking and it is to addition that "+" refers is viciously circular.[34]

---

[34] This is not the first time we see an answer to Privilege turn out to be circular. As a very helpful reviewer pointed out to me, this may be taken to provide evidence for some kind of semantic primitivism, along the lines of Boghossian's (1989), Ginsborg's (2011), Kearns and Magidor's (2012), or Verheggen's (2015) – all instances of what Kusch (2006) calls "meaning-determinist primitivism". After all, a mark of authentically primitive phenomena is their tendency to keep popping up, unanalyzed, in the midst of our attempts to explain them reductively, thereby making the reduction fail (to be clear, I do not find primitivism to provide an especially plausible picture of the nature of meaning, but I

The second problem with the notion that the key to meeting Privilege's challenge lies in the concept of ideal conditions is, therefore, that – by taking this route – we just move the difficulty to another level. When it first arises, the challenge is to justify the privileging of certain dispositions over all the others, and then it mutates into the problem of justifying the privileging of certain dispositions *in ideal conditions* over all our other dispositions *in ideal conditions*.

Here, of course, I am assuming that there is more than one legitimate way to flesh out the notion of ideal conditions. If only dispositions in ideal conditions$_3$ were really dispositions in *ideal* conditions – if neither the conjunction of (a) and (b) nor that of (a) and (b*) were legitimate precisifications of the notion of ideal conditions – then one could just say (bracketing for the time being my third reason to be wary of the ideal-condition strategy) that dispositions in ideal conditions$_3$ are privileged over dispositions in ideal conditions$_1$ and dispositions in ideal conditions$_2$ simply because they are the only ones which are really dispositions in *ideal* conditions. That being said, the assumption in question seems to me to be eminently plausible,[35] which means that this second reason for being wary of ideal-condition answers to Privilege must be taken seriously.

The third reason to be skeptical of the potential of the ideal-condition strategy I just mention, but I hope that this will not give the reader the impression that I take it to be less important than the previous ones – it is just that I really do not know what a supporter of the ideal-condition strategy could say to address this worry, and without some idea of what one could say to answer the problem I have in mind there just is not that much to discuss. Anyway, this third problem is just that, even though I am willing to grant that the idea has some intuitive content, that ideal conditions can provide the normative oomph the dispositionalist is after – and, more in general, that they suffice to set apart certain dispositions as the reference-determining ones – is far from clear. Dispositionalists need to back up their

---

do agree with the reviewer's point about the evidential value of the fact that several somewhat natural strategies to try to answer Privilege are ultimately question-begging).

[35] For one, denying it commits you to offering some account of the truth conditions of the – by the way, rather convoluted (see, e.g., Lewis 1997 and Hauska 2009) – counterfactuals related to disposition ascriptions such as "Saul is disposed to answer "125" if asked for "68 + 57" in ideal conditions" that does not make any use either of the notion of ideal conditions or of semantic notions (keep in mind the no-circularity requirement of section 1!), which is no easy task (see Lance and O'Leary-Hawthorne 1997, pp. 317-322; Lance and O'Leary-Hawthorne focus on conditionals such as "If we were to answer questions involving huge numerals, we would give response "Z"", but their discussion applies, mutatis mutandis, also to the conditionals we are interested in here).

intuitions with some kind of argument, and I myself have no idea how such an argument could go.

3.3. Standard

Some may think that the most promising way to try to rebut the Privilege Argument might be to focus on *standard* conditions, i.e. on the conditions in which, as a matter of fact, we usually produce our utterances. We must, however, be careful here. As long as we focus on simple and complex dispositions, it is just false that, for any pair of numbers X and Y, I am disposed to answer "Z" if queried about "X + Y" in standard conditions (or "disposed$_s$ to answer "Z" if queried about "X + Y"", for short). With regard to authentically huge numbers, the natural thing to say is that there is no answer I am disposed$_s$ to give. And there will also be cases, involving very large even though not exactly huge numbers, in which I am disposed$_s$ to give an incorrect answer – or a variety of incorrect answers. Both the Cardinality Argument and the Mistake Argument would still be with us.

In order to address this problem, we can make use of Warren's apparatus and say that the privileged, meaning-determining dispositions are the speakers' *general* and *composite* dispositions to *stably* give a certain response when queried about a given problem in *standard* conditions – which, as a matter of fact, are normal in Warren's sense. More concisely: the meaning-determining dispositions are the speakers' dispositions$_{gcss}$.

Is this enough to rebut Privilege? There are two problems here. The first one is that the notion of standard conditions cannot really provide the required normative oomph. Here, just as in the case of the analogous worry I raised while discussing the ideal-condition strategy, I do not have much to say besides the rather obvious remark that the dispositionalist owes us an argument here – and I myself would not know how to argue for such a point. I will therefore focus on the second worry. This second worry arises the moment one realizes that if we identify the meaning-determining dispositions not just with the speakers' dispositions$_s$, but with their dispositions$_{gcss}$, then it is not sufficient to show that *the standard conditions* should be privileged over *the non-standard ones*: what we have to show is that *dispositions$_{gcss}$* should be privileged over *all the other dispositions speakers possess* – some of which will themselves be dispositions$_s$. And this is no easy task.

The point is that a composite disposition – and, a fortiori, a disposition$_{gcss}$ – *is not really a disposition*. According to Warren's definition, S has a composite disposition to perform A if and only if A would be the output of the manifestation of certain linked dispositions of S's. Therefore, saying that I have a certain composite disposition is just short for saying that I have the relevant linked dispositions; saying that I have the composite disposition to correctly compute

23

"714,231,723,055,439 + 231,135,230,349,825" is just short for saying that I am disposed to correctly compute "9 + 5", "3 + 2 + 1", "4 + 8", etc.

Here it is important to keep in mind the difference between composite and complex dispositions. Saying that I have the composite disposition to correctly compute "714,231,723,055,439 + 231,135,230,349,825" is *not* saying that if asked for "714,231,723,055,439 + 231,135,230,349,825" I would answer "945,366,953,405,264", just not directly, but as a result of the execution of the relevant intermediate steps (i.e. adding 5 to 9 etc). That would be a *complex* disposition. Saying that I have the composite disposition to correctly compute "714,231,723,055,439 + 231,135,230,349,825" is really just saying that I have a bunch of much simpler dispositions, which *could* compose the relevant complex disposition, but – as a matter of fact – *do not*. The difference between having a complex and a composite disposition is like that between buying a house and buying the bricks with which to build it, which I think we can agree is a pretty radical difference.

If we keep this in mind, it is not clear why we should believe that the correct response to "X + Y" is determined by my composite "disposition" to answer "Z" (which is not even a real disposition!) rather than by my complex disposition to give some other answer, inconsistent with the addition function. After all, it seems fairly clear that, ceteris paribus, a complex disposition should be given much more weight than a composite one – whose entire psychological reality consists in the fact that, scattered throughout our dispositional landscape, lie the building blocks from which it could be built.[36]

The standard-condition strategy does not, therefore, fare any better than the other strategies I discussed. These strategies do not, of course, exhaust the dispositionalist's options, and both the ideal-condition and the standard-condition strategy can no doubt be developed in other ways, even substantially different from the ones I focused on. However, the arguments I put forward in this section are fairly easy to generalize to other cases – such as that of evolutionary versions of the dispositional approach, in the vein of Millikan (e.g. 1989) – and so I am skeptical that semantic dispositionalism has the resources to satisfactorily answer the Privilege Argument. That being said, I would love to see semantic

---

[36] One might be tempted to answer this argument by saying that, in the case we are interested in, the ceteris paribus clause is never satisfied, i.e. that non-addition-tracking dispositions$_s$ never differ from dispositions$_{gcss}$ only by being non-composite. That is quite clearly false, though; for one, it is pretty clear that our non-composite but general dispositions to stably give a certain response when asked about a certain problem in standard conditions do not track addition – for they cover only a finite segment of that function.

dispositionalists at least *try* to address this problem, instead of just ignoring it – which, so far, seems to have been their weapon of choice.

References

Båve, Arvid (2020). Semantic Dispositionalism without Exceptions. *Philosophical Studies* 177 (6), 1751-1771.

Bird, Alexander; Handfield, Toby (2008). Dispositions, Rules, and Finks. *Philosophical Studies* 140 (2), 285-298.

Blackburn, Simon (1984). The Individual Strikes Back. *Synthese* 58 (3), 281-301.

Boghossian, Paul A. (1989). The Rule-Following Considerations. *Mind* 98 (392), 507-549.

Boghossian, Paul A. (2003). The Normativity of Content. *Philosophical Issues* 13 (1), 31-45.

Boghossian, Paul A. (2005). Is Meaning Normative?. In Christian Nimtz, Ansgar Beckermann (eds.), *Philosophie und/als Wissenschaft – Hauptvorträge und Kolloquiumsbeiträge zu GAP.5*. Paderborn: Mentis, 205-218.

Boghossian, Paul A. (2015). Is (Determinate) Meaning a Naturalistic Phenomenon?. In Steven Gross, Nicholas Tebben, Michael Williams (eds.), *Meaning without Representation – Essays on Truth, Expression, Normativity, and Naturalism*. New York-Oxford: Oxford University Press, 331-358.

Chalmers, David J. (1996). *The Conscious Mind – In Search of a Fundamental Theory*. New York-Oxford: Oxford University Press.

Cheng, Kai-Yuan (2009). Semantic Dispositionalism, Idealization, and *Ceteris Paribus* Clauses. *Minds & Machines* 19 (3), 407-419.

Cheng, Kai-Yuan (2010). Intrinsic Finks and Attributions of Rule-Following Dispositions. *Grazer Philosophische Studien* 80 (1), 209-220.

Gibbard, Allan (2012). *Meaning and Normativity*. New York-Oxford: Oxford University Press.

Ginsborg, Hannah (2011). Primitive Normativity and Skepticism about Rules. *Journal of Philosophy* 108 (5), 227-254.

Glüer, Kathrin; Wikforss, Åsa Maria (2009). Against Content Normativity. *Mind* 118 (469), 31-70.

Guardo, Andrea (2012). Rule-Following, Ideal Conditions and Finkish Dispositions. *Philosophical Studies* 157 (2), 195-209.

Guardo, Andrea (2014). Semantic Dispositionalism and Non-Inferential Knowledge. *Philosophia* 42 (3), 749-759.

Guardo, Andrea (2018). *Seguire una regola*. Milano-Udine: Mimesis.

Hattiangadi, Anandi (2006). Is Meaning Normative?. *Mind & Language* 21 (2), 220-240.

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

Hauska, Jan (2009). Dispositions Unmasked. *Theoria* 75 (4), 304-335.

Heil, John; Martin, C. B. (1998). Rules and Powers. *Philosophical Perspectives* 12 (1), 283-312.

Johnston, Mark (1992). How to Speak of the Colors. *Philosophical Studies* 68 (3), 221-263.

Jones, Nicholaos J.; Podlaskowski, Adam C. (2012). Idealizing, Abstracting, and Semantic Dispositionalism. *European Journal of Philosophy* 20 (1), 166-178.

Kaplan, David (1989). Demonstratives – An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In Joseph Almog, John Perry, Howard Wettstein (eds.), *Themes from Kaplan*. New York-Oxford: Oxford University Press, 481-563.

Kearns, Stephen; Magidor, Ofra (2012). Semantic Sovereignty. *Philosophy and Phenomenological Research* 85 (2), 322-350.

Kowalenko, Robert (2009). How (not) to Think about Idealisation and *Ceteris Paribus* Laws. *Synthese* 167 (1), 183-201.

Kripke, Saul (1981) [1982]. *Wittgenstein on Rules and Private Language – An Elementary Exposition*. Oxford: Blackwell.

Kusch, Martin (2005). Fodor v. Kripke: Semantic Dispositionalism, Idealization and Ceteris Paribus Clauses. *Analysis* 65 (2), 156-164.

Kusch, Martin (2006). *A Sceptical Guide to Meaning and Rules – Defending Kripke's Wittgenstein*. Montreal-Kingston: McGill-Queen's University Press.

Lance, Mark Norris; O'Leary-Hawthorne, John (1997). *The Grammar of Meaning – Normativity and Semantic Discourse*. Cambridge: Cambridge University Press.

Lewis, David (1986). Causal Explanation. In David Lewis, *Philosophical Papers – Volume II*. New York-Oxford: Oxford University Press, 214-240.

Lewis, David (1997). Finkish Dispositions. *Philosophical Quarterly* 47 (187), 143-158.

Machamer, Peter; Darden, Lindley; Craver, Carl F. (2000). Thinking about Mechanisms. *Philosophy of Science* 67 (1), 1-25.

Martin, C. B. (1994). Dispositions and Conditionals. *Philosophical Quarterly* 44 (174), 1-8.

Miller, Alexander (2010). The Argument from Queerness and the Normativity of Meaning. In Martin Grajner, Adolf Rami (eds.), *Truth, Existence and Realism*. Paderborn: Mentis, 107-124.

Millikan, Ruth Garrett (1989). Biosemantics. *Journal of Philosophy* 86 (6), 281-297.

Podlaskowski, Adam C. (2012). Simple Tasks, Abstractions, and Semantic Dispositionalism. *Dialectica* 66 (4), 453-470.

Schacter, Daniel L.; Wagner, Anthony D. (2013). Learning and Memory. In Eric R. Kandel, James H. Schwartz, Thomas M. Jessell, Steven A. Siegelbaum, A.

J. Hudspeth (eds.), *Principles of Neural Science – Fifth Edition*. New York: McGraw-Hill, 1441-1460.

Schlosser, Markus E. (2011). The Metaphysics of Rule-Following. *Philosophical Studies* 155 (3), 345-369.

Shogenji, Tomoji (1993). Modest Scepticism about Rule-Following. *Australasian Journal of Philosophy* 71 (4), 486-500.

Skyrms, Brian (1996) [2014]. *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Skyrms, Brian (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.

Skyrms, Brian (2010). *Signals – Evolution, Learning & Information*. New York-Oxford: Oxford University Press.

Sultanescu, Olivia; Verheggen, Claudine (2019). Davidson's Answer to Kripke's Sceptic. *Journal for the History of Analytical Philosophy* 7 (2), 7-28.

Verheggen, Claudine (2015). Towards a New Kind of Semantic Normativity. *International Journal of Philosophical Studies* 23 (3), 410-424.

Warren, Jared (2020). Killing Kripkenstein's Monster. *Noûs* 54 (2), 257-289.

Wasserman, Ryan (2018). *Paradoxes of Time Travel*. New York-Oxford: Oxford University Press.

Whiting, Daniel (2007). The Normativity of Meaning Defended. *Analysis* 67 (2), 133-140.

Whiting, Daniel (2009). Is Meaning Fraught with Ought?. *Pacific Philosophical Quarterly* 90 (4), 535-555.

Wikforss, Åsa Maria (2001). Semantic Normativity. *Philosophical Studies* 102 (2), 203-226.

Wittgenstein, Ludwig (1921) [1974]. *Logisch-philosophische Abhandlung*. English translation *Tractatus logico-philosophicus*. London-New York: Routledge.

Wright, Larry (1973). Functions. *Philosophical Review* 82 (2), 139-168.

Zalabardo, José (1997). Kripke's Normativity Argument. *Canadian Journal of Philosophy* 27 (4), 467-488.